

Evaluation of the Missoula-VITAS Quality of Life Index—Revised: Research Tool or Clinical Tool?

CAROLYN E. SCHWARTZ, Sc.D.,¹⁻⁴ MELANIE P. MERRIMAN, Ph.D., M.B.A.,^{5,6}
GEORGE REED, Ph.D.,³ and IRA BYOCK, M.D.⁷

ABSTRACT

Background: Quality of life (QOL) is a central outcome measure in caring for seriously ill patients. The Missoula-VITAS Quality of Life Index (MVQOLI) is a 25-item patient-centered index that weights each of five QOL dimensions (symptoms, function, interpersonal, well-being, transcendence) by its importance to the respondent. The measure has been used to assess QOL for hospice patients, and has been found to be somewhat complex to use and analyze.

Objective: This study aimed to simplify the measure, and evaluate the reliability and validity of a revised version as either a research or clinical tool (i.e., “psychometric” versus “clinimetric”).

Design: Two data collection efforts are described. The psychometric study collected QOL data from 175 patients at baseline, 3–5 days, and 21 days later. The implementation study evaluated the feasibility and utility of the MVQOLI-R during over six weeks of use.

Setting/subjects: End-stage renal patients on dialysis, hospice, or long-term care patients participated in the psychometric study. The implementation study was done in hospice, home health, and palliative care settings.

Measurements: The MVQOLI-R and the Memorial Symptom Assessment Scale.

Results: The psychometric and implementation studies suggest that the MVQOLI-R performs well as a clinical tool but is not powerful as an outcome research instrument. The MVQOLI-R has the heterogeneous structure of clinimetric tools, and demonstrated both relevance and responsiveness. Additionally, in a clinical setting the MVQOLI-R was useful therapeutically for stimulating communication about the psychosocial and spiritual issues important to the tasks of life completion and life closure.

Conclusions: The MVQOLI-R has clinical utility as a patient QOL assessment tool and may have therapeutic utility as a tool for fostering discussion among patients and their clinicians, as well as for helping patients identify sources of suffering and opportunities during this time in their lives.

¹QualityMetric Incorporated, Waltham, Massachusetts.

²Health Assessment Lab, Waltham, Massachusetts.

³Department of Medicine, University of Massachusetts Medical School, Worcester, Massachusetts.

⁴DeltaQuest Foundation, Inc., Concord, Massachusetts.

⁵Touchstone Consulting, Inc. North Bay Village, Florida.

⁶University of Miami School of Nursing, Miami, Florida.

⁷The Palliative Care Service, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire.

INTRODUCTION

THE FIELD OF RESEARCH into palliative and end-of-life (EOL) care has blossomed in recent years with the emergence and increased utilization of standardized tools for assessing quality of life (QOL) and parameters of quality of care through the EOL. These tools can focus on an array of relevant dimensions. QOL tools for EOL care research and practice begin with the standard World Health Organization conceptualization of QOL: physical, social, and emotional well-being.¹ They may also expand into more pragmatic aspects of quality care, for example, symptom management,² relations with the health care system,³ or may focus on spiritual aspects of well-being.^{4,5} When there is no hope of cure, however, these domains may shift their meaning for the respondent. For example, the domain of social well-being may change from meaning how much the respondent is able to engage in social activities deemed "normal" for a well person, to reflecting how much of a burden the respondent feels he or she is on the family, friends, and health care providers in their social network.

In addition to changing their meaning, the proportional relevance of these QOL domains may increase or decrease as physical health and function change. These shifts impact the approach QOL researchers need to consider in construct validation. For example, a standard approach to validating a measure of QOL in general, and health-related QOL in particular, is to evaluate its association with an objective clinical measure of health. For patients whose illnesses are progressive and incurable, however, nonphysical domains of QOL may become more important,⁶ so a measure of QOL in far-advanced stages of illness might not be expected to be as responsive to clinical change in health status as in a healthier population.

In response to the need for a QOL measure that assesses the individual experience of people nearing the EOL, Byock and Merriman (7) created the Missoula-VITAS Quality of Life Index (MVQOLI). This 25-item patient-centered tool was aimed at measuring adaptation to and integration of their physical and functional decline, as well as attainment of tasks of life completion and life closure.⁸ The tool addressed five QOL domains that are clinically relevant for EOL care (i.e., symptom control, function, interpersonal issues, well-being, and transcendence). It also contained items

to measure satisfaction and importance within each domain, and the summary score was a weighted sum that reflected the importance of each dimension to the patient. Data from the initial validation study⁷ supported the idea that subjective QOL and objective function are no longer linearly related among seriously ill patients. This finding challenges the prevailing assumption that QOL varies in proportion to function, especially activities of daily living and independence.

An individual patient's scores on the MVQOLI are most easily interpreted using a bar graph called the QOL Profile. The domain scores range from -30 to $+30$ and therefore produce bars of varying length both above and below the x-axis. The length of each bar corresponds to the overall impact of the domain on QOL; the direction of the bar shows whether the domain detracts from or enhances quality of life. Professional caregivers can use the QOL Profile to identify foci for interventions to alleviate suffering and improve QOL for an individual patient. From a clinical and conceptual perspective, this tool has great appeal.

From a psychometric perspective, however, the tool had not been studied comprehensively. First, the initial psychometric validation provided suggestive evidence for the internal consistency reliability ($\alpha = 0.77$), and broad construct validity of the total score ($r = 0.43$ with global QOL), but did not address more specific details about characteristics important for research purposes (e.g., the item distributions, the empirical factor structure, the test-retest reliability of the tool, or the responsiveness of the tool to clinically important change). Second, experience with the tool over the past 8 years has highlighted two issues that limit practical use. It is somewhat long (i.e., 25 items) for the seriously ill patients for whom it is intended. Because the 5 importance items are related to weighting dimensions for the total score, we sought to evaluate how much information and variability the importance items contributed to the total score. A growing body of research⁹⁻¹² supports Nunnally and Burnstein's¹³ statement that weighted indices correlate quite highly with unweighted indices, suggesting the weighting items may be unnecessary in the MVQOLI-R.

