

Quality of Life Measures for Nursing Home Residents

Rosalie A. Kane,¹ Kristen C. Kling,² Boris Bershadsky,¹ Robert L. Kane,¹
Katherine Giles,¹ Howard B. Degenholtz,³ Jiexin Liu,¹ and Lois J. Cutler¹

¹Division of Health Services Research and Policy, School of Public Health, University of Minnesota, Minneapolis.

²Department of Psychology, St. Cloud State University, St. Cloud, Minnesota.

³Department of Health Services Administration, Center for Bioethics and Health Law,
University of Pittsburgh, Pittsburgh, Pennsylvania.

Background. Quality of life (QOL) is a goal for nursing home residents, but measures are needed to tap this phenomenon.

Methods. In-person QOL interviews were attempted for 1988 residents, stratified by cognitive functioning, from 40 nursing homes in five states. Likert-type response options were used with reversion to dichotomous responses when necessary; *z*-score transformations were used to combine the formats. Tests of internal consistency and confirmatory factor analysis were performed; cluster analysis was used to shorten the scales. Correlations between domain scores were examined, and tests of convergent validity performed. Analyses were repeated for subgroups based on cognitive functioning levels.

Results. Long QOL scales were constructed for 1316 of the 1988 residents, including many with substantial cognitive impairment. Confirmatory factor analysis confirmed 10 QOL domains. Cronbach alphas ranged from .76 to .52. The majority (93%) of the 45 possible interscale correlations among domains were below .14 and the rest were between .4 and .5. QOL scales were correlated with, but distinct from, residents' emotions ratings and overall satisfaction, and each was correlated with a corresponding summary rating for the domain.

Conclusions. QOL can be feasibly measured from resident self-report for much of the nursing home population, including cognitively impaired residents. Additional research is suggested on the measures, but the approach has promise for regulation, continuous quality improvement, and public information.

BESIDES achieving good quality of care and good physical and mental health outcomes, long-term care (LTC) should preserve and promote quality of life (QOL). Insufficient research has been devoted to QOL in nursing facilities (NFs). Moreover, although health care research tends to use a restricted concept, health-related quality of life (HRQOL), for general and disease-specific measures (1,2), the outcomes of NFs should include elements of living as well as care. NF life can be sterile, regimented, devoid of both privacy and meaningful association, and deadening to the human spirit (3–6), and health care providers are thus challenged to monitor QOL defined broadly. The distinction between quality of care and QOL is perhaps spurious because the former contributes to the latter. However, the psychological and social aspects of QOL have not yet been measured in a widespread way for NF residents.

In response to the NF regulatory reforms of 1987 (7), the Centers for Medicare and Medicaid Services (CMS) commissioned a standardized minimum data set (MDS) to assess NF residents (8). By 1998, all NFs that accepted federal funds were using the MDS. MDS-derived quality indicators were developed to guide CMS and state inspections (9,10). Although a mandated standard for NFs since 1987, QOL is less clearly articulated in quality indicators, and regulators give relatively few QOL deficiencies (11).

CMS contracted with us to develop QOL measures that reflected psychosocial domains that were either omitted from or not directly emphasized in the MDS to counterbalance the necessary attention to quality of care, which is

largely captured by MDS-derived indicators. Regulatory QOL expectations include the following: individualized care plans that reflect choice and accommodate individual needs and preferences; activity programs maximizing individual interests; a comfortable, clean, homelike environment; ability to retain personal belongings; and a range of specific residents' rights (e.g., privacy, association, information, and refusing care). Our measure included those domains of QOL explicitly referenced in NF regulations (i.e., dignity, privacy, choice, and individuality) as well as other constructs related broadly to a good QOL.

QOL can be assessed by using different data sources and reporters (12). For NF residents, reporters have typically been staff members (the synthesizers of information to enter into the MDS) and occasionally family members; these sources of data are typically correlated with each other but are far from identical (13,14). Direct observation of the resident is also feasible, with specific observations inferred as reflecting a positive or negative QOL (15). Because QOL is inherently subjective, residents themselves, arguably, are the best reporters of their QOL. Moreover, QOL is not a technical matter in which a professional's expertise sometimes overrides the resident's. Like Rubinstein (16), we argue that residents' own appraisal of their lived experiences should be integral to QOL assessments. Therefore, we considered the residents' self-reports to be the "gold standard" for residents' QOL, and one of our goals was to determine the extent to which these self-reports could be collected in NFs.

