NATIONAL QUALITY FORUM

Measure Evaluation 4.1
January 2010

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The sub-criteria and most of the footnotes from the evaluation criteria are provided in Word comments and will appear if your cursor is over the highlighted area (or in the margin if your Word program is set to show revisions in balloons). Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each sub-criterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the* sub-criteria *(yellow highlighted areas)*.

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the sub-criterion, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few sub-criteria as indicated)

(for NQF staff use) **NQF Review** #: OT2-017-09 **NQF Project:** Patient Outcomes Measure Submissions

MEASURE DESCRIPTIVE INFORMATION

De.1 Measure Title: Functional Assessment of Cancer Therapy - Breast (FACT-B)

De.2 Brief description of measure*:* Adult Breast Cancer patients

1.1-2 Type of Measure: outcome

De.3 If included in a composite or paired with another measure, please identify composite or paired measure

The FACIT Measurement System is a collection of QOL questionnaires targeted to the management of chronic illness. "FACIT" (Functional Assessment of Chronic Illness Therapy) was adopted as the formal name of the measurement system in 1997 to portray the expansion of the more familiar "FACT" (Functional Assessment of Cancer Therapy) series of questionnaires into other chronic illnesses and conditions. Thus, FACIT is a broader, more encompassing term that includes the FACT questionnaires under its umbrella.

The measurement system, under development since 1987, began with the creation of a generic CORE questionnaire called the Functional Assessment of Cancer Therapy-General (FACT-G). The FACT-G (now in Version 4) is a 27-item compilation of general questions divided into four primary QOL domains: Physical Well-Being, Social/Family Well-Being, Emotional Well-Being, and Functional Well-Being. It is considered appropriate for use with patients with any form of cancer, and has also been used and validated in other chronic illness conditions (e.g., HIV/AIDS and multiple sclerosis) and in the general population (using a slightly modified version).

Validation of a core measure allowed for the evolution of multiple disease, treatment, condition, and non-cancer-specific subscales. FACIT subscales are constructed to complement the FACT-G, addressing relevant disease-, treatment-, or condition-related issues not already covered in the general questionnaire. Each is intended to be as specific as necessary to capture the clinically-relevant problems associated with a given condition or symptom, yet general enough to allow for comparison across diseases, and extension, as appropriate, to other chronic medical conditions.
In the case of the FACT-B, it is comprised of the aforementioned FACT-G plus the 9-item BCS (Breast Cancer Subscale). Combined, the questionnaire is called the FACT-B. All results presented in this submission are for the FACT-B, the FACT-G plus the Breast Cancer Subscale.

De.4 **National Priority Partners Priority Area:** population health, Palliative and End of Life care
De.5 **IOM Quality Domain**: patient-centered
De.6 **Consumer Care Need:** Living With Illness

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** proprietary measure A.3 **Measure Steward Agreement:** agreement signed and submitted A.4 **Measure Steward Agreement attached:** FACIT.org_StewardAgreement.pdf	**A** **Y**☐ **N**☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** **Y**☐ **N**☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** public reporting, quality improvement The Functional Assessment of Cancer Therapy - Breast (FACT-B) is a Health-Related Quality of Life (HRQOL) measure used for Breast Cancer Patients	**C** **Y**☐ **N**☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** **Y**☐ **N**☐
(for NQF staff use**)** **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** **Y**☐ **N**☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	Eval Rating
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** affects large numbers, a leading cause of	**1a**

morbidity/mortality, severity of illness, patient/societal consequences of poor quality, frequently performed procedure, high resource use

1a.2

1a.3 Summary of Evidence of High Impact: FACT-B is a widely reported disease in medical literature. It impacts a large number of people, has a high impact on patient and family Quality of Life, and has significant impact on emotional well-being. Treatment of breast cancer consumes significant financial resources. Given recent successes in the treatment of breast cancer, measuring quality of life and the quality of life during survival could be as important as the length of survival.

1a.4 Citations for Evidence of High Impact:

Brady, M.J., Cella, D.F., Mo, F., Bonomi, A.E., Tulsky, D.S., Lloyd, S.R., Deasy, S., Cobleigh, M., & Shiomoto, G. (1997). Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. Journal of Clinical Oncology, 15(3), 974-986.

Fetting, J.H., Gray, R., Fairclough, D.L., Smith, T.J., Margolin, K.A., Citron, M.L., Grove-Conrad, M., Cella, D., Pandya, K., Robert, N., & Abeloff, M.D. (1998). Sixteen-week multidrug regimen vs. Cyclophosphamide, Doxorubicin, and Fluorouracil (CAF) as adjuvant therapy for node positive, receptor negative breast cancer: An intergroup study. Journal of Clinical Oncology, 16(7), 2382-2391.

Marcus, A.C., Garrett, K.M., Cella, D., Wenzel, L.B., Brady, M.J., Crane, L.A., McClatchey, M.W., Kluhsman, B.C., & Pate-Willig, M. (1998). Telephone counseling of breast cancer patients after treatment: A description of a randomized clinical trial. Psycho-Oncology, 7, 470-482.

Audrain, J., Rimer, B., Cella, D., Stefanek, M., Garber, J., Pennanen, M, Helzlsouer, K., Vogel, V., Lin, T.H., & Lerman, C. (1999). The impact of a brief coping skills intervention on adherence to breast self-examination among first-degree relatives of newly diagnosed breast cancer patients. Psycho-Oncology, 8, 220-229.

Wenzel, L., Fairclough, D., Brady, M., Cella, D., Garrett, K., Kluhsman, B. & Marcus, A. (1999). Age-related differences in the quality of life of breast carcinoma patients after treatment. Cancer, 86(9), 1768-1774.

