The New School Psychology Bulletin
2009, Vol. 6, No. 1

Print ISSN: 1931-793X; Online ISSN: 1931-7948

Research Report

Quality of Life Differences In an African American and Caucasian Sample of Chronic Illness Patients: Assessment of Differential Item Functioning

Scott M. Debb, M.A.
Center on Outcomes, Research and Education (CORE)

David L. Blitz, M.A.
Center on Outcomes, Research and Education (CORE)
Illinois Institute of Technology

Seung W. Choi, Ph.D.
Center on Outcomes, Research and Education (CORE)
Institute for Healthcare Studies, Feinberg School of Medicine, Northwestern University

The value of a health care intervention is judged by its impact on the quantity and quality of life (QOL) of an individual (Cella & Bonomi, 1996). When analyzing QOL data, results from the general population are often assumed to be equally representative of various subpopulations; however, establishing measurement invariance across different populations (e.g., racial and ethnic groups) is crucial for score interpretability. In this study, ordinal logistic regression (following Zumbo, 1999) was used to examine differential item functioning (DIF) in an African American and Caucasian sample of 898 cancer and chronic illness patients that had been administered a QOL questionnaire for use with chronic illness sufferers. Results indicated that significant differences in response patterns were present between the two groups. This has which have clinical implications pertaining to psychological assessment, conceptualization and intervention.

Keywords: Quality of Life (QOL), Differential Item Functioning (DIF), Cross-cultural, Race/ethnicity, cancer, assessment

Differences in health care treatment and prognosis in the United States have been extensively studied and identified when comparing groups of people across sociodemographic variables such as race, ethnicity or education. In cancer research, outcomes studies on racial disparities have investigated sociodemographic group differences in the context of mortality rates, stage of disease and treatment options (Ashing-Giwa et al., 2004; Delfino, Ferrini, Taylor, Howe, & Anton-Culver, 1998; Mettlin, Murphy, Cunningham, & Menck, 1997; Shinagawa, 2000). Quality of life (QOL) assessment also focuses on these types of disparities, although less research has been conducted focusing specifically on item level differences between racial groups in favor of reporting more global findings.

To illustrate, Hoffman and colleagues (2001) found that underemployment, poor health insurance coverage and lower socioeconomic status (SES) were significantly related to advanced-stage prostate cancer in men across racial groups. Penedo, Dahn, Shen, Schneiderman, and Antoni (2006) found that African-American and Hispanic men living with prostate cancer reported lower health-related QOL than Caucasian men living with prostate cancer. Thompson et al. (2001) found that African-American men with prostate cancer showed higher mortality rates, more extensive disease at time of diagnosis, poorer performance status, younger age, and worse prognosis than Caucasian men with the same diagnosis. Other studies have found that African-American men with prostate cancer tend to report a poorer level of physical well-being than Caucasian men (Eton, Lepone, & Helgeson, 2001; Krupski et al., 2005; Maliski, Kwan, Orecklin, Saigal, & Litwin, 2005). Lubeck, Kim, Grossfeld, Ray, and Penson (2001) found that African-American men with prostate cancer reported poorer physical well-being than Caucasian men with prostate cancer, and also slower rates of improvement in their QOL.

Cross-Cultural Concerns

Much of the research in health related QOL for chronic illness sufferers has focused on the physical well-being dimension of QOL (Cella & Bonomi, 1996; Zickler, 2009), rather than cross-cultural differences that may

---

Correspondence concerning this article should be addressed to Scott M. Debb, Center on Outcomes, Research and Education (CORE) 1033 University Place, Suite 110, Evanston, IL, 60201. Tel: +001 224–364–7533. Fax: +001 847–570–8018. E-mail: sdebb@northshore.org.

pertain to accessing healthcare or psychological services. Cultural values can affect many aspects of an individual's lifestyle, including how they interact with others in a shared physical or social context. When assessing QOL, these cultural factors must be considered because they are often pervasive aspects to a person's functioning, as values are passed from one generation to the next (Sattler, 2001; Whaley & Davis, 2007).

Although some research has examined specific cultural factors, identifying the stronger underlying cultural associations relating to QOL is vague. Fave (2006) noted that QOL research should be considered in relation to health as well as "other areas and activities of daily life" (p. 165). Related to this idea, researchers have reported on the significance of psychological factors affecting chronic pain sufferers, including subjective experience of pain, depression, stress, anxiety, and capacity to cope with their situation (Gonzales, Martelli, & Baker, 2000; Park & Adler, 2003; Turk & Okifuji, 2002; Wool & Mor, 2005).