DOMAIN IDENTIFICATION

Although, following the thinking of Lawton (17), we perceive them as part of QOL, we excluded emotional, health, and functional status from measure development because they were beyond our contract scope and because self-report measures for those constructs already exist in the literature (18–21). Using literature review, expert opinion, focus groups, and discussions with stakeholders, we identified 11 QOL domains pertinent to nursing home life. For each domain, we generated candidate items by using the aforementioned sources and refined through iterative pretesting for clarity and acceptability to residents. The 11 domains identified were as follows: comfort, functional competence, autonomy, dignity, privacy, individuality, meaningful activity, relationships, enjoyment, security, and spiritual well-being; these domains have been described and defended elsewhere (22). Each had strong arguments for inclusion, though we make no claim that they are exhaustive. Like others, we conceptualized QOL as a general construct and the various QOL domains as independent but related dimensions of overall QOL (23).

Comfort and *security* are basic building blocks in any schema of well-being (24). The *comfort* domain most overlaps with quality of care domains covered in the MDS, though we included a wider range of discomforts, such as being too hot, too cold, or in an uncomfortable position, as well as pain and its amelioration. The *security* domain referred not to safety in the sense of minimizing falls and accidents (typically construed as quality of care), but rather to an overall sense of security, safety, and order. Order and predictability are related to a sense of security, and in turn, to overall QOL (25).

Another four domains fall into the social sphere. *Relationships* was defined as any relationships that the resident found meaningful, including those with residents and staff, and family, friends, and associates outside the NF. Close friendships, and the experience of having and being a confidante, have been shown to be important elements of social well-being (26). Because some residents cannot go beyond being observers of the scene around them, whereas others can be active participants in physical and intellectual activities, *meaningful activity* was defined to encompass a range of discretionary experiences, excluding activities of daily living (ADL) tasks. *Functional competence* was defined to mean that, within their physical or cognitive abilities, residents were as independent as they wanted to be. Importantly, this domain does not refer to the ability to perform (or actual performance of) ADL activities or duplicate an ADL measure. We expected this domain to be sensitive to the adequacy of the physical environment to support competence and to the rules and prohibitions of the facility, which can foster or impede a resident's exercise of independence. We hypothesized that pleasurable experiences might constitute an *enjoyment* domain. Although we intended that *enjoyment* would include a variety of experiences, the only items in the subscale we generated relate to enjoyment of food.

Another cluster of potentially related domains deals with constructs highly related to self-worth and individual

agency: namely, *individuality*, *autonomy*, *privacy*, and *dignity*. A substantial literature asserts the importance and difficulty of residents maintaining a sense of continuity and identity, and suggested inclusion of *individuality* (6). *Autonomy* for NF residents, specified as being self-directing, exercising choice, and having perceived control, has been associated with improved morbidity and mortality (27–29). Studies of consumer preference suggest that *privacy*, which is control of information about oneself and experience of solitude or private interaction with people of one's own choosing, is highly desired; in theory, privacy is a prerequisite for autonomy (30). *Dignity* is both a feeling experienced by residents (e.g., dignity being respected) and a status conferred on residents (e.g., residents allowed their dignity). Some residents may be incapable of perceiving indignities (though they should nonetheless be treated with dignity), whereas others may be intensely conscious that their dignity is compromised. Indeed, some ethicists believe that all residents' QOL is diminished when any resident is belittled or has his or her dignity violated (31,32). Finally, we tried to tap the domain of *spiritual well-being*, one receiving attention in the literature (33,34).

METHODS

Sample

We tested the measures in five states (California, Florida, Minnesota, New Jersey, and New York) that varied widely in Medicaid reimbursement rates, deficiency patterns, NF supply, ownership patterns, and labor force characteristics. State catchments were defined (each containing 300–400 NFs). A random sample of NFs was drawn within each catchment area, stratified by size, and urban–rural location; NFs with fewer than 50 beds were excluded. NF participation was voluntary; recruitment visits were made to 49 NFs before 40 were selected; we omitted 6 facilities because of an atypical population (e.g., mostly psychiatric) or unusual circumstances (e.g., a large ongoing building project), and 3 refused because of staff shortages or management changes. Within a NF, we chose up to five nursing units; in the 6 with more than five units, we first selected any Alzheimer's Special Care Unit or Medicare unit (to maximize variation and ensure we did not systematically exclude residents assessed as cognitively impaired) and randomly chose the rest.

We excluded residents who were under the age of 65, comatose, or in a vegetative state. We then strove for a random stratified sample of 50 residents per NF, evenly stratified by cognitive abilities. To assess cognitive abilities, we modified Lawton's previously validated summative cognitive function score (35). The resultant 6-point scale is made up of binary items on short-term and long-term memory loss and a 4-point rating on decision-making capacity; a score of 0 represents no impairment and 5 represents severe impairment. This modified cognitive score is correlated at .88 with the longer version and at .93 with the Cognitive Performance Score developed by Morris and colleagues (36). The cognitive score and birth date were abstracted from the facility records 2–3 weeks ahead of the

expected data collection. The sample was then selected to choose evenly between a cognitive score of 0–2 and 3–5 and, if possible, across participating nursing units. We randomly selected a sample of 2000 (50 per facility) and a replacement group using this stratifying procedure, with the latter to replace sample members who died, were discharged, were absent for the weeks of data collection, or refused to participate. From the original sample, we were unable to include 179 residents; 55 had been discharged, 48 had died, 18 were hospitalized, 13 were temporarily absent, and 45 (2% of all those approached) refused to participate. In 2 small NFs, with less than full occupancy and younger residents to be dropped, insufficient replacements were available to reach 50, and the final sample is composed of 1988 residents.