Fairclough, D.L., Fetting, J.H., Cella, D., Wonson, W., Grove-Conrad, M., & Moinpour, C. (2000). Quality of life and quality-adjusted survival for breast cancer patients receiving adjuvant therapy. Quality of Life Research, 8, 723-731.

Arora, N.K., Gustafson, D.H., Hawkins, R.P., McTavish, F., Mendenhall, J., Pingree, S., Cella, D.F. (2001). Impact of surgery and chemotherapy on the Quality of life of younger women with breast Carcinoma: A prospective study. Cancer, 92(5): 1288-1298.

Gustafson, D.H., Hawkins, R., Pingree, S., McTavish, F., Arora, N.K., Mendenhall, J., Cella, D.F., Serlin, R.C., Apantaku, F.M., Stewart, J., Salner, A. (2001). Effect of computer support on younger women with breast cancer. Journal of General Internal Medicine. 16(7), 435-445.

Hahn, E.A., Holzner, B., Kemmler, G., Sperner-Unterweger, B., Hudgens, S., Cella, D. (2005) Cross-cultural evaluation of health status using item response theory: FACT-B comparisons between Austrian and U.S. breast cancer patients. Evaluation & the Health Professions, 28 (2), 233-259.

Yost, K., Yount, S., Eton, D., Silberman, C., Broughton-Heyes; Cella, D. (2005) Validation of the Functional Assessment of Cancer Therapy-Breast Symptom Index (FBSI). Breast Cancer Research and Treatment, 90, 295-298.

Cella, D., Fallowfield, L., Barker, P., Cuzick, J., Locker, G., Howell, A. (2006) Quality of life of postmenopausal women in the ATAC trial after completion of 5 year's adjuvant treatment for early breast cancer. Breast Cancer Research and Treatment, 100 (3), 273-284.

C☐
P☐
M☐
N☐

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure:

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
n/a

1b.3 Citations for data on performance gap:
n/a

1b.4 Summary of Data on disparities by population group:
The disparities in cancer care by population group is widely published. In fact the NIH and AHRQ have whole funding initiatives specifically identified to address these issues. Dr. Cella's group has considerable expertise in researching and measuring response differences across groups, including publications on literacy and cross cultural assessment issues.

1b.5 Citations for data on Disparities:
ISOQOL Article of the Year 2007, Hahn, E., et al The impact of literacy on health-related quality of life measurement and outcomes in cancer outpatients. Quality of Life Research, 16(3), 495-507.
Hahn E, Cella D. Health outcomes assessment in vulnerable populations: measurement challenges and recommendations. Archives of Physical Medicine and Rehabilitation 2003; 84(Suppl 2):S35-S42.

1b
C☐
P☐
M☐
N☐

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** This questionnaire was specifically developed and validated for breast cancer patients. As described previously and in the literature citations, it provides a multidisclinary measure of a patient's well being including emotional, physical. It is a measure that's responsive to change over time. It can also be used to measure response to treatment (it measures change to clinical status). It can be used to demonstrate that a change in treatment plan is warranted, and it can be used to demonstrate the effectiveness of palliative care.

Item content was determined by combined expert and breast cancer patient input, ensuring that clinically important issues relevant to patients were included. Content validity has been ensured by use of a rigorous, peer-reviewed procedure for determining the relevance and relative importance of each of the many issues raised by breast cancer patients as having a bearing upon their HRQOL. There are over 25 published reports detailing its performance. Thus, there is a solid reference literature to which one can compare results. Finally, there is a growing body of research that illustrates clinically significant differences and changes in scores in the FACT-B scale, aiding in study sample size determination and interpretation results.

1c.2-3. Type of Evidence: cohort study, evidence based guideline, expert opinion, observational study, randomized controlled trial, systematic synthesis of research

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
Since the FACT-B was developed, it has been widely used by clinicians and the clinical trial industry to assess HRQOL for breast cancer patients.

1c.6 Method for rating evidence: See 1 c 4 response

1c
C☐
P☐
M☐
N☐

1c.7 Summary of Controversy/Contradictory Evidence: n/a

1c.8 Citations for Evidence (*other than guidelines***):** Brady, M.J., Cella, D.F., Mo, F., Bonomi, A.E., Tulsky, D.S., Lloyd, S.R., Deasy, S., Cobleigh, M., & Shiomoto, G. (1997). Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. Journal of Clinical Oncology, 15(3), 974-986.

Fetting, J.H., Gray, R., Fairclough, D.L., Smith, T.J., Margolin, K.A., Citron, M.L., Grove-Conrad, M., Cella, D., Pandya, K., Robert, N., & Abeloff, M.D. (1998). Sixteen-week multidrug regimen vs. Cyclophosphamide, Doxorubicin, and Fluorouracil (CAF) as adjuvant therapy for node positive, receptor negative breast cancer: An intergroup study. Journal of Clinical Oncology, 16(7), 2382-2391.

Marcus, A.C., Garrett, K.M., Cella, D., Wenzel, L.B., Brady, M.J., Crane, L.A., McClatchey, M.W., Kluhsman, B.C., & Pate-Willig, M. (1998). Telephone counseling of breast cancer patients after treatment: A description of a randomized clinical trial. Psycho-Oncology, 7, 470-482.

Audrain, J., Rimer, B., Cella, D., Stefanek, M., Garber, J., Pennanen, M, Helzlsouer, K., Vogel, V., Lin, T.H., & Lerman, C. (1999). The impact of a brief coping skills intervention on adherence to breast self-examination among first-degree relatives of newly diagnosed breast cancer patients. Psycho-Oncology, 8, 220-229.

Wenzel, L., Fairclough, D., Brady, M., Cella, D., Garrett, K., Kluhsman, B. & Marcus, A. (1999). Age-related differences in the quality of life of breast carcinoma patients after treatment. Cancer, 86(9), 1768-1774.