A second limitation of the original tool was that its format proved difficult for very sick people to complete. Some items were single statements to which the respondent indicated agreement on a Likert scale (e.g., "My symptoms are adequately

controlled.”) Other items presented two contrasting statements for a respondent to indicate which statement better reflected his/her personal situation. The contrasting statement format was used inconsistently across dimensions, and statements varied in the extent to which they represented direct opposites of one another, (e.g., “I feel x ” and “I feel *not-x*” versus “I feel x ” and “I feel y ”). These sources of variation may have increased the cognitive complexity of the respondent’s task, and potentially added random error attributable to method variance. We thus sought to simplify and abbreviate the MVQOLI, and to evaluate more rigorously the psychometric characteristics of the revised tool.

This research also sought to explore the utility of the MVQOLI-R as a research and/or clinical tool (i.e., a psychometric versus clinimetric instrument). Experience over the past decade has suggested that clinicians deem the MVQOLI useful for helping them understand the QOL of very

sick patients and explore with patients issues that may contribute to or detract from their experienced QOL, and thus that the tool has clinical utility. Its usefulness in outcomes research is yet to be reported.

The term “clinimetrics,” originally described by Feinstein,¹⁴ refers to ratings, scales, indices, and instruments used to measure clinical phenomena that cannot be measured with the customary methods of laboratory data. Although both psychometric and clinimetric tools must be deemed relevant to the populations for which they are intended (i.e., ecologically valid) and responsive to clinically important change, clinimetric tools are distinct from psychometric tools in important ways (Table 1). These include the goal of the tool, the ways scores are aggregated, the homogeneity of the items, and the appropriate way of demonstrating an instrument’s validity for its intended purpose.¹⁵ Clinimetric tools evaluate a phenomenon such as QOL by assess-

TABLE 1. PSYCHOMETRIC VERSUS CLINIMETRIC DISTINCTIONS FOR QOL TOOLS

<i>Issue</i>	<i>Psychometric</i>	<i>Clinimetric</i>
Goal of tool	To measure an underlying latent variable.	To identify a clinical phenomenon.
Desired item characteristics	Unidimensional and normally distributed.	Clinically relevant to the phenomenon being measured. Non-normal distribution or rare endorsement of item not a problem.
Indicators of reliability	Internal consistency and stability when health status has not changed.	Stability when health status has not changed.
Indicators of validity	Construct, content, face, ecological, and discriminant validity; responsiveness to clinically significant change in QOL.	Face, content, and ecological validity and responsiveness to clinically significant change in QOL.
Use in clinical practice	Can monitor and screen QOL to provide feedback for use in healthcare encounter; Improve provider–patient communication. Screen for provider–patient communication, and monitor quality of care.	Can screen for clinical phenomenon of interest and provide tool for provider–patient discussion. May be good predictor of outcomes.
Use in clinical research	Can document QOL in cross-sectional and longitudinal studies, where probability-based inference is desired.	Can be used to identify people who exhibit the clinical phenomenon of interest. May also be useful as an “intervention” to identify issues impacting quality of life and stimulating discussion between patient and providers.

QOL, quality of life.

ing the occurrence of multiple, heterogeneous attributes, such as symptoms that affect the phenomenon. They may create scores by summing a set of ratings or frequency/proportion data.¹⁵ In contrast, psychometric tools measure an underlying construct by assessing attributes that are relatively homogenous¹⁵ and are affected by changes in the latent construct. Psychometric indices traditionally aggregate scores using summative approaches that are derived from empirical studies of the tool's structure (e.g., adding items on a unidimensional subscale to generate a subscale score). Finally, the validity of a clinimetric tool is primarily based on its correlation with and responsiveness to a criterion (i.e., external gold standard assessment of the same or related phenomenon). In contrast, psychometric tools focus on a number of indices of construct validity that specify the structure of the tool, the theoretical basis of the construct,¹⁶ its ability to predict differences between known groups, and its responsiveness to related, but not overlapping, constructs.¹⁷

Fayers and Hand^{18,19} note that standard psychometric techniques are not appropriate to validating clinimetric tools because psychometric theory assumes that items are effect indicators, meaning in this instance that questionnaire items would reflect changes in QOL. Instead, items in clinimetric measures are causal indicators, meaning they assess experiences, events or symptoms that *cause* changes in QOL. Fayers and Hand^{18,19} provide examples of how using psychometric methods on clinimetric tools can lead to inaccurate conclusions about the validity or lack of validity of the tool.

An example from the context of palliative care might prove useful here. Symptoms such as constipation may create psychological distress, anxiety, or depression, and these psychological or social issues may lead to other physical or functional outcomes. Thus, symptoms (often assessed in clinimetric tools) are causal indicators rather than effect indicators, and utilizing psychometric analytic methods (e.g., looking at item distributions, factor structure, and construct validity correlations) is not appropriate for clinimetric tools. Consequently, if one is working with a clinimetric tool, standard psychometric evaluations may be misleading. The present work evaluates the MVQOLI-R from both psychometric and clinimetric perspectives.

METHODS

This study included two data collection efforts. The first effort was quantitative, focusing on evaluating the psychometric characteristics of the MVQOLI-R. The second effort was qualitative in which the tool was implemented in clinical sites, and the perceived clinical usefulness of the instrument was investigated using semi-structured interviews.

The psychometric study

Participants. We sought to include patients who were either (1) seriously ill, but relatively stable or (2) seriously ill and deteriorating. Accordingly, eligible participants were likely to see their provider within 3 weeks of study entry, were over the age of 18, not pregnant, and had either progressive, chronic, and life-threatening illness, or a prognosis between 6 weeks and 3 years. We recruited patients from three types of sites: dialysis clinics, hospices, and long-term care facilities. Potential participants were screened for dementia and excluded if they scored lower than 25 on the Folstein Mini-Mental Status Exam.²⁰ Written informed consent was obtained. Participants were paid \$10 for each interview.

Design and sites. Data were collected up to three times over 3 weeks: baseline (day 0), and subsequently at 3 to 5 days, and 21 days later. All data were collected using semistructured interviews with trained research assistants. Interviews were conducted at the most convenient location for the participant. End-stage renal patients were interviewed at the clinic during dialysis, hospice patients at home or in a residential hospice, and long-term care patients at the nursing home or rehabilitation hospital, or at home if they were discharged from the long-term care facility during the study follow-up period.