Specific to the chronic illness population, Gallagher (2001) noted that using a biomedical model for conceptualizing chronic pain is not sufficient due to the circumstances imposed on the person by the illness. Rather, Gallagher noted that a biopsychosocial model of assessment is more effective because it aids in identifying functioning related to QOL as well as taking into considering cultural and developmental factors, which can help establish more appropriate treatment goals and therapeutic interventions. To this end, research studies often utilize only a "tiny proportion of the world's cultural richness (and) don't take into account the societal factors that shape behavior and preferences" (Meyers, 2007, p. 36).

Measurement of QOL is accomplished by applying psychometrically validated data pertaining to the physical and mental health of an individual's well-being (The Health Institute, 1993). Despite this, it may not be feasible to fully assess the severity of the interaction between symptoms and the person's ability to function because of the possibility of upper and lower end extremes canceling each other out making the data seem that a person is functioning at a reasonable level. When extremes such as this are present, it can be due to socio-cultural differences, which can lead a clinician to accept false hypotheses (Dana, 1993). Further, while the use to which a measure is put may be psychometrically valid, a chronic illness patient may be hesitant to convey all the necessary information because they might feel embarrassed, ashamed or be unable to effectively communicate with their healthcare provider due to language, education or cultural differences (Roberts, 2001; Tchen et al., 2003).

For example, Vazquez (2007) stated that in African American men in particular, fears of medicine as an establishment in our society often contribute to a poorer medical health status in general. Green et al. (2003) and Hahn and Cella (2003) noted that African Americans may have more difficulty accessing quality healthcare related services due to a lack of community or personal resources. This has the potential to have a negative impact on the effectiveness of the therapeutic relationship between patient and clinician (Day, 2004; Hoshmand, 2001).

Measuring probabilities of item endorsement

A frequently used and widely accepted method of measuring QOL in healthcare is by patient self-report (Barofsky & Sugarbarker, 1990; Fayers & Hays, 2005; Loge et al., 2004; Skeel, 1989). Patient reported outcomes (PRO) measures can consist of qualitative or quantitative components to gauge the different components of ability to function daily. The scaled questionnaire items are said to be fair by examining whether or not people who belong to different groups (e.g., African American and Caucasian), assuming the same latent trait, respond comparably on that item. This is to say that fairness is defined by assessing if the probability of endorsing any given item in a particular fashion is dependent only on an underlying trait, or also on group membership. One method to assess variability of item endorsement is by examining differential item functioning (DIF). As defined by Zumbo (1999), "DIF occurs when examinees from different groups show differing probabilities of success on (or endorsing) the item after matching on the underlying ability that the item is intended to measure" (p. 12).

There are various techniques that have been proposed in the measurement literature for the assessment of DIF. Selection of a technique is based in part on the format of the data being analyzed. For example, the Mantel and Haenszel's (1959) chi-square procedure, which was initially developed for use in epidemiological research, is used to detect DIF for dichotomous data. The Mantel procedure (1963) is the extension of this procedure for polytomous data, although it can only detect uniform DIF. Simultaneous Item Bias Test (SIBTEST; Shealy & Stout, 1993) and Poly-SIBTEST (Chang, Mazzeo, & Roussos, 1995) are nonparametric methods of DIF detection, using the latent trait as the matching criterion.

Another approach is based on logistic regression. Swaminathan and Rogers (1990) originally proposed the use of logistic regression for DIF detection with dichotomous items, and this approach was extended for polytomous items by Miller and Spray (1993), French and Miller (1996), and Zumbo (1999). The use of logistic regression in this context is based on the comparisons of nested models, where their likelihood values of predictors such as a total score or group membership are compared. For example, where total score is referred to as (TS), group membership is referred to as (G), and you have an item score Y with K categories and a cumulative probability of k, a baseline model (Model 0, or "M0") might be specified as $\text{logit}[P(Y > k)] = \alpha_k$. Three nested models might be $\text{logit}[P(Y > k)] = \alpha_k + \beta_1(TS)$ (Model 1 or "M1"), $\text{logit}[P(Y > k)] = \alpha_k + \beta_1(\text{TS}) + \beta_2(\text{G})$ (Model 2 or "M2"), and $\text{logit}[P(Y > k)] = \alpha_k + \beta_1(TS) + \beta_2(G) + \beta_3[(TS) \text{ x } (G)]$ (Model 3, or "M3"). Following is a more detailed description of the model from Kim, Cohen, Alagoz, & Kim (2007), although a more thorough treatment of the procedure can be found in their recent article.