Measures

Domains.—Eighty-eight candidate items were generated for the 11 QOL domains with 4–13 items assigned per domain. Seventy-two of these potential QOL items provided a 4-point Likert response format, usually with the choice, “often, sometimes, rarely, or never.” The protocol required that the interviewers encourage the respondent to reply by using Likert formats. After three attempts to evoke a Likert-type response, the interviewer offered respondents a response choice of “mostly yes” or “mostly no.” Of the remaining 13 QOL candidate items, 7 were asked in a dichotomous format from the outset. These included social relationship questions (i.e., considering *any* resident to be a *close* friend, having a confidante, and being a confidante), 3 meaningful activity items (i.e., developing a new interest, continuing with an old interest, and leaving the facility grounds for a nonmedical reason), and 1 individuality item (bringing something to the nursing home that made resident feel at home). Four questions were posed in terms of whether some activity (e.g., taking walks, bathing, being outdoors, and participating in religious observances) happened too much, not enough, or the right amount; these were treated dichotomously with the right amount contrasted to “too much” or “not enough.” Finally, 1 item had five possible choices and 1 had three.

Emotional well-being.—We adapted part of the Dementia QOL scale (37,38) to measure emotional well-being. Residents were asked how often they experienced 10 feeling states (lonely, happy, bored, angry, contented, worried, interested in things, sad, afraid, and looking forward to the future) during the past few weeks (often, sometimes, rarely, or never). As with the QOL items, we developed an alternative format for the emotions, which asked whether or not the residents had felt the particular emotion in the past 2 weeks. This combined scale yielded a Cronbach alpha of .80. According to Lawton’s model of QOL, self-reported QOL (one of his four dimensions) should be correlated with emotional health, another dimension (17).

Satisfaction.—Also for comparison with our QOL measures, we asked four global questions, each on a 4-point scale, about satisfaction tapping the NF’s services and pro-

grams, their own room, and bathroom; the physical setting other than their room and bathroom; and the likelihood of recommending the NF to a friend. A comparable yes–no fall-back format was used for residents who could not use the Likert option. The combined scale had an alpha of .75.

Summary ratings.—Residents rated the QOL in their NF according to 11 summary items (each reflecting a QOL domain) and his or her life as a whole. For example, we asked, “considering your life at _____, how would you rate the quality of your life as far as feeling physically comfortable: excellent, good, fair, or poor?” and “Having the privacy you want?” These ratings also offered both Likert and binary response formats.

Data Collection

Data were collected between January and June, 2000 by specially trained data collectors. In the week-long training, attention was given to practicing the movement between the two response formats in role-playing and in actual training interviews. A detailed manual was provided with examples of how to keep the respondent engaged during the long interview and how to obtain responses without biasing the respondent.

No sampled residents were eliminated for interviews without in-person screening; 325 residents who could not be roused or respond coherently to a screening protocol on the basis of responses to a greeting were not interviewed. Interviews were attempted for all others, but interviewers were permitted to abort the interview if residents gave non-responsive or incoherent answers to 8 of the 13 comfort items on the first page of the questionnaire. Interviews lasted 45 to 90 minutes and were ordinarily conducted in residents’ rooms. Occasionally, because of residents’ schedules or fatigue levels, interviews were done in several sittings; divided data collection was needed for only 5% of the QOL interviews. Using the Likert responses did not account for longer interviews; interview length was increased when the resident was gregarious or in need of help in becoming oriented to the interview.

Analyses

Developing a scoring system.—Because residents could shift between the Likert-type and binary response options during the interview, we generated a procedure to combine the two response formats. As a first step in achieving comparability across the two response options, we applied *z* transformations for each item, separately for the Likert and dichotomous response options. The results of these *z* transformations were combined into one variable that included information from all residents who had answered the question, regardless of response option chosen. For each of these combined *z*-transformed items, we examined the “yes” and “no” responses relative to the Likert-type responses. For the majority of items, the “yes” and “no” responses corresponded to 3.8 and 1.5, respectively, in the metric of the Likert-type responses. Because our goal was to create a simple method of interpolating these dichotomous responses into the Likert-type responses, we adopted this

recoding scheme for all items that offered both response sets. Using this information, we used the following transformation to generate item scores for all individuals who provided a valid response (i.e., Likert-type or “yes/no”): most positive Likert = 4; positive binary = 3.8, positive Likert = 3, negative Likert = 2, negative binary = 1.5, and most negative Likert = 1. For the small number of dichotomous items, we coded yes as 4 and no as 1, so as not to count positive responses on important questions (e.g., having any resident as a close friend, having a confidante, or developing a new interest) as less than the Likert-style questions that were scored between 4 and 1. The five possible responses to the question “How often have you been outside?” were scored by using equal intervals between 4 and 1, and the item “Do you consider one or more staff member to be a friend?” was scored as 4 for “more than one,” 3 for “one,” and 1 for “none.” We required that 75% of items be completed for scales with four items and at least 66% for scales with five or more items, in which case we imputed the missing items to the mean of that individual for the scale. No missing items were allowed for the three-item scale.