Data were collected from 12 sites in Massachusetts, New York, and Florida. They included: five dialysis clinics (participant numbers in brackets): Hahneman [91] and University [6], in Worcester, Massachusetts; North County [23], in Fitchberg, Massachusetts; Blackstone Valley [12], in Milford, Massachusetts; Hopedale [4], in Hopedale, Massachusetts; five hospices: (UMass Memorial [12] and Visiting Nurses Association Care Network [7], of Worcester, Massachusetts; Harrington [1], of Southbridge, Massachusetts;

Hospice of the North Shore [1], of Danvers, Massachusetts; and VITAS Healthcare Corporation [2], Miami, FL; and two rehabilitation hospitals (Spaulding [1], of Boston, Massachusetts; and Fairlawn [5], of Worcester, Massachusetts). A total of 184 participants were recruited, of whom 167 provided complete, useable data for all three time points, 175 had complete data at the first and second time point, 14 had complete data from two interviews, and three had complete data from only one interview.

Hypotheses. We hypothesized that the MVQOLI-R would demonstrate: (1) high test-retest reliability over three to five days in patients with stable health status; (2) a factor structure consistent with subscales for symptoms, function, interpersonal, well-being, and transcendence; (3) high internal consistency within subscales and for the total score; (4) convergent and divergent validity, as evidenced by moderately strong associations with measures of psychological well-being, global symptom-related distress, and global QOL; but weak correlations with mood and total symptom burden; (5) high relevance or ecological validity, evidenced by high participant ratings on an item regarding the pertinence of the MVQOLI-R questions to their QOL; and (6) responsiveness to clinically significant change in QOL, as evidenced by significant associations between MVQOLI-R change scores and global symptom-related distress over 21 days, independent of changes in mood.

Measures. In addition to the above-described MVQOLI-R, standardized tools were used to assess health status, psychological well-being and mood. Patient-reported health status was measured using the Memorial Symptom Assessment Scale (2). This measure asks respondents to rate the frequency, severity, and bothersomeness over the past week of 33 symptoms, and yields summary scores of psychological, physical, global symptom-related distress, and total symptom scores. It has documented reliability and validity, and has been used in QOL studies of patients with breast cancer and patients with acquired immune deficiency syndrome (AIDS).

Psychological well-being was measured by the Ryff Psychological Well-Being measure-short-form.²¹ This 17-item standardized measure has documented reliability and validity, and contains subscales for purpose in life, personal growth, au-

tonomy, environmental mastery, positive relations with others, and self-acceptance.

Mood was measured by the Brief Profile of Mood States (Brief POMS).²² This 17-item scale is an abbreviated version of the 58-item POMS.²³ It has documented internal consistency reliability ($r = 0.91$), and correlates highly ($r = 0.93$) with the longer version of the measure.

Data were also collected on participant sociodemographic characteristics, such as date of birth, gender, ethnicity, marital status, education, religion, and importance of spiritual practice.

Statistical analysis

Data analyses evaluated the item distributions, as well as the reliability and validity of the MVQOLI-R.

Reliability analyses. Cronbach α ²⁴ was used to assess internal consistency in subscales and in the total score. Test-retest stability was computed using the intraclass correlation coefficient²⁵ comparing scores from baseline and 3–5 day post-baseline.

Validity analyses. Construct validity was assessed with the Pearson correlation coefficient to look at associations between MVQOLI-R total score and subscale measures from validated measures of psychological well-being, global symptom-related distress, global QOL; mood and total symptom burden. Effect size for these analyses was defined using Cohen's²⁶ criteria, where small effects would be a correlation coefficient between 0.10 and 0.29, a moderate effect would be between 0.30–0.49, and a large effect would be a correlation coefficient larger than 0.50.

Responsiveness was assessed using regression models to estimate the association of MVQOLI-R scores with global symptom-related distress scores on the MSAS. Using data collected at baseline and 21 days, we used a mixed model approach²⁷ with global symptom-related distress as the outcome, total MVQOLI-R and mood as predictors, and respondent as a random effect. The model simultaneously accounts for the association between MVQOLI-R and global distress scores for all respondents (cross-sectional effects) and the association between changes in MVQOLI-R and global distress scores over time for individual respondents (longitudinal effects.) In order to estimate whether the cross-sectional

and longitudinal effects are different, we separate these components in the model using procedures described by Neuhaus and Kalbfleisch²⁸ and most recently Begg and Parides.²⁹ Two coefficients result: one for the mean value that represents the cross-sectional effect; and another, termed the residual factor, that represents the within-subject longitudinal effect.

The Implementation Study

Participants and sites. The implementation study was conducted with staff and patients in three different settings at the Visiting Nurse Service of New York (VNS-NY): hospice, home health, and palliative care. Eligible patients had advanced chronic illness with an expected prognosis of 3 months to 3 years. Additionally, patients had to be interacting regularly with staff (i.e., daily, weekly, monthly, or quarterly), and had to be capable of providing answers for the MVQOLI-R, either by completing the instrument on their own or by answering when staff read the items. Managers at each site identified 3–5 staff to participate in the pilot study, and each staff person was asked to utilize the MVQOLI-R with 5–10 patients.

Procedures and measure. The implementation study began with onsite meetings at the VNS-NY where we worked with provider staff to develop procedures for integrating administration of the MVQOLI-R into their current clinical work flow. Patient tracking forms and other materials were prepared and used to train patient care staff who would administer the MVQOLI-R and data entry staff who would use the computer scoring tools and produce the graphic QOL profile for each patient. Staff then used the instrument with their assigned patients over the next 6 weeks, and kept notes on their experience with the tool on the tracking forms provided.

At the end of the 6 weeks, interviews were conducted with staff who had been using the tool, the data entry staff, and the executive leaders of the programs we were working with. The Implementation Feedback Survey asked about: (1) challenges in following the implementation procedures and suggestions for improvement, (2) patient response to the instrument, and (3) usefulness of the MVQOLI-R data (including the graphic profile) for patient care. A total of 10 interviews, reflecting clinicians' experience with 40 patients, were completed.