The first test (M1 vs. M2) is used to detect uniform DIF. In general, uniform DIF can be construed as DIF occurring in the same direction across the trait continuum; more concretely, this can be seen as the difference in the cumulative category probability across groups as being consistent across all levels of the total score. The second test (M1 vs. M3) is a simultaneous test of uniform and nonuniform DIF, consistent with the methodology originally proposed by Swaminathan and Rogers (1990). In other words, while it is not a 'pure' test of nonuniform DIF, by convention it is what is used. Zumbo (1999) proposed that the comparison of M2 vs. M3 should strictly yield a test of nonuniform DIF; however, "this strategy needs to be tested further with real data … [but] it may be useful as a data analytic strategy" (p.26). In general, nonuniform DIF can be construed as an item favoring one group at one level and another group at another level; more concretely, this can be seen as the difference in the cumulative category probability across groups as being inconsistent across all levels of the total score. The detection of both uniform and nonuniform DIF is one of the advantages of the logistic regression technique over some of the earlier techniques, such as Mantel's (1963) procedure, which can only detect uniform DIF.

The approach used in this study most closely approximates logistic regression with cumulative logits for ordinal responses (Zumbo, 1999; see also Agresti, 1990). This technique was chosen because of its capacity to detect both uniform and nonuniform DIF by way of nested models. This includes detecting the presence of both forms of DIF simultaneously. Further, logistic regression is a parametric procedure that utilizes observed score as the matching criteria, rather than latent trait.

Purpose of study

The purpose of this study was to investigate racial group response differences in various domains of health related QOL. Using a large and diverse sample of participants with breast, colon, head and neck, and lung cancer, as well as HIV/AIDS related malignancies, we investigated group differences between African Americans and Caucasians for symptoms reported on a health related QOL questionnaire. This included an item-level analysis of questions assessing the physical, social/family, emotional, and functional well-being dimensions of QOL as identified by Cella et al. (1993).

Methods

A secondary analysis was conducted using QOL data that were originally collected as part of the Bilingual Intercultural Oncology Quality of Life (BIOQOL) project: a cross-sectional validation study of the FACT-G (Cella et al., 1993) and Functional Assessment of Human Immunodeficiency Virus Infection (FAHI; Cella & Bonomi, 1996; Cella, McCain, Peterman, Mo, & Wolen, 1996). Aims of the project were to validate the measure's scales across language (English, Spanish), race (African-American non-Hispanic, Caucasian non-Hispanic, Hispanic), literacy (high, low), and mode of administration (interview, self-administration). For this study, the scope of the analysis was limited to data from the African American and Caucasian participants that spoke English.

Participants

During the initial data collection process for the BIOQOL study, 3,329 chronic illness patients were recruited from cancer centers in Atlanta, Chicago and Puerto Rico. Participants were eligible to be included in the study if they had a cancer or AIDS-related diagnosis, were at least 18 years of age, spoke English or Spanish, and provided informed consent in accordance with institutional review board requirements. Table 1 outlines the relevant demographic data that was collected as part of the BIOQOL project.

Functional Assessment of Cancer Therapy-General (FACT-G)

At the time of field testing for the BIOQOL project, the FACT-G Version 3; (Cella et al., 1993) was the most current version being used. This measure was originally developed as a 34-item questionnaire, with the majority of the items consisting of Likert-scaled response categories with scores ranging from 0-4. Response categories include "Not at all," "A little bit," "Somewhat," "Quite a bit" and "Very much." Negatively worded items (e.g. "I have a lack of energy" or "I have nausea") are reverse-scored so that patients who experience a higher severity of the symptom receive a lower point value; therefore, patients with higher scores are reporting a higher QOL in relation to that symptom, and lower scores correspond to a lower QOL.

As reported in Cella et al., a principal components factor analysis was conducted during the initial validation of the FACT-G measure, which yielded a four-factor structure that resulted in the Physical, Social/ Family, Emotional and Functional Well-being subscales. Cella and colleagues also provided evidence of construct validity by analyzing the association between FACT-G scores and other similar measures available at that time, and the observed Pearson correlations were within the researchers' expected range for all measures. The authors also assessed reliability with previously untested outpatients (n=70) with cancer related diagnoses. Participants were administered the FACT-G and then asked to complete the measure again after three to seven days. Eighty-six percent of the participants completed the second administration and the test-retest correlations ranged from 0.82-0.92, indicating that the measure was reliable.