Validation and item reduction.—We specified a priori which items we thought should load on each of the 11 domains, with the domain assignments ranging from 14 to 4 items. Although we fielded 88 candidate items, we eliminated 12 items from analyses because of highly skewed distributions or high rates of nonresponse. Confirmatory factor analysis (CFA) on the remaining 76 items suggested that the individuality and relationship domains were practically indistinguishable ($r = .99$). Because the reliability of the individuality scale was poor ($\alpha = .56$), we dropped this domain. The resultant model had 10 factors and fit the data well ($\chi^2 = 6024$, $df = 2310$, $p < .000$; root mean square error of approximation or RMSEA = .044; CFI [comparative fit index] = .973; analysis available upon request.)

We next used cluster analysis to produce a short version of the instrument that could be even more useful in a practical NF setting while maintaining the hypothesized domain structure. A list of all 45 possible pairings of the 10 remaining domains was constructed. The items for each pair were then subjected to a cluster analysis by using the squared Euclidean distance as the similarity metric, and Ward’s method was the clustering algorithm. The resulting dendrogram was examined to determine whether the items did indeed form two clusters that corresponded to the a priori classifications of the items. Items that were frequently assigned to domains that did not match the a priori assignment were eliminated. This approach makes few assumptions about the distribution of the data; and, because it uses a limited number of items in each step, it reduces the loss of power that is due to item nonresponse. In addition, cluster analysis enabled us to maximize the independence of the domains. This effort produced a parsimonious subset of 42 items. Finally, we used CFA to validate the domain structure of the short version of the scale. We tested a 10-factor congeneric model with a second-order factor structure in which all 10 of the QOL domains loaded onto 1 higher-order factor (i.e., the latent variable QOL).

We assessed concurrent validity by testing the hypothesis that each QOL domain is positively correlated with overall satisfaction and with better emotional well-being. We also tested the hypothesis that each domain would be associated with the single summary rating of that domain by regressing each domain score on all summary items. We performed analyses separately for the subgroups with better (0–2) and worse (3–5) cognition scores.

RESULTS

Table 1 provides descriptive statistics for the entire sample (1988) and 4 subsamples: those for whom we could calculate no domains, those for whom we could calculate 1–4 domains, those for which we could calculate 5–8 domains, and those for whom we could calculate 9 or more domains ($n = 1125$). Because the computation rules vary with the length of the domains, n s differ slightly and are sometimes larger for the 11 short domains, the 10 short domains, and the 10 long domains, but the patterns shown in Table 2 hold up with all variations of the scale. Sixty percent or more of the residents could respond sufficiently for scale calculation, and residents who could complete any domains were likely to complete all or most. Gender made no difference in completion rates, but older age, poorer ADL functioning, poorer cognitive functioning, and longer length of stay were all highly significantly related to a lesser likelihood of being able to complete the questionnaire. Although older residents were less likely to complete, nonetheless we were able to construct QOL scores from large numbers of very old residents; 71% of the whole sample were 81 years of age or older, and 67% of those for whom we could complete 9–11 domain scores were 81 or older (not shown in the table). The relationship between MDS cognition score and completion was strong, but again some residents with poor cognition were able to complete the tool. Of the 1988 in the sample, 45% had cognitive scores of 4 or 5; we could calculate 9 or more domain scales from 24% of that group, compared with 45% of those with scores of 0 and 1. Conversely, 24% of the 1225 residents for whom we could calculate 9 or more domain scores fell in the cognitive score range of 4–5. Because the sampling method was based on enumeration of the census 2–3 weeks prior to interviewing, very short-stay residents are underrepresented and only 14% had been admitted within the past 3 months.

Residents in the upper half for cognitive functioning (0–2) were statistically significantly more likely to use Likert responses than those scoring 3–5 ($p < .01$). Cognitive scores were related to the type of response format (Likert or binary) used; 54% of the whole sample and 79% of those who with cognitive scores of 0–2 completed 85% or more of the items with the Likert responses. When we relaxed the standard to 75% of the possible Likert response patterns completed by using the Likert scales, the proportions were 58% for the whole sample and 85% for those with a cognitive score of 0–2.