Data analysis. The qualitative data were reviewed by two investigators (M.P.M. and K.E.), and analyzed for recurring themes. In particular, the qualitative analysis aimed at: (1) identifying facilitators and barriers for integrating the MVQOLI-R into current patient care processes in multiple health care delivery settings; (2) assessing staff and management satisfaction with MVQOLI-R administration procedures; and (3) determining how useful the MVQOLI-R information is to the staff in clinical practice.

RESULTS

The psychometric study

Participants. Patient accrual was implemented from June 2001 through September 2002. We recruited patients with 144 end-stage renal disease on dialysis patients and 31 hospice or long-term care patients (hereafter referred to as "hospice"). Participants had a mean age of 66.3 (standard deviation [SD] = 14.6), and tended to be Caucasian, high school graduates, married, and Catholic (Table 2). The majority indicated that spiritual practice was very important to them.

Item distributions

Frequency distributions on the items demonstrated significant skewness (i.e., scores grouped at one end of the scale) (Table 3). This tendency was exacerbated in four of the five importance items, where respondents overwhelmingly (87%–94%) endorsed strong agreement with the importance of a domain (Table 3). These four domain items all contained the word "important" in ways that ostensibly increased demand characteristics. Accordingly, the importance items were eliminated from the remainder of the psychometric analyses. Skewness was high in several other items as well. There were 7 of 20 items with 50% or more answers at one end of the scale.

Revising the MVQOLI-R scoring protocol

Our decision to drop the importance items required a revision of the scoring algorithm for the MVQOLI-R. Rather than using importance items to weight the severity scores in generating a total score, a simplified algorithm added satisfaction and severity items to generate subscale scores. Additionally, we standardized the scores

TABLE 2. BASELINE PSYCHOMETRIC STUDY PATIENT CHARACTERISTICS ($n = 175$)

	n	%
Group		
Hospice	31	17.7
Dialysis	144	82.3
Gender		
Male	98	59.4
Female	67	40.6
Race/ethnicity		
Caucasian	152	86.9
African American	13	7.4
Asian/Pacific Islander	1	0.6
Hispanic	4	2.3
Other	5	2.9
Education		
1–8 grade	13	7.4
9–11 grade	29	16.6
High school graduate	57	33.6
Some college	45	25.7
College degree (4 years)	19	10.9
Advanced degree	12	6.9
Marital status		
Single, never married	23	13.1
Married or living with significant other	93	53.1
Separated	8	4.6
Divorced	17	9.7
Widowed	34	19.4
Importance of spiritual practice		
A great deal	110	62.9
Somewhat	41	23.4
A little	14	8.0
Not at all	10	5.7
Religious preference		
Catholic	103	61.3
Other Christian denomination	52	31.0
Jewish	3	1.8
Buddhist	1	0.6
Other	9	5.3

so that they ranged from 1 to 5, rather than having the satisfaction items range from -2 to $+2$, and the severity items range from -4 to $+4$. This new scoring algorithm led to scores that were more highly correlated ($r = 0.94$ – 0.97) than with the more complicated scoring algorithms. This algorithm reverses polarity of the negatively worded items (#2, 6, 7, 16, 19, 21, and 24). As a result, higher scores on the MVQOLI-R reflect better QOL.

Factor analysis

Factor analysis with a Promax rotation was conducted on the all 25 items, and on the item pool after dropping the importance items. Factor

analysis results supported the decision to drop the importance items because the importance scores grouped together, loading on the same factor. Eigenvalues indicated use of five factors was appropriate and 52% of the variation was explained by the first five factors (Table 3). An examination of the factor loadings in the 20-item analysis did not support the theoretical structure of the instrument. In general, the highest loading items within each factor were not items that were hypothesized to load together. This anomaly led us to wonder whether the MVQOLI-R items were not effect indicators, but rather were causal indicators (i.e., items reflected factors that changed QOL rather than items that are a sign of QOL),¹⁸ and thus not appropriate for correlational analyses.

Reliability

α Reliability estimates for the four subscale items within each *a priori* domain ranged from 0.23 to 0.70 (Table 4). Only the Symptom domain appeared to be measuring a unidimensional concept. Closer examination revealed that the α coefficient could be increased by deleting one or more items from a subscale, but that the resulting coefficient would still not be in a range normally considered adequate for a psychometric tool. An examination of test–retest stability over 3–5 days demonstrated high stability in responses over time: the intraclass correlation coefficient for the subscales ranged from 0.59–0.70, and was 0.77 for the Total score (Table 4).

Validity

Relevance. The MVQOLI-R demonstrated relevance (i.e., ecological validity) in these samples: 89% of the study participants reported that the tool was strongly or moderately relevant to their QOL (mean = 1.8 on 5-point scale where 1 = strongly agree; SD = 0.87).

Construct validity. Because the *a priori* subscales were not supported by the factor analytic results, the construct validity of the MVQOLI-R was assessed only for the total score. Patterns of association suggested that the MVQOLI-R total score was moderately correlated with global QOL, the total score and most subscales of the Ryff Psychological Well-Being measure, global symptom-related distress and total symptom burden on the Memorial Symptom Assessment Scale, and mood