Table 1
Demographic data of sample

	Total N=898
Ethnicity	
African American	502 (55.9%)
Caucasian	396 (54.1%)
Sex	
Female	588 (65.5%)
Male	310 (34.5%)
Education	
< High School	257 (28.6%)
High School/ GED	478 (53.2%)
Bachelors degree	99 (11%)
Graduate education	64 (7.1%)
Literacy	
High Literacy	685 (76.3%)
Low Literacy	213 (23.7%)
Diagnosis	
Breast cancer	390 (43.4%)
Colon cancer	113 (12.6%)
Head/ Neck cancer	73 (8.1%)
Lung cancer	210 (23.4%)
AIDS-related malignancies	112 (12.5%)
Insurance Status	
No current insurance	167 (18.6%)
Any insurance coverage	731 (81.4%)
Performance Status	
Normal activity	358 (39.9%)
Some symptoms	281 (31.3%)
Bed rest <50% of the day	189 (21.0%)
Bed rest >50% of the day	68 (7.6%)

At the time of the present study's secondary analysis of the data collected for the BIOQOL project, version 4 of the FACT-G had been developed. The revised version was developed "to enhance clarity and precision of measurement without compromising its established reliability and validity" (Lent, Hahn, Eremenco, Webster, & Cella, 1999, p. 695). Item content did not change between the two versions (Center on Outcomes, Research and Education, 1997), although the total number of items in the measure reduced from 34 to 27. One additional item was not included in the present study's analysis because it exhibited a poor response rate relative to the other items. Table 2 shows the items included in the present study.

Procedure

Data from 898 English-speaking African American and Caucasian participants from the initial BIOQOL recruitment were analyzed. The number of months since the patients in sample had been diagnosed with cancer averaged 32.75 (*SD* = 46.04; *Mdn* = 15.42), ranging from 1 month to approximately 37 years. To confirm that the dimensionality of our sample was consistent with the findings from Cella et al., a principal components analysis (PCA) with varimax rotation was conducted, as was Kaiser-Meyer-Olkin (KMO) and Bartlett's tests. Item dimensionality was determined by assessing the scree plot (clean elbow), Kaiser-Guttman criterion (Eigen value greater than or equal to 1.0), total percent of variance accounted for by the set of components (75%) and the percent of variance accounted for by each component (10%).

A DIF analysis by ethnicity was conducted using a parametric, observed-score method ordinal logistic regression (following Zumbo, 1999, as implemented using a custom macro in SAS 9). One of the features of the custom macro involved the implementation of the recommendations by Shtatland, Moore, and Barton (2000) regarding obtaining pseudo-R^2 effect size measures; in this case, McFadden's (1974) pseudo-R^2 (also known as R^2_{DEV}) was specifically obtained as part of the output as an indication of impact at the item level. FACT-G subscale scores were regressed onto ethnicity as an indication of impact at the subscale level, with Caucasians as the reference group and African Americans as the focal group.

Table 2
FACT-G items by subscale

Physical Well-being

1. I have a lack of energy
2. I have nausea
3. Because of my physical condition, I have trouble meeting the needs of my family
4. I have pain
5. I am bothered by side effects of treatment
6. I feel ill
7. I am forced to spend time in bed

Social/Family Well-being

8. I feel distant from my friends
9. I get emotional support from my family
10. I get support from my friends and neighbors
11. My family has accepted my illness
12. Family communication about my illness is poor
13. I feel close to my partner (or the person who is my main support)

Emotional Well-being

14. I feel sad
15. I am satisfied with how I am coping with my illness
16. I am losing hope in the fight against my illness
17. I feel nervous
18. I worry about dying
19. I worry that my condition will get worse

Functional Well-being

20. I am able to work (include work in home)
21. My work (include work in home) is fulfilling
22. I am able to enjoy life
23. I have accepted my illness
24. I am sleeping well
25. I am enjoying the things I usually do for fun
26. I am content with the quality of my life right now

The matching criteria were specified as the observed total score for each subscale, rather than for the entire measure since the FACT-G is multidimensional in nature. To control for an inflated Type I error during DIF analysis, the Holm (1979) correction was employed, such that obtained *p*-values were compared to Holm-corrected cutoffs.

Results

Results from the KMO test for measuring sampling adequacy (0.90) and Bartlett's test of sphericity ($p<.001$) both indicated that it was acceptable to proceed with interpreting PCA results. Dimensionality was found to be generally consistent with the sample used to validate the FACT-G. The scree plot (Figure 1) produced an elbow at four components accounting for a total of 52.47% of the variance (Table 3). Each of the four factors identified accounted for over 10% of the variance. Examination of the rotated component matrix (Table 4) generally supported the loadings of items onto the four established subscales of the FACT-G.

To examine DIF between the African American and Caucasian samples, ordinal logistic regression analysis was performed on the FACT-G, with the items grouped by subscale. Item level analysis of the four subscales yielded Uniform DIF, Pseudo R^2 and Holm correction cutoff statistics. By comparing the Uniform DIF statistic and the Holm cutoff statistic, DIF detection is accomplished. This information can be found in Table 5.