Results of the CFA on the final version of the short scales supported a 10-factor structure with 42 items. The data fit the model well ($\chi^2 = 2441$, $df = 1024$, $p < .000$; RMSEA = .041; CFI = .985). Table 2 presents the standard-

Table 1. Description of Sample by Scale Completion for Long Scales

Trait	Total Sample (%) (N = 1988)	No. of Domains Completed (%)			
		0 (n = 672)	1-4 (n = 104)	5-8 (n = 87)	9-11 (n = 1125)
Female	76.7	80	75	76	75
Mean age*	85.3	86.6	86.7	87.4	84.2
Cognition score [†]					
0 (best)	19	4	9	9	29
1	11	4	5	10	16
2	13	5	15	14	17
3	13	10	14	18	14
4	29	38	40	40	22
5	26	39	16	8	2
Cognitive group*					
0-2	13	29	33	62	43
3-5	87	71	67	48	57
Mean LOS*	32.56	40.9	33.52	26.68	27.94
ADL group					
Needs help w/ 0-2 ADLs	75	55	73	82	86
Needs help w/ 3-5 ADLs	25	45	27	18	14
Bed mobility and/or eating*					
Needs help w/ neither	79	60	78	86	91
Needs help w/ 1 or both	21	40	22	14	9

Notes: ADL = activities of daily living; those used in this scale are bed mobility, eating, transferring, toilet use, and dressing. LOS = length of stay.

[†] Percentage may be more or less than 100% because of rounding.

* All significant differences are at $p < .000$. Chi-squares were used for categorical variables (gender, ADL group, cognitive group, and needing help with bathing or toileting); analyses of variance were used for age and LOS.

ized CFA regression weights for each item. All parameters were statistically significant (Figure 1). The paths from the second-order overall QOL factor (labeled QOL) to each domain were statistically significant. The regression weights ranged from .194 for comfort to .471 for functional competence.

Table 3 shows the alpha reliability for the shortened scales for the entire group and by cognition status. The n varies for each analysis because the number of respondents for whom we could construct the domains varies. All standardized domain scores range from 4 (better QOL) to 1 (worse QOL). The mean scores tended to be positive (>2), but considerable variation was present. The measures of internal consistency (Cronbach's alpha) of each scale ranged from .77 for functional competence to .53 for meaningful activity. We compared the alpha reliability for the high-cognitive group to the low-cognitive group by calculating 95% CI around the alpha from the high-cognitive group (39). The alpha coefficients from the low-cognitive group were not statistically different than the high-cognitive group, except for privacy and enjoyment. The reliability of these two scales in the low-cognitive group was .66.

Correlations among domains reinforced the CFA in demonstrating that the domains are correlated but sufficiently independent to reinforce that different constructs are being measured. Only 3 of 45 possible correlations between

Table 2. Standardized Regression Coefficients of Observed Variables on Latent QOL Factors

Domain	Abbreviated Items	Coefficient
Comfort	Too cold	.347
	So long in same position it hurts	.537
	In physical pain	.516
	Bothered by noise in own room	.512
	Bothered by noise elsewhere in nursing home	.488
Security	Get a good night's sleep	.500
	Possessions are safe	.458
	Clothes lost or damaged in laundry	.388
	Confident can get help when needed	.799
	Can get doctor or nurse quickly	.687
Meaningful activity	Afraid because of how you or others treated	.465
	Get outdoors as much as you want	.424
	How often you get outdoors	.437
	Enjoyable things to do at nursing home on weekends	.576
	Enjoys organized activities at nursing home	.552
Relationships	Gives help to others	.380
	Easy to make friends at nursing home	.628
	Considers any resident to be close friend	.380
	Staff stop just to have friendly conversation	.557
	Consider 1 or more staff to be a friend	.574
Functional competence	Nursing home makes it easy for family & friends to visit	.577
	Easy to get around room by self	.674
	Easily can reach things you need	.692
	Can get to bathroom quickly anywhere in nursing home	.583
	Can easily reach toilet articles	.750
Enjoyment	Take care of things & room as much as wanted	.603
	Like the food here	.771
	Enjoy mealtimes at nursing home	.746
	Get favorite foods here	.592
	Can be alone when want to	.624
Privacy	Can make a private phone call	.476
	Can visit with someone in private	.744
	Can be together with other resident in private	.731
	Staff knock & wait before entering	.436
	Staff treats you politely	.690
Dignity	Staff treats you with respect	.731
	Staff handles you gently	.697
	Staff respects your modesty	.596
	Staff takes time to listen to you	.569
	Go to bed at the time you want	.554
Autonomy	Get up in the morning when you want	.453
	Can you decide what clothes to wear	.459
	Successful in making changes at nursing home	.470
	Participate in religious activities	.390
	Religious observances have meaning	.413
Spiritual well-being	Feel your life has meaning	.759
	Feel at peace	.678

Notes: QOL = quality of life; item text has been shortened because of space considerations.

pairs of domains exceeded .4 (between dignity and security, functional competence and autonomy, and spiritual well-being and relationships) (Table 4). None of the possible pairs had a correlation of .5 or more, and 30 of them were .3 or less.