Fairclough, D.L., Fetting, J.H., Cella, D., Wonson, W., Grove-Conrad, M., & Moinpour, C. (2000). Quality of life and quality-adjusted survival for breast cancer patients receiving adjuvant therapy. Quality of Life Research, 8, 723-731.

Arora, N.K., Gustafson, D.H., Hawkins, R.P., McTavish, F., Mendenhall, J., Pingree, S., Cella, D.F. (2001). Impact of surgery and chemotherapy on the Quality of life of younger women with breast Carcinoma: A prospective study. Cancer, 92(5): 1288-1298.

Gustafson, D.H., Hawkins, R., Pingree, S., McTavish, F., Arora, N.K., Mendenhall, J., Cella, D.F., Serlin, R.C., Apantaku, F.M., Stewart, J., Salner, A. (2001). Effect of computer support on younger women with breast cancer. Journal of General Internal Medicine. 16(7), 435-445.

Hahn, E.A., Holzner, B., Kemmler, G., Sperner-Unterweger, B., Hudgens, S., Cella, D. (2005) Cross-cultural evaluation of health status using item response theory: FACT-B comparisons between Austrian and U.S. breast cancer patients. Evaluation & the Health Professions, 28 (2), 233-259.

Yost, K., Yount, S., Eton, D., Silberman, C., Broughton-Heyes; Cella, D. (2005) Validation of the Functional Assessment of Cancer Therapy-Breast Symptom Index (FBSI). Breast Cancer Research and Treatment, 90, 295-298.

Cella, D., Fallowfield, L., Barker, P., Cuzick, J., Locker, G., Howell, A. (2006) Quality of life of postmenopausal women in the ATAC trial after completion of 5 year's adjuvant treatment for early breast cancer. Breast Cancer Research and Treatment, 100 (3), 273-284.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):** n/a

1c.10 Clinical Practice Guideline Citation: n/a

1c.11 National Guideline Clearinghouse or other URL: n/a **1c.12 Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** n/a **1c.13 Method for rating strength of recommendation** (*If different from* USPSTF system*, also describe rating and how it relates to USPSTF)***:** n/a **1c.14 Rationale for using this guideline over others:** This questionnaire and its shortened version the Functional Breast Symptom Index (FBSI) is widely used by clinicians, clinical trialists and cooperative group trials. It has been well validated and is in widespread use for assessing breast cancer patients QOL. It is also widely accepted for use in clinical decision-making. It is available in over 45 languages.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report***, met?** **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome***):** Individual items ask breast cancer patients about how true certain symptoms have been for them. The composite score of all the items gives a Quality of Life (QOL) score which can be used by clinicians and in clinical trials to determine certain clinical indicators. All FACIT scales are scored so that a high score is good. As each of the items on the FACT-G and the BCS ranges from 0-4, the range of possible scores is 0-144, with 0 being the worst possible score and 144 the best. To obtain the 0-144 score each negatively-worded item response is recoded so that 0 is a bad response and 4 is good response. All responses are added with equal weight to obtain the total score. In cases where some answers may be missing, a total score is prorated from the score of the answered items, so long as more than 50% of the items were answered. Computer programs written in SPSS and SAS for the FACT-B Scale are available. **2a.2 Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** Respondents are requested to look back on the previous 7 days. **2a.3 Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** The FACT-B scores are individual respondent scores. Responses are "Not at All", "A Little Bit", "Somewhat", "Quite a Bit" "Very Much". Each item is scored as being either a positive or negative item, depending on if the response would be positive or negative to the patient's quality of life. (ie "I have a lack of energy" Response:"Very Much" is considered a negatively scored item vs. "My thinking is clear" Response: "Very Much" is considered a positively worded item.	**2a-specs** **C☐** **P☐** **M☐** **N☐**

Here is the scoring algorithm for the FACT-B:

FACT-B Scoring Guidelines (Version 4) – Page 1

Instructions:* 1. Record answers in "item response" column. If missing, mark with an X
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. Add subscale scores to derive total scores (TOI, FACT-G & FACT-B).
5. The higher the score, the better the QOL.

Subscale		Item Code	Reverse item?	Item response	Item Score
PHYSICAL WELL BEING (PWB)	GP1	4	-	_____	=_____
	GP2	4	-	_____	=_____
	GP3	4	-	_____	=_____
	GP4	4	-	_____	=_____
	GP5	4	-	_____	=_____
	GP6	4	-	_____	=_____
	GP7	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=PWB subscale score

SOCIAL/FAMILY WELL BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
	GS4	0	+	_____	=_____
	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=SWB subscale score

EMOTIONAL WELL BEING (EWB)	GE1	4	-	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	-	_____	=_____
	GE4	4	-	_____	=_____
	GE5	4	-	_____	=_____
	GE6	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 6: _____
Divide by number of items answered: _____=EWB subscale score

FUNCTIONAL WELL-BEING (FWB)	GF1	0	+	_____	=_____
	GF2	0	+	_____	=_____
	GF3	0	+	_____	=_____
	GF4	0	+	_____	=_____
	GF5	0	+	_____	=_____
	GF6	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=FWB subscale score

FACT-B Scoring Guidelines (Version 4) – Page 2

Subscale	Item Code	Reverse item?	Item response	Item Score
BREAST	B1	4 -	_____	=_____
CANCER	B2	4 -	_____	=_____
SUBSCALE	B3	4 -	_____	=_____
(BCS)	B4	0 +	_____	=_____
B5	4 -	_____	=_____	
B6	4 -	_____	=_____	
B7	4 -	_____	=_____	
B8	4 -	_____	=_____	
B9	0 +	_____	=_____	
P2	NOT CURRENTLY SCORED			

Sum individual item scores:_____
Multiply by 9: _____
Divide by number of items answered: _____=BC Subscale score

To derive a FACT-B Trial Outcome Index (TOI):

_____ + _____ + _____ =_____=FACT-B TOI
(PWB score) (FWB score) (BCS score)

To Derive a FACT-G total score:

_____ + _____ + _____ + _____=_____=FACT-G Total score
(PWB score) (SWB score) (EWB score) (FWB score)

To Derive a FACT-B total score:

_____ + _____ + _____ + _____ + _____ =_____=FACT-B Total score
(PWB score) (SWB score) (EWB score) (FWB score) (BCS score)

*For guidelines on handling missing data and scoring options, please refer to the Administration and Scoring Guidelines in the manual or on-line at www.facit.org.