Figure 1. Assessment of dimensionality.

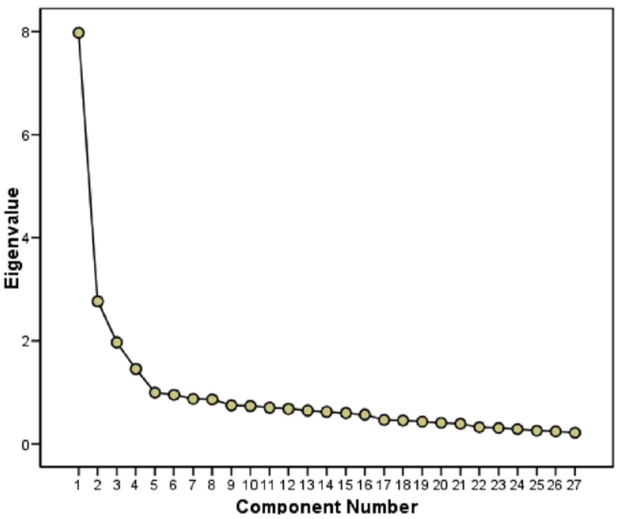

Table 3
Rotated loadings for the Eigenvalues for the four components of the FACT-G

Component	*Eigen value*	*% Variance*	*Cumulative %*
1. Physical Well-being	4.34	16.06	16.06
2. Functional Well-being	3.99	14.77	30.83
3. Emotional Well-being	2.94	10.89	41.72
4. Social/ Family Well-being	2.90	10.74	52.47

Table 4
Rotated component matrix

Item	*Component 1*	*Component 2*	*Component 3*	*Component 4*
1	**.524**	.240	.205	-.100
2	**.644**	-.066	.179	.183
3	**.574**	.369	.181	-.083
4	**.647**	.136	.083	.135
5	**.625**	.139	.035	-.012
6	**.798**	.140	.157	.044
7	**.719**	.342	.049	-.055
8	.473	.256	.244	**.266**
9	-.038	.090	.035	**.830**
10	-.002	.164	.132	**.708**
11	.065	.202	.177	**.696**
12	.221	-.132	.065	**.535**
13	.007	.290	.019	**.649**
14	.089	.531	.073	**.234**
15	.305	.203	**.553**	.153
16	.035	.455	**.294**	.339
17	.210	.043	**.573**	.126
18	.216	.171	**.673**	-.012
19	-.012	.078	**.810**	.118
20	.110	.075	**.775**	.089
21	.495	**.587**	-.106	-.024
22	.281	**.735**	-.044	.031
23	.305	**.732**	.178	.195
24	-.038	**.508**	.282	.237
25	.438	**.318**	.191	.192
26	.393	**.725**	.123	.079
27	.324	**.702**	.268	.061

Differential Item Functioning

Four items from the FACT-G exhibited significant DIF. Results indicated that African-American and Caucasian respondents exhibited different probabilities of responding to physical, emotional and functional well-being items on the measure. Specifically, on the Physical Well-being subscale, the item "I have a lack of energy" exhibited significant DIF between the African American and Caucasian samples. African American patients with comparable physical well-being were less likely to endorse this question than their Caucasian counterparts. On the Emotional Well-being subscale, the item "I worry that my condition will get worse" was less likely to be endorsed by the African American sample compared with the Caucasian sample. On the functional well-being subscale, for the item "I am able to work," Caucasian participants were more likely to endorse this item than African American participants, although African American participants were more likely to endorse the item "I am content with the quality of my life right now" than Caucasians.

Table 5

Ordinal logistic regression and Holm correction for DIF

Item	*Uniform DIF*	*Pseudo* R^2	*Holm Cutoff*	*DIF Present?*
1	0.0021	0.0331	0.0071	Y
2	0.5113	0.0091	0.0167	N
3	0.7075	0.0005	0.0250	N
4	0.1121	0.0082	0.0125	N
5	0.0606	0.0150	0.0100	N
6	0.0145	0.0132	0.0083	N
7	0.7401	0.0003	0.0500	N
8	0.8951	0.0001	0.0500	N
9	0.1212	0.0061	0.0083	N
10	0.8747	0.0001	0.0250	N
11	0.6873	0.0006	0.0167	N
12	0.1604	0.0112	0.0100	N
13	0.3661	0.0036	0.0125	N
14	0.7825	0.0002	0.0250	N
15	0.8040	0.0004	0.0500	N
16	0.0393	0.0245	0.0100	N
17	0.7254	0.0004	0.0167	N
18	0.6034	0.0008	0.0125	N
19	0.0001	0.0359	0.0083	Y
20	<0.0001	0.0455	0.0083	Y
21	0.6767	0.0005	0.0250	N
22	0.1493	0.0044	0.0125	N
23	0.0309	0.0357	0.0100	N
24	0.2814	0.0048	0.0167	N
25	0.9810	<0.0001	0.0500	N
26	<0.0001	<0.0001	0.0071	Y