TABLE 3. DESCRIPTIVE STATISTICS ON ITEMS OF MVQOL-1-R ($n = 175$)^a

Subscale	Item content	Item distributions		Factor loading for 20-item tool (Promax rotation)				
		Disagree (score = 1 out of 5)	Agree (score = 5 out of 5)	1	2	3	4	5
		n (%)	n (%)					
Symptoms	1. My symptoms are adequately controlled.	5 (2.84)	71 (40.34)	0.58	-0.07	-0.08	0.10	0.36
	2. I feel sick all the time.	91 (51.70)	6 (3.41)	0.77	0.06	0.08	-0.06	-0.04
	3. I accept my symptoms as a fact of life.	6 (3.41)	117 (66.48)	0.19	-0.10	-0.04	-0.09	0.80
	4. I am satisfied with the current control of my symptoms.	15 (8.52)	86 (48.86)	0.73	-0.06	-0.00	-0.06	0.24
	5. Physical discomfort overshadows any opportunity for enjoyment.	55 (30.39)	26 (14.36)	—	—	—	—	—
Function	6. I am dependent on others for personal care.	84 (47.73)	28 (15.91)	-0.01	-0.03	-0.05	0.84	-0.17
	7. I am no longer able to do many of the things I like to do.	20 (11.36)	81 (46.02)	0.16	-0.04	0.02	0.68	-0.29
	8. I am satisfied with my ability to take care of my basic needs.	9 (5.11)	93 (52.84)	-0.07	-0.14	-0.12	0.64	0.24
	9. I accept the fact I can not do many of the things that I used to do.	11 (6.25)	97 (55.11)	0.13	-0.10	-0.06	-0.21	0.74
	10. My contentment with life depends upon being active and being independent in my personal care.	5 (2.76)	105 (58.01)	—	—	—	—	—
Interpersonal	11. I have recently been able to say important things to the people close to me. ^b	14 (7.95)	81 (46.02)	-0.25	0.37	0.16	-0.08	0.32
	12. I feel closer to others in my life now than I did before my illness.	32 (18.18)	67 (38.07)	-0.03	0.73	0.01	-0.26	-0.06
	13. In general, these days I am satisfied with relationships with family and friends.	3 (1.70)	119 (67.61)	-0.02	0.01	0.54	-0.01	-0.04
	14. At present, I spend as much time as I want to with family and friends. ^b	17 (9.66)	87 (49.43)	-0.37	0.16	0.39	0.27	0.16
	15. It is important to me to have close personal relationships.	3 (1.67)	124 (68.89)	—	—	—	—	—
Well-being	16. My affairs are not in order; I am worried that many things are unresolved.	66 (37.50)	26 (14.77)	0.35	-0.01	0.46	0.03	-0.02
	17. If I were to die suddenly today, I would feel prepared to leave this life. ^b	26 (14.77)	91 (51.70)	0.15	-0.08	0.84	-0.13	-0.13
	18. I am more satisfied with myself as a person now than I was before my illness.	49 (27.84)	33 (18.75)	0.32	0.65	-0.09	-0.08	-0.23

TABLE 3. DESCRIPTIVE STATISTICS ON ITEMS OF MVQOL-1-R ($n = 175$)^a (CONT'D)

Subscale	Item content	Item distributions		Factor loading for 20-item tool (Promax rotation)				
		Disagree (score = 1 out of 5)	Agree (score = 5 out of 5)	1	2	3	4	5
		n (%)	n (%)					
	19. The longer I am ill, the more I worry about things "getting out of control." ^b	63 (35.80)	36 (20.45)	0.55	0.13	0.22	0.18	0.01
	20. It is important to me to be at peace with myself.	0 (0)	138 (76.24)	—	—	—	—	—
Transcendence	21. I feel more disconnected from all things now than I did before my illness.	62 (35.23)	32 (18.18)	0.47	0.12	0.03	0.26	0.20
	22. I have a better sense of meaning in my life now than I have had in the past. ^b	27 (15.34)	59 (33.52)	-0.04	0.75	-0.24	0.13	-0.04
	23. I am comfortable with the thought of my own death. ^b	33 (18.75)	71 (40.34)	0.00	-0.19	0.78	-0.01	0.05
	24. Life has lost all value for me; every day is a burden.	123 (68.89)	5 (2.84)	0.16	0.38	0.01	0.14	0.36
	25. It is important to me to feel that my life has meaning.	0 (0)	131 (72.38)	—	—	—	—	—
	Eigenvalue	—	—	3.79	1.87	1.80	1.50	1.41
	% Explained variance	—	—	19.0%	9.3%	9.0%	7.5%	7.0%

^aImportance items are in shaded rows. These items had highly skewed distributions and were thus dropped from subsequent psychometric analyses.

^bItems identified as hard to understand or difficult to answer by patients or staff.

(Table 5). This pattern did not support our convergent and divergent validity hypotheses, but did suggest that the tool tapped a phenomenon related to symptom-related QOL and psychological well-being. There was a small positive association with age ($r = 0.21$, $p < 0.01$), but no association with gender or education ($p = 0.23$ and 0.58 , respectively; data not shown).

Responsiveness. The original conceptualization of the MVQOLI was that QOL for individuals with advanced incurable illness is less contingent on physical and functional status, and more related to the patient's adaptation to and integration of their physical and functional decline, and tasks of life completion and closure. If we consider the MVQOLI-R as a clinimetric tool, then we would hypothesize that the total score and individual subscales would be related to change on

the global symptom-related distress subscale of the Memorial Symptom Assessment Scale, after adjusting for baseline mood and changes in mood. In other words, we would expect that the MVQOLI-R scores are associated with reduced QOL (as measured by global symptom-related distress). This hypothesis was supported by random effects modeling, controlling for baseline and change in mood. We found that the total MVQOLI score had an association with the global symptom-related distress score, both cross-sectionally and longitudinally, after adjusting for mood (Table 6; Fig. 1). The between-subjects (mean factor) coefficient and within-subjects (residual factor) coefficient are similar (-0.023 and -0.022 , respectively), implying that global distress decreases with increasing MVQOLI-R total score, both across individuals and within individuals over time.

TABLE 4. DESCRIPTIVE STATISTICS OF MVQOL-I-R SUBSCALES AND TOTAL SCORE ($n = 175$)

	At baseline							Test-retest stability (95% CI)
	Mean (SD)	Median	Skewness	Kurtosis	Possible range	Respondent range	Internal consistency	
Symptoms	16.73 (3.26)	17	-1.46	4.96	4-20	5-20	0.70	0.64 (0.55, 0.73)
Function	14.04 (3.28)	14	-0.17	2.53	4-20	5-20	0.49	0.59 (0.48, 0.68)
Interpersonal	16.11 (2.95)	16	-0.75	3.31	4-20	6-20	0.41	0.61 (0.51, 0.70)
Well-being	13.53 (3.88)	14	-0.54	2.76	4-20	4-20	0.48	0.70 (0.62, 0.77)
Transcendence	14.95 (3.12)	15	-0.55	3.40	4-20	4-20	0.23	0.68 (0.59, 0.76)
Total score	75.36 (10.64)	76	-0.50	3.64	20-100	33-97	0.71	0.77 (0.70, 0.83)

SD, standard deviation; CI, confidence interval.