We compared each QOL domain score with scales measuring two constructs that we expected to be related but not redundant: emotional well-being and satisfaction. Table 5 shows the results for the emotions scale. The cor-

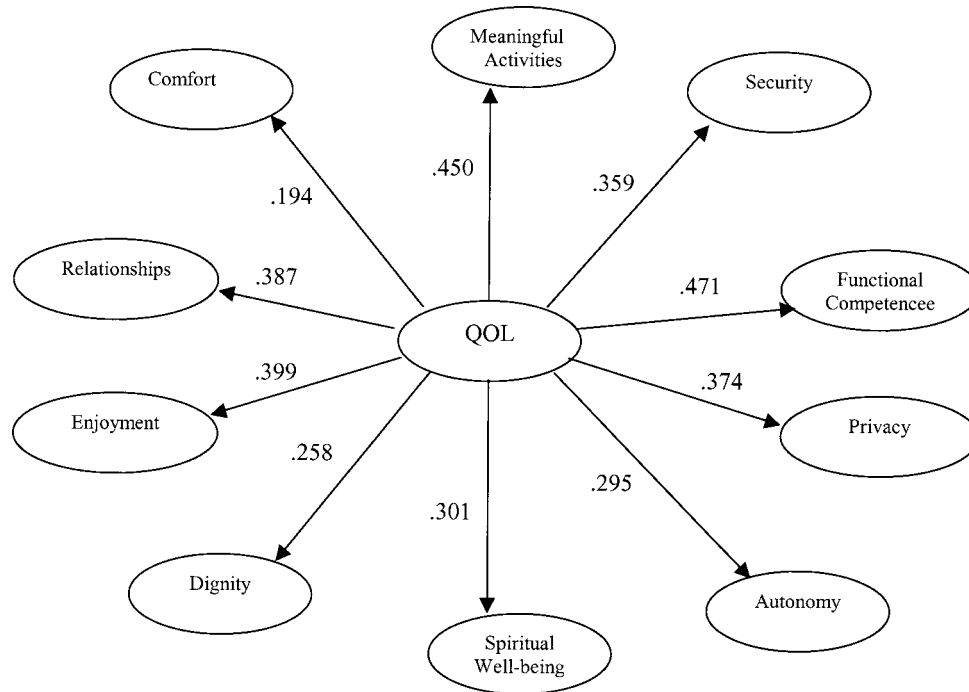


Figure 1. Domain loadings on overall quality of life (QOL) based on the second-order confirmatory factor analysis of the short scales.

relation coefficients were all statistically significant, but the patterns of correlations suggest that the emotions score was sufficiently different from the QOL scales that different constructs are being measured. Table 6 shows the same result when the domain scores are correlated with the global satisfaction scale. Correlations were performed separately by cognitive stratum. Again correlations were in the desired direction, statistically significant, and at level that suggests the constructs are related but different.

We regressed each QOL domain score against the full array of summary ratings (not tabled). The domain scale score was always significantly associated with the matching summary item at the .001 level. The domain score was sometimes also significantly correlated with one or more additional domain summary items, but the largest regression coefficient was with the summary rating for the appropriate domain.

DISCUSSION

Summary

We were able to interview large numbers of NF residents and create measures of relevant aspects of QOL for them. Interviews sufficient to develop QOL scales were completed for approximately 60% of residents (the percentage varying among the 11 long and short scales and the 10 long and short scales remaining after factor reduction, because the different scale lengths led to different imputation conventions). This result was achieved even though at least half of the sample included the more impaired levels on a cognitive performance scale; only 19% of the sample had a perfect

score cognitive score, and 17% had the worst possible cognitive score. (We expect somewhat better completion rates for our subsequent fielding of the measures, because the domains are well defined and grouped and the total questionnaire will be shorter.) By comparison, among all residents in the sample NFs from which the samples were drawn, 13% had a perfect score and 29% had a score of 5, suggesting that we somewhat oversampled those with higher cognition. A 1993 national sample of NF residents that used the same scale found slightly better rates of cognitive functioning; 16% of the residents had no cognitive impairment and 23% were severely impaired (40).

We confirmed 10 distinct factors that each related to an overall construct of QOL in both longer and shorter scales.

Table 3. Reliability of Domain Scores by Cognitive Function

Domain	n	Mean (SD)	Alpha		
			All	High Cognition	Low Cognition
Comfort	1066	3.02 (.06)	.62	.63	.59
Security	931	3.40 (.56)	.65	.66	.62
Meaningful activity	907	2.69 (.72)	.53	.53	.53
Relationships	992	3.07 (.70)	.64	.63	.66
Functional competence	962	3.25 (.74)	.77	.76	.79
Enjoyment	1081	3.22 (.73)	.71	.73	.66*
Privacy	849	3.34 (.63)	.70	.72	.66*
Dignity	1076	3.67 (.43)	.76	.76	.75
Autonomy	766	3.30 (.64)	.59	.58	.60
Spiritual well-being	966	3.15 (.70)	.64	.65	.61

* Alpha in the low-cognition group was significantly different from the high-cognition group at $p < .05$.