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
n/a

2a.5 Target population gender: Female, Male
2a.6 Target population age range: Adults with Breast Cancer

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator***):**
n/a

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
n/a

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population***):** n/a

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**
n/a

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
n/a

2a.12-13 Risk Adjustment Type: no risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method***):**
n/a

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Other (specify) See detailed scoring in 2.a.3
2a.20 Interpretation of Score: better quality = higher score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**
See 2.a.3.

2a.22 Describe the method for discriminating performance *(e.g., significance testing)***:**
MID's for FACT-B

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
The sample size for the study in question is dependent on the how the scale will be used. It can be used with single patients for clinical decision-making. Or it can be used for clinical trial QOL scores to be tied to a clinical response. If IRT (item response theory)/Rasch analyses will be used, that will also impact the sample size.

The sample can be any individual or group of patients being treated, or having previously been treated for breast cancer.

The questionnaire can be administered by RN's or research personnel directly instructing the participants, or it can be administered electronically online or via telephone CATI (computer adaptive telephone interview). Each assessment method will impact the sample in terms of accessibility.

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested)*
Survey: Patient

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
The FACT-B questionnaire is currently being used by investigators from medical and educational institutions, industry sponsors, and cooperative clinical trial groups. Application includes use in Phase I, II, and III, clinical trials, in health-practice, for symptom management, for psychological intervention, and in other disease- or symptom- treatment evaluations. The FACT-B is most commonly used in the clinical trial and treatement settings, but has also been used in screening, survivorship and end-of-life evaluations.

2a.26-28 Data source/data collection instrument reference web page URL or attachment: Attachment FACT-B Validation.pdf

2a.29-31 Data dictionary/code table web page URL or attachment: Attachment FACT-B_ENG_Final_Ver4_16Nov07.doc

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested)*
Can be measured at all levels

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested)*
all settings

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)***:** The data used in the original validation report were derived from two samples. The first sample was tested twice over a 2-month periodand was therefore used for the sensitivity to change analysis. Because this sample also completed a QL instrument, the Functional Living Index-Cancer (FLIC), 2 ° data from this sample were used to assess concurrent validity as well. This first sample consisted of 47 breast cancer patients who completed version 1 of the FACT-B as part of a funded psychosocial QL intervention project targeting patients
with advanced (stage III or IV) breast cancer. All patients were being treated at Rush-Presbyterian-St. Luke's Medical Center in Chicago, IL. Patients completed the FACT-B once at baseline, and again 2 months later. Demographic and clinical data for these patients are listed in Table I in the FACT-B publication attached previously.
The second sample consisted of 295 breast cancer patients who completed version 3 of the FACT-B as part of a large 3-year validation study of the FACT Measurement System. Data from this administration of the FACT-B were used for reliability analyses, as well as to further assess validity. These patients were recruited from the following medical centers: Rush-Presbyterian-St. Luke's Medical Center, Chicago, IL; Grady Memorial Hospital (Emory University),
Atlanta, GA; and Cook County Hospital, Chicago, IL. To participate in this study, patients had to meet the following criteria: (1) be diagnosed with breast cancer (any stage of disease); (2) be older than 18 years; (3) have no metastases to the brain or CNS; and (4)not be using psychotropic medication. Demographic and clinical data for these patients are listed in Table 2 of the FACT-B publication. A subset (n = 32) of these patients completed the FACT-B on a second occasion, 3 to 7 days after the first administration, to assess test-retest reliability (stability of the instrument over time). These patients were a consecutive subsample, drawn from Cook County Hospital and Rush-Presbyterian-
St. Luke's Medical Center.

2b
C☐
P☐
M☐
N☐

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
The FACT-G total score provides a useful summary of overall quality of life across a diverse group of patients. The FACT-B questionnaire total score further augments the FACT-G summary score by adding the breast cancer-specific subscale. Two alternative approaches are also noteworthy. One is to separately analyze the FACT-G total score and the breast cancer-specific subscale score. Another is to select subscales of the FACT which are most likely to be changed by the intervention being tested. For example, the Physical, Functional, and breast Cancer-specific subscales would be most likely to change in a chemotherapy clinical trial. One could also consider creating a separate a priori index summing two or three subscales into a 21-item Trial Outcome Index (Cella et al, 1997). On the other hand, the Emotional or Social Well-being subscale would be expected to change most when evaluating a psychosocial intervention.

Standardized scores are also available for the FACT-G portion of the FACT-L. In order to derive standardized scores (ranging from 0-100) for each scale (Physical Well-Being, Social/Family Well-Being, Emotional Well-Being, and Functional Well-Being) and total score, 27 FACT-G items were analyzed using the Rasch rating scale model (Wright & Masters, 1982). The data were from a heterogeneous group of cancer and HIV patients. Individual quality of life measurements for the separate FACT-G scales were calibrated using the BIGSTEPS computer program (Wright & Linacre, 1997). The obtained scaled measures, expressed in logits with internal measurement properties, were then transformed linearly, to a 0-100 scale (standardized scores), with zero representing worst quality of life and 100 representing best quality of life. The values on both scales (logits or transformed logits) are interval. The standardized scores can easily be derived from raw scores by using a conversion table (provided in the scoring section of the FACIT manual). Work on validating the standardized scores and their usefulness in research and clinical trials is continuing.