Discussion

Relative to the Caucasian patients from this sample, African Americans generally reported better emotional well-being, but lower functional well-being. African-Americans reported experiencing relatively less symptoms than Caucasians in areas related to vitality and treatment side effects. African Americans also reported less worry about the worsening of their condition and generally appearing more comfortable with their current health status. Caucasians reported experiencing less severity of symptoms on items reflecting feelings about their ability to work. This suggests that the underlying dynamic related to the subjective perception of chronic illness may be interpreted differently by African Americans as compared to their Caucasian counterparts.

One hypothesis is that African-Americans may have larger emotional support networks, through community resources. Similar to our findings, Bloor, Sandler, Martin, Uchino, and Kinney (2006) found that individuals with greater social support systems were positively associated with better perceptions of their physical health but poorer health status. Utsey, Chae, Brown, and Kelly (2002) noted that ethnic identity has also been found to be a positive predictor of higher QOL for African-Americans. Further, Bynum, Burton, and Best (2007) stated that cultural and community resources can aid in reducing psychological stress.

This suggests that the medical community may not be adequately meeting client's needs at their client's level. Schwartz (2007) noted that "at every level, African Americans fare worse than Whites (and) in many cases so do other racial and ethnic groups" (p.36). Issues such as this are far reaching and can potentially affect any group or subgroup of patients, whether it is a racial or ethnic group, or a specific subgroup of patients that have been stigmatized due to a medical or psychological condition. It also relates to people's coping ability in the face of possible sociodemographic inequities, such as occupational status, income, education level or community resources.

Treatment planning and aftercare can also be affected, since effective therapy should focus on concerns from the patient's perspective (Cella, Lloyd, & Wright, 1996). For chronic illness sufferers, intervention may need to focus more on the psychological aspects of their prognosis, in addition to the necessary medical interventions as determined by their medical doctor. Uncovering differential styles of responding to health related QOL items help make healthcare professionals aware of latent dynamics existing between QOL and functionality, and facilitates the process of clinical conceptualization from the patient's perspective.

Limitations and Future Research

While many different methods to detect DIF are available, only one method was utilized in this study. Since the presence or absence of DIF is dependent on the technique used, which in turn depends on the specific definition ascribed by the technique, assessing DIF using multiple methods of detection is warranted. Also, this study focused solely on African American and Caucasians cancer patients. To provide a more comprehensive framework for health related QOL pertaining to psychological well-being, mental health data could be incorporated into a study with similar parameters. Further, other subgroups need to be examined, extending the findings beyond the inclusion criteria of this study in an effort to develop a more comprehensive model of the dynamic relationships between the various constructs associated with QOL and well-being from a cross-cultural perspective.

Conclusions

Research has examined the more specific aspects of well-being related to the QOL of chronic illness patients, although there is less data assessing differences at the item level. Our results indicate that item level differences exist between African American and Caucasian chronic illness patients within multiple domains of QOL, suggesting that effective intervention beyond medical treatment must take into account each individual patient's phenomenological interpretation of their situation. Beyond assessing health related QOL in a medical context, clinical conceptualization, treatment planning and aftercare are all impacted when racial subgroups appear to be performing similarly at the group level but differently at the item level.

References

Agresti, A. (1990). *Categorical data analysis.* New York: Wiley.

Ashing-Giwa, K., Padilla, G., Tejero, J., Kraemer, J., Wright, K., Coscarelli, A., et al. (2004). Understanding the breast cancer experience of women: A qualitative study of African American, Asian American, Latina, and Caucasian cancer survivors. *Psycho-Oncology, 13,* 408-428.

Barofsky, I., & Sugarbarker, P.H. (Eds.). (1990). *Quality of Life Assessments in Clinical Trials.* New York: Raven.

Bloor, L.E., Sandler, R.S., Martin, C., Uchino, B.N., & Kinney, A.Y. (2006). Associations between emotional support and health-related quality of life among a population-based sample of blacks and Caucasians. *Journal of Social and Clinical Psychology, 25,* 96-116.

Bynum, M.S., Burton, E., & Best, C. (2007). Racism experiences and psychological functioning in African American college freshmen: Is racial socialization a buffer? *Cultural Diversity & Ethnic Minority Psychology, 13,* 64-71.