Using similar models for each subscale we find similar results and some evidence that these particular subscales separate into those that provide stronger cross-sectional effects and some longitudinal effect (symptom, function, well-being), and those that provide stronger longitudinal effects but less cross-sectional effects (interpersonal and transcendence). We view these subscale models as hypothesis generating. All significant associations were in the expected direction: lower levels of MVQOLI-R indicators of QOL were associated with worse global symptom-related distress.

Clinical tool or research tool?

The factor structure of the MVQOLI-R on the samples included in this study did not support the theoretical structure of the instrument. In general, the highest loading items within each factor were not items that were hypothesized to load together. Similarly, the internal consistency analyses did not support the unidimensionality of the construct. The tool did, however, exhibit good test-retest reliability in the 3-5 day follow-up. Additionally, the total score correlated in the expected direction and magni-

TABLE 5. CONSTRUCT VALIDITY CORRELATIONS

Scale or variable	<i>r with MVQOLI-R total score</i>
Ryff Happiness Scale	
Positive relation to others	0.36**
Purpose in life	0.006
Personal growth	0.27**
Environmental mastery	0.63**
Autonomy	0.27*
Acceptance	0.40**
Total Ryff score	0.47**
Memorial Symptom Assessment Scales (MSAS)	
Global Distress	-0.42**
Total MSAS	-0.35**
Mood Status	
POMS	0.46**
Global QOL	0.35**
Age	0.21**

* $p < 0.05$ for test the null hypothesis: $\rho = 0$; ** $p < 0.01$.
POMS, Profile of Mood States; QOL, quality of life.

TABLE 6. RANDOM EFFECTS MODELS FOR ASSESSING CLINICAL RESPONSIVENESS

MVQOLI-R Subscale	Parameter ^a	Coefficient	95% Confidence Interval	p value
Total score	Mean total score factor	- 0.023	- 0.034, - 0.013	0.000
	Residual total score factor	- 0.022	- 0.035, - 0.0077	0.002
	Mean mood factor	- 0.032	- 0.058, - 0.0063	0.015
	Residual mood factor	- 0.034	- 0.062, - 0.0052	0.021
	Constant	4.32	3.64, 4.997	0.000
Symptoms	Mean symptom factor	- 0.09	- 0.12, - 0.06	0.000
	Residual symptom factor	- 0.04	- 0.08, - 0.00	0.04
Function	Mean function factor	- 0.07	- 0.10, - 0.04	0.000
	Residual function factor	- 0.03	- 0.07, 0.01	0.184
Interpersonal	Mean interpersonal factor	0.01	- 0.02, 0.05	0.478
	Residual interpersonal factor	- 0.03	- 0.07, 0.01	0.105
Well-Being	Mean well-being factor	- 0.05	- 0.08, - 0.03	0.000
	Residual well-being factor	- 0.02	- 0.06, 0.01	0.167
Transcendence	Mean transcendence factor	- 0.02	- 0.06, 0.01	0.210
	Residual transcendence factor	- 0.04	- 0.08, 0.00	0.073

^aThe coefficient for the mean value represents the cross-sectional effect and the coefficient for the residual factor (subject's value - mean) represents the longitudinal effect.

tude with the selected indicator of QOL in advanced illness (i.e., global symptom-related distress), and demonstrated responsiveness to clinically important change. This impact reflected different components of the MVQOLI-R cross-sectionally and longitudinally, suggesting that all of the subscales play a substantive role in elucidating the causal indicators of QOL in the context of progressive decline. We thus conclude that the MVQOLI-R may be more adequately characterized as a clinical tool than a psychometric research tool. The next phase of this research thus focused on evaluating its perceived utility as a clinical tool.

Implementation study results

The implementation pilot study provided valuable information about the use of the MVQOLIR, and suggestions for how using the MVQOLI-R and its implementation might be improved. The following emerged as the major benefits of using the MVQOLI-R for this group of caregivers, and indications for preparations necessary to enhance its clinical utility.

- *The tool opened the door for discussion.* The tool provided a framework and language for discussion of psychosocial and spiritual issues that may otherwise be awkward for both patients and pro-

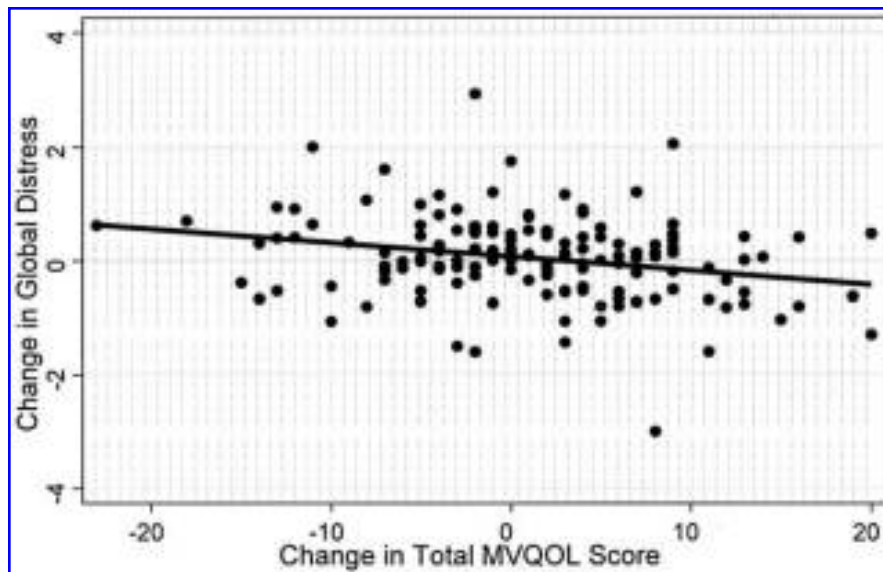


FIG. 1. Responsiveness of the MVQOLI-R. The total score on the MVQOLI-R had an impact on global symptom-related distress, both cross-sectionally and longitudinally, after adjusting for mood. This supports the hypothesis that the tool is clinimetric, such that the items and subscales cause changes in quality of life (QOL) in advanced illness, defined as global symptom-related distress.

professional caregivers to address. Use of the tool stimulated in-depth discussions that enhanced caregivers' understanding of patient concerns, even if they had previously been aware that the concerns existed. It was noted, however, that the tool did not work for very sick patients. In the palliative care setting, a much larger percentage of patients were able to use the tool than in the hospice setting, where 50% to 75% of patients were excluded, primarily because of dementia or other cognitive deficits. This suggests that the tool may be less useful in patients with prognosis of only a few weeks or very poor performance status, but quite useful in working with patients at earlier stages of illness.