MIDs can be determined using both distribution-based and anchor-based methods (Lydick, et al. 1993 & Crosby, et al. 2003). Distribution-based measures are based on statistical distributions, and include effect size measures (Cohen, et al. 1988, Deyo, et al. 1991 & Kazis, et al. 1989), the standard error of measurement (SEM) (Wyrwich, et al. 1999 & Wyrwich, et al. 1999), the responsiveness index (Guyatt, et al. 1987) and the reliable change index (Jacobson, et al. 1991). Anchor-based methods 'anchor' or map score differences onto differences in clinical measures. Clinical measures can be objective indicators (e.g., response to treatment) or subjective assessments of patient status (e.g., performance status rating, global ratings of change in health-status). Anchor-based differences can be determined either cross-sectionally at a single time point or longitudinally across multiple time points.

The following discusses the internal consistency of the FACT-B as published in the first validation manuscript. Chronbach's alpha for the nine-item BCS was .63. Alpha coefficients for the other subscales of the FACT-G were in the 0.69 to 0.86 range. FACT-G and FACT-B total score alpha coefficients were both 0.90. The alpha coefficient for the TOI-PFB was 0.88

Test-retest reliability. Test-retest correlation coefficients for subscales and aggregate scores never previously published were .88 for the BCS, .89 for the TOI-PFB, and .85 for the FACT-B total score, indicating a high degree of stability over time periods (3 to 7 days) during which no change would be expected.
FACT-B Descriptive Statistics: Sensitivity to Change Sample'
Scale Meant SD
PWB (7-item) 21.0 5.8
SWB (7-item) 23.3 4.2
RWD (2-item) 7.1 0.9
EWB (5-item) 15.6 2.9
FWB (7-item) 20.3 5.1
FACT-G total score (28-item) 87.2 13.7
BCS (9-item) 24.6 4.6
FACT-B total score (37-item)t 111.8 16.3
FACT-B TOI-PFB (23-item)§ 65.9 12.5
*n = 47.
tHigher scores reflect better QL.
tFACT-G plus BCS.

§PWB plus FWB plus BCS.

Sensitivity of FACT-B to 2-Month Changes in PSR
Scale Group No. of Patients (N = 47) Mean Change Score (S.D) t (1,44) P
PWB (7-item) Declined PSR 8 -3.3 (1.6)
No change 29 1.2 (3.3) -4.37 < .001
Improved PSR 10 3.1 (4.2)
SWB (7-item) Declined PSR 8 0.1 (4.0)
No change 29 0.2 (2.8) 0.70 .491
Improved PSR 10 -1.0 (4.4)
RWD (2-item) Declined PSR 8 0.5 (0.5)
No change 29 -0.2 (1.0) 1.21 .250
Improved PSR 10 0.1 (1.4)
EWB (5-item) Declined PSR 8 0.0 (1.5)
No change 29 0.7 (2.8) -0.68 .499
Improved PSR 10 0.8 (2.2)
FWB (7-item) Declined PSR 8 -1.9 (4.5)
No change 29 -0.1 (3.1) 2.22 .031
Improved PSR 10 1.9 (4.2)
FACT-G total score (28-item) Declined PSR 8 -4.5 (4.1)
No change 29 1.8 (8.2) -2.68 .010
Improved PSR 10 4.7 (6.0)
BCS (9-item) Declined PSR 8 -3.0 (4.6)
No change 29 0.1 (3.8) -2.61 .012
Improved PSR 10 1.9 (3.8)
Total FACT-B (37-item)* Declined PSR 8 7.5 (5.9)
No change 29 1.9 (8.8) -3.53 .001
Improved PSR 10 6.6 (8.8)
TOI-PFB (23-item)t (PWB + FWB + BCS) Declined PSR 8 -8.1 (5.0)
No change 29 1.2 (6.5) -4.92 < .001
Improved PSR 10 6.9 (7.3)
*FACT-G plus BCS.
tPWB plus FWB plus BCS.

It should be noted that additional work has been done in recent years to modify and shorten existing measures for more brief assessments to lessen respondent burden (and reduce clinical trial costs). The 6 item FBSI (Breast Symptom Index) was developed. Herewith are the descriptives of those analyses which are very similar to analyses done with all FACIT questionnaires.

For the FBSI, Internal consistency reliability of the FBSI-6 was assessed
with Cronbach's alpha coefficient. Convergent and divergent validity were assessed with Spearman correlation coefficients between FBSI scores and FACTB
scale scores. Higher correlation coefficients were expected with physically-oriented FACT-B scales such as PWB, FWB and TOI-PFB, whereas lower correlations
were expected with non-physical scales such as EWB, SWB, and TOI-ESB. Spearman correlation coefficients were also calculated between the FBSI-6 score and two
clinical variables, pain and ECOG PSR.

The ability of the FBSI-6 to differentiate clinically distinct patients according to their pre-treatment pain and ECOG PSR was investigated cross-sectionally using ANOVA. Longitudinal analyses (one-way ANCOVA) were also conducted to determine the ability of FBSI change scores to differentiate patients by best overall treatment response.