Cella, D., & Bonomi, A. (1996). The Functional Assessment of Cancer Therapy (FACT) and Functional Assessment of HIV Infection (FAHI) quality of life measurement system. In B. Spilker (Ed.). *Quality of Life and Pharmacoeconomics in Clinical Trials.* Philadelphia: Lippincott-Raven.

Cella, D.F., Lloyd, S.R., & Wright, B. (1996). Cross-Cultural Instrument Equating: Current Research and Future Directions. *Quality of Life and Pharmacoeconomics in Clinical Trials, 73,* 707-715.

Cella, D.F., McCain, N.L., Peterman, A.H., Mo, F., & Wolen, D. (1998). Development and validation of the Functional Assessment of Human Immunodeficiency Virus Infection (FAHI) quality of life instrument. *Quality of Life Research, 5,* 450-463.

Cella, D.F., Tulsky, D.S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., et al. (1993). The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. *Journal of Clinical Oncology, 11,* 570-579.

Center On Outcomes, Research and Education (CORE). (1997). *Manual of the Functional Assessment Of Chronic Illness Therapy (FACIT) Measurement System (Version 4).* Evanston, IL: David Cella.

Chang, H.-H., Mazzeo, J., Roussos, L. (1995). Detecting DIF for polytomously scored items: An adaptation of the SIBTEST procedure. *Journal of Educational Measurement, 33(3),* 333-353.

Dana, R.H. (1993). *Multicultural Assessment Perspectives for Professional Psychology.* Needham Heights, MA: Simon & Shuster.

Day, S.X. (2004). *Theory and Design in Counseling and Psychotherapy.* Boston: Houghton Miffllin.

Delfino, R.J., Ferrini, R.L., Taylor, T.H., Howe, S., & Anton-Culver, M. (1998). Demographic differences in prostate cancer incidence and stage: An examination of population diversity in California. *American Journal of Preventive Medicine, 14,* 96-102.

Eton, D., Lepore, S., & Helgeson, V. (2001). Early quality of life in patients with localized prostate carcinoma. *Cancer, 92,* 1451-1459.

Fave, A.D. (2006). The impact of subjective experience on the quality of life: A central issue for health professionals. In M. Csikszentmihalyi & I.S. Csikszentmihalyi (Eds.). *A life worth living: Contributions to positive psychology.* New York: Oxford University Press.

Fayers, P.M., & Hays, R.D. (2005). *Assessing quality of life in clinical trials: Methods and Practice.* Oxford: Oxford University Press.

French, A.W., & Miller, T.R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement, 33,* 315-332.

Gallagher, R.M. (2001). Clinical pain research: Challenges across cancer, cultures, and disciplines. *Pain Medicine, 2,* 1-2.

Gonzales, V.A., Martelli, M.F., & Baker, J.M. (2000). Psychological assessment of persons with chronic pain. NeuroRehabilitation, 14, 69-83.

Green, C.R., Anderson, K.O., Baker, T.A., Campbell, L.C., Decker, S., Fillingim, R.B., et al. (2003). The unequal burden of pain: Confronting racial and ethnic disparities in pain. *Pain Medicine, 4,* 277-294.

Hahn, E., & Cella, D. (2003). Health outcomes assessment in vulnerable populations: Measurement challenges and recommendations. *Archive of Physical Medical Rehabilitation, 84,* 35-41.

Hoffman, R.M., Gilliland, J., Elay, W., Stephenson, R.A., Stanford, J.L., et al. (2001). Racial and ethnic differences in advanced-stage prostate cancer: The Prostate Cancer Outcomes Study. *Journal of the National Cancer Institute, 93,* 388-395.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65-70.

Hoshmand, L.T. (2001). Psychotherapy as an Instrument of Culture. In B.D. Slife, R.N. Williams, & S.H. Barlow (Eds.). *Critical Issues in Psychotherapy- Translating New Ideas Into Practice* (pp. 99-113). Thousand Oaks, CA: Sage.

Kim, S.H., Cohen, A.S., Alagoz, C., & Kim, S. (2007). DIF detection and effect size measures for polytomously scored items. *Journal of Educational Measurement, 44,* 93-116.

Krupski, T.L, Fink, A., Kwan, L., Maliski, S., Connor, S.E., Clerkin, B., et al. (2005). Health-related quality of life in low income uninsured men with prostate cancer. *Journal of Health Care for the Poor and Underserved, 16,* 375-390.

Lent, L., Hahn, E., Eremenco, S., Webster, K., & Cella, D. (1999). Using Cross-Cultural Input to Adapt the Functional Assessment of Chronic Illness Therapy (FACIT) Scales. *Acta Oncologica, 38,* 695-702.