Table 4. Correlations Among Domain Scores

Domain	Cmf	Sec	MAct	Rel	Fcomp	Enj	Priv	Dig	Aut	SWB
Comfort	1.00	.36	.12	.08	.23	.26	.20	.26	.19	.11
Security		1.00	.25	.25	.29	.35	.26	.49	.33	.23
MAct			1.00	.39	.35	.30	.27	.24	.29	.31
Rel				1.00	.22	.34	.32	.35	.23	.41
Fcomp					1.00	.22	.35	.30	.46	.18
Enj						1.00	.18	.30	.26	.33
Privacy							1.00	.37	.37	.22
Dignity								1.00	.34	.24
Aut									1.00	.16
SWB										1.00

Notes: Cmf = comfort; Sec = security; MAct = meaningful activity; Rel = relationships; Fcomp = functional competence; Enj = enjoyment; Priv = privacy; Dig = dignity; Aut = autonomy; SWB = spiritual well being.

The merit of the shorter scales is practicality; although for some domains the difference in length is no more than 1 item, the security, comfort, and autonomy scales were shortened respectively by 8, 7, and 4 items. Longer scales will still be useful for more in-depth exploration of or quality improvement efforts related to a particular domain.

The scales performed well in several tests of concurrent validity (correlation with a summary item for the particular scale, correlation with ratings of emotion, and correlation with global satisfaction ratings). We make no claim, however, that we have tapped the entire construct of QOL. For some purposes, measures of affect, functional status, and self-perceived health should be fielded along with our measures to produce a more fully rounded picture of QOL. Self-reported measures of such domains are readily available in the gerontological literature (35–38) and could be added to our battery for research or quality-assurance purposes. Indeed, we did include affect measures in this study for purposes of concurrent validity testing and found them correlated, as expected. Continued pursuit of a measure of individuality seems warranted given its theoretical importance.

Table 5. Correlations Between QOL Short Scales and Emotions by Cognition Group

QOL Domain	Correlations* With Emotions Scale		
	All Residents	Residents w/ high cog. function (0–2)	Residents w/ low cog. function (3–5)
Comfort	.40	.42	.35
Security	.42	.38	.48
Meaningful activity	.29	.30	.26
Relationships	.27	.24	.30
Functional competence	.29	.27	.31
Enjoyment	.35	.33	.39
Privacy	.24	.22	.29
Dignity	.33	.33	.34
Autonomy	.29	.30	.27
Spiritual well-being	.34	.37	.29

Notes: QOL = quality of life; emotions summed for the scale are happy, sad, contented, angry, afraid, worried, bored, interested in things, lonely, and looking forward to the future. The alpha for this summative scale is .80.

* All correlations are significant at the .01 level, using two-tailed tests.

Limitations

A limitation of the study is the lack of test–retest data for the QOL measures. For a variety of technical reasons we could not field such a test at Wave 1, though we now have work in progress to accomplish that task and to examine the effect of response pattern on test–retest reliability. To achieve finer gradations in our QOL measure than a dichotomous yes–no answer allows, we used Likert response options. However, to preserve the input of those who could not cope with the complexity of Likert responses, we allowed a simpler response pattern when necessary. Further work will refine the interpolation of Likert and binary responses and explore the relative merits of the two approaches. The Likert approach seems particularly useful for tracking changes over time and for minimizing positivity bias. It seems likely that residents would be more willing to move from often to sometimes, for example, than from “mostly yes” to “mostly no.” If test–retest data show that the stability of Likert responses is less satisfactory than with binary responses, the greater precision of Likert responses is undermined.

Table 6. Correlations Between QOL Short Scales and Global Satisfaction by Cognition Group

QOL Domain	Correlations* With Satisfaction Scale		
	All Residents	Residents w/ high cog. function (0–2)	Residents w/ low cog. function (3–5)
Comfort	.30	.36	.20
Security	.45	.50	.37
Meaningful activity	.31	.33	.27
Relationships	.37	.36	.38
Functional competence	.23	.24	.20
Enjoyment	.38	.42	.32
Privacy	.31	.30	.35
Dignity	.44	.47	.41
Autonomy	.25	.26	.24
Spiritual well-being	.32	.30	.37

Notes: QOL = quality of life; The satisfaction scale was developed from four questions rating satisfaction with the programs and services, satisfaction with one’s own room and bath, satisfaction with the rest of the facility environment, and likelihood of recommending this facility. The alpha for this summative scale is .75.