The minimally important difference (MID) in FBSI-6 scores, representing the smallest score difference that is clinically significant and likely to be meaningful to patients and physicians, was determined by distribution and
anchor-based techniques. Distribution-based measures included 1/3 and 1/2 standard deviations for pretreatment and week 16 scores and for longitudinal change scores. We also computed the standard error

of measurement (SEM)at pre-treatment and week 16. Results from the cross-sectional and longitudinal analyses provided anchor-based MID estimates. FBSI score differences for patients in adjacent clinically distinct groups (e.g., ECOG 1 versus ECOG 0) were estimates of the MID. Similarly, FBSI-6 change scores across categories of response to treatment were estimates of the MID. **2b.3 Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*: The FACT-B was found to be reliable and valid. In the years since this paper was published in 1997, this questionnaire has been used in a significant number of clinical and cooperative group trials. The items are all now well validated and translated into over 50 languages. Further the items are now part of the PROMIS item banks and as such, the item parameters are well documented.	
2c. Validity testing **2c.1 Data/sample** *(description of data/sample and size)*: Patients were eligible for this study if they met the following criteria: (1) presented with advanced (stage III or IV) breast cancer; (2) were currently receiving chemotherapy and/or radiation therapy; (3) could read and speak English; and (4) were not impaired cognitively with CNS metastasis, overt psychosis, major depression, or delirium. The data used in this report were derived from two samples. The first sample was tested twice over a 2-month periodand was therefore used for the sensitivity to change analysis. Because this sample also completed a QL instrument, the Functional Living Index-Cancer (FLIC), 2 ° data from this sample were used to assess concurrent validity as well. This first sample consisted of 47 breast cancer patients who completed version 1 of the FACT-B as part of a funded psychosocial QL intervention project targeting patients with advanced (stage III or IV) breast cancer. All patients were being treated at Rush-Presbyterian-St. Luke's Medical Center in Chicago, IL. Patients completed the FACT-B once at baseline, and again 2 months later. The second sample consisted of 295 breast cancer patients who completed version 3 of the FACT-B as part of a large 3-year validation study of the FACT Measurement System. Data from this administration of the FACT-B were used for reliability analyses, as well as to further assess validity. **2c.2 Analytic Method** *(type of validity & rationale, method for testing)*: See discussion of various analytic methods we use in 2.b.2 **2c.3 Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*: Chronbach's alpha for the nine-item BCS was .63. Alpha coefficients for the other subscales were in the 0.69 to 0.86 range. FACT-G and FACT-B total score alpha coefficients were both 0.90. The alpha coefficient for the TOI-PFB was 0.88. Test-retest reliability. Test-retest correlation coefficients for subscales and aggregate scores never previously published were .88 for the BCS, .89 for the TOI-PFB, and .85 for the FACT-B total score, indicating a high degree of stability over time periods (3 to 7 days) during which no change would be expected.	**2c** C☐ P☐ M☐ N☐
2d. Exclusions Justified **2d.1 Summary of Evidence supporting exclusion(s):** n/a **2d.2 Citations for Evidence:** n/a **2d.3 Data/sample** *(description of data/sample and size)*: n/a **2d.4 Analytic Method** *(type analysis & rationale)*: n/a	**2d** C☐ P☐ M☐ N☐ NA☐

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*: n/a	
2e. Risk Adjustment for Outcomes/ Resource Use Measures **2e.1 Data/sample** *(description of data/sample and size)*: n/a **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: n/a **2e.3 Testing Results** *(risk model performance metrics)*: n/a **2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:** Not needed.	**2e** C☐ P☐ M☐ N☐ NA☐
2f. Identification of Meaningful Differences in Performance **2f.1 Data/sample from Testing or Current Use** *(description of data/sample and size)*: Data were from ECOG study E1193, a clinical trial of female breast cancer patients with progressing regional or metastatic disease (see Sledge et al. 2003 for details. This three-arm Phase III trial compared treatment with (a) doxorubicin versus (b) paclitaxel versus (c)doxorubicin + paclitaxel + filgrastim (G-CSF). HRQL data were collected at pre-treatment and 16-week follow- up using the Functional Assessment of Cancer Therapy – Breast (FACT-B). **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: Please note these MID's are for the FBSI-6, not the FACT-B (the longer version of the questionnaire. At the time of this submission, MID analyses for FACT-B were not able to be found, so we are submitting FBSI MID analysis results). Using distribution-based criteria, the MIDs for the FBSI-6 ranged from 1.5 to 2.5 points with a median of 2.2 points. Cross-sectional anchor-based criteria yielded similar estimates, with score differences between adjacent categories of PSR of 2.8 and 3.2 points, and score differences between adjacent categories of pain of 2.4 and 2.3 points (Table 2). The results from the longitudinal analysis were not informative in determining MIDs. Based on the distribution-based and cross-sectional anchor-based analyses, 2–3 points on the FBSI-6 can be interpreted as an MID. **2f.3 Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*: FBSI: The FBSI-6 was significantly correlated with all FACT scales ($p < .0001$), most strongly with PWB, FWB,and TOI-PFB and to a lesser extent with SWB, EWB, BCS, and TOI-ESB (Table 1). FBSI-6 demonstrated stronger associations with PSR and pain (r ¼)0.42 and r ¼)0.44, respectively) relative to all other FACT scales (mean r ¼)0.22 for PSR and mean r ¼)0.27 for pain). Differentiation between known groups In cross-sectional analyses, the FBSI-6 distinguished across ECOG PSR and pain ratings (Table 2). In longitudinal analyses, there was little change in FBSI-6 scores, and the changes were not consistently in the expected direction. Minimally Important Difference (MID) Using distribution-based criteria, the MIDs for the FBSI-6 ranged from 1.5 to 2.5 points with a median of 2.2 points. Cross-sectional anchor-based criteria yielded similar estimates, with score differences between adjacent categories of PSR of 2.8 and 3.2 points, and score differences between adjacent categories of pain of 2.4 and 2.3 points (Table 2). The results from the longitudinal analysis were not informative in determining MIDs. Based on the distribution-based and cross-sectional anchor-based analyses, 2–3 points on the FBSI-6 can be interpreted as an MID.	**2f** C☐ P☐ M☐ N☐
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: The FACT-B is widely used in different studies in many different patient populations and in many different ways. There is not a known compilation of	**2g** C☐ P☐ M☐

different studies' comparitive results.