- *The tool facilitated holistic, collaborative care.* The MVQOLI-R facilitated more holistic care by providing information about patient needs related to emotional, social and spiritual concerns. It provided individual, patient-centered data to which caregivers could respond, in much the same way that a physical assessment identifies the need for pharmacologic or other medical intervention. Staff also reported that using the MVQOLI-R increased collaboration between psychosocial, spiritual, and medical caregivers, generally providing more of a role for psychosocial and spiritual caregivers.

- *Organizational support is needed to maximize the tool's impact.* Mechanisms for sharing the QOL

information with other caregivers need to be developed. This infrastructure would support the use of the MVQOLI-R. As one leader in the palliative care setting suggested, the MVQOLI-R might provide the core of a "referral report" to physicians who had referred patients to the service. Creating an infrastructure to support the use of the tool would facilitate this core function.

- *Practice makes perfect.* Staff noted that the tool was easier to use and generated better information on a second administration. It thus appears to be important to utilize repeated administrations of the MVQOLI-R over time. This observation also underlines the importance of training in all aspects of use of the MVQOLI-R to help staff to become familiar and more comfortable with administering the tool. This kind of assessment is very different from others to which staff is accustomed. We observed that staff seemed unsure about how to use the information.

Staff also needed coaching in the use of the information to accomplish their care objectives. While staff found the MVQOLI-R information to be useful, they did not utilize the graphic QOL profile to any significant degree. This profile has previously been the primary mechanism for sharing the results of the MVQOLI with other care team members.

- *Items require further refinement.* Finally, the pilot study underscored findings from the psycho-

metric study: there is a need to continue to refine several items to make them more understandable and acceptable to patients, and perhaps to caregivers. Items identified as difficult to understand by patients or staff are noted in Table 3. It will be important to engage in iterative cognitive testing of MVQOLI items.³⁰

DISCUSSION

Results from the psychometric and implementation studies suggest that the MVQOLI-R is not psychometrically robust but performs well in use as a clinical tool. The measure appears to have the heterogeneous structure of clinimetric tools, and demonstrated both relevance and responsiveness. Additionally, the tool was reported to be useful in a clinical setting for stimulating communication about the psychosocial and spiritual issues that comprise many of the tasks of life completion and life closure.

The evidence concerning responsiveness of MVQOLI-R scores to changes in global distress (Memorial Symptom Assessment Scale) suggest that differences in distress between individuals are associated primarily with the symptom, function, and well-being domains, while changes in global distress for one individual over time are associated primarily with the symptom and transcendence domains. The magnitude of the mean and residual coefficients are similar for all domains, however, and thus the model supports the interpretation that all five domains work together to reveal a detectable impact on the patient's experience of illness burden.

This study provided important insights into the application of this tool, but the study's limitations must be acknowledged. First, there are inherent challenges in research involving people with advanced illness (recruiting, lack of energy, attrition due to illness or death) that make it difficult to conduct routine tests of reliability and validity. If we were attempting to validate the original tool in a fully functional population, some of the psychometric problems would have been much less notable. Conversely, diminished physical and functional capacity makes it problematic to apply any subjective, patient-centered QOL tool in a hospice population. Second, the sample size for the hospice group was smaller than originally intended. We had hoped to have equal sample sizes in the two groups. Difficulty

recruiting and maintaining hospice patients in the study proved to be a substantial methodological challenge that researchers in palliative and end-of-life care frequently encounter. As a result of this small sample, we had limited power in the responsiveness analyses aimed at hospice patients ($n = 31$) or in patients with significant health state changes because there were not many major health changes in the end-stage renal patients over 21 days. Despite this limited power, significant differences were detected in the responsiveness analyses. A third limitation of this work is that many of the conclusions drawn about the MVQOLI-R were related to skewed item distributions. These items may be skewed because of strong demand characteristics in the item wording, and the problem of skewness limited the tool's ability to perform well on many psychometric tests. For example, factor analysis and internal consistency analyses both require adequate variability (i.e., normal distributions) to optimize their performance and interpretability. In retrospect, it would have been useful to hone item wording prior to psychometric and clinimetric testing. The item wording remained suboptimal in clinical use as well. Future research may undertake cognitive testing of revised items prior to starting a large data collection effort, and then collect data on a small sample ($n = 50$) to ensure that item distributions are normally distributed.

In summary, our psychometric and implementation studies support the notion that the MVQOLI-R is a valuable clinical tool, but is not useful in its current form as a psychometric instrument for measuring outcomes in patients with advanced illness. This notion has been exemplified by qualitative research^{31,32} in which the MVQOLI was found to be useful in better understanding patient QOL and needs. Based on these reports and our own findings, we believe that the MVQOLI-R may be useful for encouraging and supporting collaborative care that address multiple dimensions of patients' comfort and quality of life. It may thus be useful in clinical research for advancing individual patient assessment and improving care planning and implementation by the clinical team. This may help advance patient-centered palliative approaches earlier in the course of progressive, ultimately incurable illness. Recent research on collaborative care suggests that concurrent care models are more effective in reducing symptoms and distress in patients with progressive, incurable illness.³³

Future research might evaluate the use of the MVQOLI-R as an adjunct health delivery “intervention” to collaborative care. Our results suggest that the MVQOLI-R has clinical utility as a patient quality of life assessment tool and may have therapeutic utility for fostering discussion among patients, their families and clinicians and in identifying sources of suffering as well as opportunities during this time in their lives.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of Noreen Basque, R.N., M.S., Christine McCluskey, R.N., M.S., and Rebecca Feinberg, B.A., for their assistance in coordinating the study; Kathy Egan, R.N., M.A., C.H.P.N., for assistance in designing and conducting the implementation study; Phillip Merriam for his helpful suggestions for participant recruitment; the study interviewers and site coordinators who implemented data collection; Jean Edmunds for her helpfulness in many aspects of study implementation and literature review; Yunsheng Ma, M.P.H., Ph.D., for his assistance with data analysis; Jerry Waldman for assistance with manuscript preparation; and Peter Fayers, Ph.D.; and Maria Silveira, M.D., M.A., M.P.H., and an anonymous reviewer who provided helpful discussions and feedback on ideas presented herein. Funding for this work was provided by grants from the Partnership for Caring initiative of The Fan Fox and Leslie R. Samuels Foundation and by VITAS HealthCare Corporation.