Loge, C., Trudeau, E., Marquis, P., Kahrilas, P., Stanghellini, V., Talley, J., et al. (2004). Cross-cultural development and validation of a patient self-administered questionnaire to assess quality of life in upper gastrointestinal disorder: The PAGI-QOL. *Quality of Life Research, 13,* 1751-1762.

Lubeck, D.P., Kim, H., Grossfeld, G., Ray, P., Penson, D.F., Flanders, S.C., et al. (2001). Health related quality of life differences between Black and Caucasian men with prostate cancer: Data from the cancer of the prostate strategic urological research endeavor. *The Journal of Urology, 166,* 2281-2285.

Maliski, S.L., Kwan, L., Orecklin, J.R., Saigal, C.S., & Litwin, M.S. (2005). Predictors of fatigue after treatment for prostate cancer. *Urology, 65,* 101-108.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association, 58,* 690-700.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22,* 719-748.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.). *Frontiers in Econometrics* (pp.105-142). New York: Academic Press.

Mettlin, C.J., Murphy, G.P., Cunningham, M.P., & Menck, H.R. (1997).The National Cancer Data Base report on race, age, and region variations in prostate cancer treatment. *Cancer, 80,* 1261-1266.

Meyers, E. (2007). Why hasn't psychology studied that? *Monitor on Psychology, 38,* 36.

Miller, T.A., & Spray, J.A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement, 30,* 107-122.

Park, C.L. & Adler, N.E. (2003). Coping style as a predictor of health and well-being across the first year of medical school. *Health Psychology, 22,* 627-631.

Penedo, F.J., Dahn, J.R., Shen, B., Schneiderman, N., & Antoni, M.H. (2006). Ethnicity and determinants of quality of life after prostate cancer treatment. *Urology, 67,* 1022-1027.

Roberts, C.M. (2001). Meeting the Needs of Patients With Limited English Proficiency. *The Journal of Medical Practice Management, 17,* 1-5.

Sattler, J.M. (2001). *Assessment Of Children: Cognitive Applications* (4 ed.). San Diego: Jerome M. Sattler Publisher.

Schwartz, D. (2007). Is inequality making us sick? *Monitor on Psychology, 38,* 36.

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/ DIF from group ability differences and detects test bias/ DTF as well as item-bias/ DIF. *Psychometrika, 58,* 159-194.

Shinagawa, S.M. (2000). The excess burden of breast carcinoma in minority and medically underserved communities: Application, research, and redressing institutional racism. *Cancer 88,* 1217-1223.

Shtatland, E.S., Moore, S., & Barton, M.B. (2000, April). Why we need an R2 measure of fit (and not only one) in PROC LOGISTIC and PROC GENMOD. *Proceedings of the 25th annual SAS® users group international conference*, Indianapolis, IN, USA, Paper 256-25.

Skeel, R.T. (1989). Quality of life assessment in cancer clinical trials- It's time to catch up. *Journal of the National Cancer Institute, 81,* 472-473.

Swaminathan, H., & Rogers, H.J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27,* 361-370.

Tchen, N., Bedard, P., Yi, Q-L., Klein, M., Cella, D., Eremenco, S., & Tannock, I.F. (2003). Quality of life and understanding of disease status among cancer patients of different ethnic origin. *British Journal of Cancer, 89,* 641-647.

The Health Institute, New England Medical Center. (1993). *SF-36 Health Survey: Manual & Interpretation Guide.* Boston: John Ware.

Thompson, I., Tangen, C.M., Tolcher, A., Crawford, E.D., Eisenberger, M., & Moinpour, C.M. (2001). Association of African American ethnic background with survival in men with metastatic prostate cancer. *Journal of the National Cancer Institute, 93,* 219-225.

Turk, D.C., & Okifuji, A. (2002). Psychological factors in chronic pain: Evolution and revolution. *Journal of Consulting and Clinical Psychology, 70,* 678-690.

Utsey, S.O., Chae, M.H., Brown, C.F., & Kelly, D. (2002). Effect of ethnic group membership on ethnic identity, race-related stress, and quality of life. *Cultural Diversity & Ethnic Minority Psychology 8,* 366-377.

Vazquez, E. (2007, September/ October). Complications and Emerging Infections: Monitoring is important. *Positively Aware.*

Whaley, A.L., & Davis, K.E. (2007). Cultural Competence and Evidence-Based Practice in Mental Health Services: A Complementary Perspective. *American Psychologist, 62,* 563-574.

Wool, M.S., & Mor, V. (2005). A multidimensional model for understanding cancer pain. *Cancer Investigation, 23,* 727-734.

Zickler, P. (2009). When the body suffers, the mind suffers. *Monitor on Psychology, 40,* 31-34.

Zumbo, B.D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores.* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.