N☐ NA☐

2g.2 Analytic Method *(type of analysis & rationale)*:

2g.3 Testing Results *(e.g., correlation statistics, comparison of rankings)*:

2h. Disparities in Care

2h.1 If measure is stratified, provide stratified results *(scores by stratified categories/cohorts)*: n/a

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:

The FACT-B can help identify disparities in care/treatment regime as outlined in the literature. We have also done significant work in identifying challenges for low-literacy patients and in cross-cultural populations. We have also assessed different methods of administration to reduce patient burden, all with the hope of reducing assessment burden across all populations.

All FACT scales are designed for patient self-administration, but can also be administered by interview format. Interview administration is considered appropriate after adequate training of interviewers so as to elicit non-biased patient responses. Technical (mode of administration) and statistical equivalence of similar scales in our measurement system have been demonstrated, providing the user with some flexibility as to mode of assessment (self versus interviewer administration) literacy level (high versus low) and language (English versus Spanish). One of the aims of a recently completed large multicenter study of cancer (n = 2356) patients was to test the psychometric properties and statistical equivalence of the English and Spanish language versions of the FACT subscale across literacy level (low vs. high) and mode of administration (self vs. interview). This sample included breast cancer patients. Technical equivalence across mode of administration was demonstrated in the high literacy patients; there were no differences in data quality or in mean QOL scores, after adjustment for performance status rating, socioeconomic status, gender and age. Technical equivalence between modes of administration with the FACT permits unbiased assessment of the impact of chronic illnesses and their treatments on patients from diverse backgrounds (Hahn & Cella, 1997).

We have additional data to support the appropriateness of computer-administered versions of the questionnaire, including a multimedia touch screen program (Hahn & Cella, 2003). We are currently developing other novel administration methods such as computer-assisted telephone and web-based administration. Across these modes of administration, our preliminary data suggest that while there are small differences in the way people respond based on mode of administration, these alternate formats are essentially equivalent, particularly when deriving group statistics (e.g., means and variances).

There has been much work done with all the FACT scales to assess differences in responses between Latinos, patients with low literacy issues, different cultures, treatment regimens, genders, and many other characteristics. (Wan, G.J., Counte, M.A., Cella, D., Hernandez, L., McGuire, D., Deasy, S., Shiomoto, G., & Hahn, E. (1999) The impact of socio-cultural and clinical factors on health-related quality of life reports among Hispanic and African-American cancer patients. Journal of Outcome Measurement, 3(3), 200-215 a118 and Wan, G.J., Counte, M.A., Cella, D., Hernandez, L., Deasy, S., Shiomoto, G. (1999). An analysis of the impact of demographic, clinical and social factors on health-related quality of life. Value in Health, 2(4), 308-318, to name two such publications from our group). Current efforts in Item Response Theory (IRT) through the NIH-funded PROMIS (Patient Reported Outcomes Measurement Information System - U01 AR 052 177), under the statistical direction of David Cella, developer of the FACIT system, are significantly strengthening the ability of clinicians and researchers to detect differences at the item level across these groups with the specific intent of measuring and reducing disparities which result from socio-economic, literacy and language issues.

2h C☐ P☐ M☐ N☐ NA☐

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Scientific Acceptability of Measure Properties?*

2

Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, **met?**
Rationale:

2 C☐ P☐

	M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	Eval Rating
3a. Meaningful, Understandable, and Useful Information 3a.1 **Current Use:** in use 3a.2 **Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*)**:** The FACIT items are currently being used in several NIH-funded initiatives which are being used in public and general health status assessments. Included in these initiatives are PROMIS (U01 AR 052 177), NeuroQOL (HHSN 265200436), Toolbox (AG-260-06-01) and others. 3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** The FACT-B (and other FACIT questionnaires) are widely used in clinical trials and clinics to improve the quality of clinical care for cancer patients. In addition to the aforementioned PROMIS, NeuroQOL and Toolbox projects, the use of these questionnaires is mainstream in cooperative group oncology trials for assessing the impact of treatment on patients' QOL. Most noteably the PROMIS project's Assessment Center (www.nihpromis.org) is now available for widespread public use. Assessment Center is an online publicly available system which clinicians and researchers can use to capture patient-reported data. It allows for CAT and contains specific items and item parameters (including the FACT and FACIT items. To date there are over 13 different item banks (questions/items in domains such as Social Well Being, Fatigue, Pain, etc), the measurement characteristics of which have already been calculated by Dr. Cella and colleagues in the PROMIS initiative. Dr. Cella is also one of the founding members of the PROMIS Health Organization, a non-profit organization developed to support the ongoing PROMIS initiative. Other participants include faculty from the NIH, researchers from academic institutions, clinicians and representatives of the pharmaceutical industry. Dr. Cella has granted the PROMIS, Toolbox and NeuroQOL item banking projects permission to use all FACIT system items. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) 3a.4 **Data/sample** *(description of data/sample and size)***:** The data samples and publications on FACT-B data in previous sections of this submission demonstrate the widespread use and acceptance of this questionnaire by clinicians and researchers. 3a.5 **Methods** *(e.g., focus group, survey, QI project)***:** Data and analyses from FACT-B surveys in all forms of clinical interventions, clinical trials and NIH-funded initiative such as PROMIS demonstrate the various uses of the FACT-B. 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)***:** Data and analyses from FACT-B surveys in all forms of clinical interventions, clinical trials and NIH-funded initiative such as PROMIS demonstrate the various uses of the FACT-B.	**3a** C☐ P☐ M☐ N☐
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:**	
(for NQF staff use) **Notes on similar/related** endorsed **or submitted measures:**	