REFERENCES

- World Health Organization: World Health Organization Constitution. In: *Basic Documents*. Geneva: World Health Organization, 1948.
- Portenoy RK, Thaler HT, Kornblith AB, Lepore JM, Friedlander-Klar H, Kiyasu E, Sobel K, Coyle N, Kemeny N, Norton L, et al: The Memorial Symptom Assessment Scale: An instrument for the evaluation of symptom prevalence, characteristics, and distress. *Eur J Cancer* 1994;30A:1326–1336.
- Brown University. *ToolKit of Instruments to Measure EOL Care*. 2000. <http://www.cher.brown.edu>.
- Steinhauser KE, Bosworth HB, Clipp EC, McNeilly M, Christakis NA, Parker J, Tulsy JA: Initial assessment of a new instrument to measure quality of life at the end of life. *J Palliat Med* 2002;5:829–841.
- Brady MJ, Peterman AH, Fitchett G, Mo M, Cella D: A Case for Including Spirituality in Quality of Life Measurement in Oncology. *J Psychooncology* 1999; 8:417–428.
- O’Boyle CA, McGee HM, Browne JP: Measuring response shift using the schedule for evaluation of individual quality of life. In: Schwartz CE, Sprangers MAG (eds): *Adaptation to Changing Health: Response Shift in Quality-of-Life Research*. Washington, D.C.: American Psychological Association, 2000, pp. 123–136.
- Byock IR, Merriman MP: Measuring quality of life for patients with terminal illness: The Missoula-VITAS® quality of life index. *J Palliat Med* 1998;12:231–244.
- Byock IR: The nature of suffering and the nature of opportunity at the end of life. *J Clin Geriatr Med* 1996;12:2–10.
- Jansen SJT, Stiggelbout AM, Nooij A, Kievit J: The effect of individually assessed preference weights on the relationship between holistic utilities and non-preference-based assessment. *Qual Life Res* 2000;9: 541–557.
- Osman LM, McKenzie L, Cairns J, Friend JAR, Godden DJ, Legge JS, Douglas JG: Patient weighting of importance of asthma symptoms. *Thorax J* 2001;56:138–142.
- Mozes B, Shabtai E: The contribution of personal rating to the clinimetric functioning of a generic quality of life instrument. *J Clin Epidemiol* 1996;49:1419–1422.
- Trauer T, Mackinnon A: Why are we weighting? The role of importance ratings in quality of life measurement. *Qual Life Res* 2001;10:579–585.
- Nunnally JC, Bernstein IH: *Psychometric Theory*, 3rd ed. New York: McGraw Hill, Inc., 1994.
- Feinstein AR: Clinimetric perspectives. *J Chron Dis* 1987;40:635–640.
- Wright JG, Feinstein AR: A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol* 1992;45:1201–1218.
- Meehl PE: *Construct Validity in Psychological Tests. Psycho-diagnosis: Selected Papers*. New York: W.W. Norton & Company, Inc., 1973:3–31.
- Hays RD, Hadorn D: Responsiveness to change: An aspect of validity, not a separate dimension. *Qual Life Res* 1992;1:73.
- Fayers PM, Hand DJ: Factor analysis, causal indicators and quality of life. *Qual Life Res* 1997;6: 139–150.
- Fayers PM, Hand DJ: Causal variables, indicator variables and measurement scales: an example from quality of life. *J R Stat Soc* 2002;165(Part 2):233–261.
- Folstein MF: The mini-mental status examination. In: Crook T, Ferris S, Bartus R (eds): *Assessment in Geriatric Psychopharmacology*. New Canaan, CT: Mark Powley, 1983, pp. 50–51.
- Ryff CD: Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J Person Soc Psychol* 1989;57:1069–1081.
- Cella DFS, Jacobsen PB, Holland JC, Orav J, Silberfarb

- PM, Rafla S: A brief POMS measure of distress for cancer patients. *J Chron Dis* 1987;40:939-942.
23. McNair DM, Lorr M, Droppleman LF: *EITS Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service, 1971.
 24. Cronbach LJ: Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
 25. Kramer MS, Feinstein AR: Clinical biostatistics LII. The biostatistics of concordance. *Clin Pharmacol Ther* 1981;29:111-123.
 26. Cohen J: A power primer. *Psychol Bull* 1992;112:155-159.
 27. Littell R, Milliken G, Stroup WW, Wolfinger RD. *SAS System for Mixed Models*. Cary, NC: SAS Institute Inc., 1996.
 28. Neuhaus JM, Kalbfleisch JD: Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics* 1998;54:638-645.
 29. Begg MD, Parides MK: Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Stat Med* 2003;22:2591-2602.
 30. Miller K: Conducting cognitive interviews to understand question-response limitations. *Am J Health Behav* 2003;S264-S272.
 31. Hill N: Use of quality of life scores in care planning in a hospice setting: A comparative study. *Int J Palliat Nurs* 2002;8:540-547.
 32. Hill N: Use of quality of life scores in care planning in a hospice setting: the theme of revelation. *Int J Palliat Nurs* 2002;8:591-593.
 33. Rabow MW, Dibble SL, Pantilat SZ, McPhee SJ: The Comprehensive Care Team: A controlled trial of outpatient palliative medicine consultation. *Arch Intern Med* 2004;164:83-91.

Address reprint requests to:
Carolyn Schwartz, Sc.D.
31 Mitchell Road
Concord, MA 01742

E-mail: cschwartz@qualitymetric.com