3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?**	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **Competing Measures If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, **met?** **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Survey,	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** C☐ P☐ M☐ N☐
4c. Exclusions 4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No 4c.2 **If yes, provide justification.**	**4c** C☐ P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences 4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** Perhaps the biggest source of inaccuracies in QOL data is missing data in the questionnaires. Until recently most data was collected via paper and pencil, resulting in missed responses which were then imputed during data analysis. Recent developments in use of electronic collection of health status assessments has reduced missing data, however, those methods are subject to the budgetary constraints of the study	**4d** C☐ P☐ M☐ N☐

sponsor. In the past several years, Dr. Cella and his colleagues have made impressive advances in IRT and CAT (computerized adaptive testing) which significantly reduces patient/respondent burden by lowering the number of items/questions required to produce a QOL score. This type of assessment requires access to a computer and/or the internet, which is also dependent on sponsor funding. It also reduces the likelihood of including low socio-economic participants.	
4e. Data Collection Strategy/Implementation 4e.1 **Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** As stated above, prior to 3-4 years ago, QOL data was largely collected via paper and pencil which resulted in missing data. The missing data is dealt with via several different widely published statistical analyses methods (Bernhard, J., Cella, D., Coates, A., Fallowfield, L., Ganz, P.A., Moinpour, C., Mosconi, P., Osoba, D., Simes, J., & Hurny, C. (1998). Missing quality of life data in cancer clinical trials: Serious problems and challenges. Statistics in Medicine, 17, 517-532.) The timing and frequency of data collection is dependent on the type of disease, treatment or symptom being assessed. Patient confidentiality is handled differently according to type of assessment: if electronic, there are encryption and password protections required by HIPAA which are implemented in the database development; If paper and pencil, study coordinators are responsible for ensuring files are locked and monitored, again according to HIPAA guidelines. The largest cost of data collection for paper and pencil is the Research Assistant or questionnaire training staff time, as well as the data entry and management time. These costs are largely bypassed by ePRO (electronic Patient Reported Outcomes)assessments, however for ePRO, there are significant computer programming costs. When IRT is included, there are also significant psychometrician and biostatistician algorithm development costs. 4e.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** There is no cost for the use of any of the English versions of the FACIT measures. Licensing costs for use of the non-English multilingual versions are $1500 per subscale, per language, per trial for Roman font alphabet languages (ie French, German, Spanish) and $2000 per subscale, per language, per trial for non-Roman font languages (ie Greek, Hebrew, Russian). 4e.3 **Evidence for costs:** The evidence of these costs is 15 years' experience in NIH-funded research with these scales (including cooperative group oncology trials) as well as consulting with pharmaceutical companies who use the FACIT scales in their trials. It should be noted for the FACIT Fatigue that it is a short-form (only 13 items). Short forms/symptom indices allow for a more brief assessment which is less expensive to put into clinical practice or clinical trials. However, such a short form does not provide a full QOL measure since the other domains (such as social/family well being) are not assessed. 4e.4 **Business case documentation:** The clinical trials industry uses QOL endpoints as a secondary endpoint for label claims. NIH-funded initiatives (noteably AHRQ) are including patient perspective of treatment burden for comparitive effectiveness research initiatives.	**4e** **C☐** **P☐** **M☐** **N☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Feasibility*?	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility***, met?** **Rationale:**	**4** **C☐** **P☐** **M☐** **N☐**

RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited** ☐
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y** ☐ **N** ☐ **A** ☐

CONTACT INFORMATION
Co.1 **Measure Steward (Intellectual Property Owner)** Co.1 **Organization** David Cella at FACIT.org | 381 S. Cottage Hill Avenue | Elmhurst | Illinois | 60126 Co.2 **Point of Contact** Lauren | Lent, M.S. | l-lent@northwestern.edu | 630-531-7959
Measure Developer If different from Measure Steward Co.3 **Organization** David Cella at FACIT.org | 381 S. Cottage Hill Avenue | Elmhurst | Illinois | 60126 Co.4 **Point of Contact** Lauren | Lent, M.S. | l-lent@northwestern.edu | 630-531-7959
Co.5 **Submitter If different from Measure Steward POC** Lauren | Lent, M.S. | l-lent@northwestern.edu | 630-531-7959- |David Cella at FACIT.org
Co.6 **Additional organizations that sponsored/participated in measure development**

ADDITIONAL INFORMATION
Workgroup/Expert Panel involved in measure development Ad.1 **Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**
Ad.2 **If adapted, provide name of original measure:** The FACT-B is the original measure. We also have the FBSI-6 which is a 6 item shortened Breast Cancer symptom index, derived from the FACT-B. Ad.3-5 **If adapted, provide original specifications URL or attachment**
Measure Developer/Steward Updates and Ongoing Maintenance Ad.6 **Year the measure was first released:** 1997 Ad.7 **Month and Year of most recent revision:** 2005-04 Ad.8 **What is your frequency for review/update of this measure?** Due to the item banking work in PROMIS, all the FACIT items are under continual review Ad.9 **When is the next scheduled review/update for this measure?** 2010-03
Ad.10 **Copyright statement/disclaimers:** Copyright 1987,1997
Ad.11 -13 **Additional Information web page URL or attachment:** Attachment FACT-B_ENG_Final_Ver4_16Nov07-633978443939539718.doc
Date of Submission (*MM/DD/YY*)**:** 12/31/2009