* All correlations are significant at the .01 level, using two-tailed tests.

Finally, we cannot generalize to all NFs nationally or generate QOL norms given our sampling in this developmental phase. For example, our sampling method generated only a small short-stay sample. There is ample reason to believe that QOL is a different phenomenon among short-stay residents, and perhaps different among those admitted for rehabilitation and as opposed to those admitted near death or to begin a long stay. Future work with different samples will examine such differences.

Implications

We treated resident report as the gold standard for QOL because QOL is inherently subjective. We recognize that such a decision goes against the professional grain. Professional judgments are much more often used as quality markers than direct resident reports. Future work is needed to triangulate the results of data collected from multiple sources on the QOL of the same residents. However, if these multiple sources are discrepant, a quandary arises about which results should take precedence. Unfortunately, we know of no external criterion to validate a resident's self-report on his or her own QOL.

These findings show that it is possible to generate resident self-reported QOL data that could be added to professionally generated assessment data found on the MDS. It remains, however, to find ways to approximate the QOL of the 40% of residents too cognitively impaired to tell us about it (40). Strategies that could be explored include making direct observations of the resident, making direct observations of resident and staff interaction in areas of the NF with high proportions of people with dementia, designing a highly objective protocol to be administered to care staff, inferring QOL for the entire group from that of those who can report, or developing a method to use cognitively intact residents as sentinel reports about the QOL of those too cognitively impaired to be interviewed.

Various practical issues require resolution before wide-scale adoption of resident QOL measures occurs. First, we used external data collectors who were extensively trained, who were not part of the direct care team, and who devoted an average of approximately 1 hour to each interview (all of which entail costs). Second, our interviewers guaranteed the residents *anonymity*, a strategy that encourages truthfulness but is inconsistent with direct efforts to improve a resident's life. We intend further testing to determine whether facility staff (including nurses and other types of staff) can be trained to administer these protocols and can obtain the same results as research data collectors. Harkening back to the three uses for the QOL tools—regulatory oversight, facility-initiated quality improvement, and public reporting—we could envisage internal data collection most readily for quality improvement.

This work is relevant for clinical care, because tools shape behavior. Providing a means to assess QOL increases the attention to this important neglected area. From a research standpoint, such QOL measures (along with direct resident reports on the health and function domains that we did not investigate here) may be correlated with programmatic, environmental, and staffing attributes. Because QOL is a subtle phenomenon that must be viewed on a continuum,

QOL scores should be viewed somewhat differently in a quality-assurance scheme than readily measurable negative events such as bedsores, urinary tract infections, or being in physical restraints. It is also important that the sample of residents interviewed be large enough to obtain stable estimates of QOL in the NF.

We recognize that some professionals will resist the idea of measuring some of the QOL domains. For example, the relationship domain is customarily viewed as outside the control of the facility. We argue that facilities can influence whether the resident reports good or poor relationships both directly because the facility staff relate to the residents and indirectly because the facility provides a milieu that encourages or discourages perpetuation of the residents' existing relationships. Efforts to help NFs use the tools in quality-improvement processes are pivotal, including providing ideas and best practice information on how facilities might influence the QOL domains. More research is needed to identify how resident factors out of the control of the facility (e.g., terminal state, multiple sensory impairments) affect QOL in order for risk adjustment to be refined.

The very act of asking residents directly about their QOL could engage NF staff directly and systematically with residents' opinions about their daily existence in a way that seldom occurs in a typical NF. Such a process militates against the tendency to depersonalize residents, and to view them merely as care recipients rather than people who live out their lives in difficult circumstances.

Measuring QOL requires hubris. To reduce a resident's QOL to a series of answers on standardized scales may seem overly mechanistic. Yet, it seems better to make the effort to talk to residents about these subjects than the alternative of ignoring the topics and allowing comparative judgments about NF quality to be made solely on the basis of information about narrowly construed health outcomes.

ACKNOWLEDGMENTS

This research was supported by a contract from the Centers for Medicare and Medicaid Services (Master Contract 500-96-008). The opinions expressed here are solely those of the authors and should not be interpreted to reflect the views of the funding or employing organizations.

We acknowledge the enormous contributions of M. Powell Lawton, who was a coinvestigator on this project until his sudden illness in late 2000. The authors gratefully dedicate the papers emerging from Wave 1 of this data collection to his memory. We thank our CMS project officers Mary Pratt and Karen Schoeneman, though conclusions should be attributed to the authors, not CMS.

Address correspondence to Rosalie A. Kane, Division of Health Services Research and Policy, School of Public Health, University of Minnesota, 420 Delaware Street, SE, Box 197, D-527 Mayo Building, Minneapolis, MN 55455. E-mail: kanex002@umn.edu

REFERENCES

1. Frytak JR. Assessment of quality of life in older adults. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000: 200–236.
2. Kuyken W, Orly J. Quality of life assessment across cultures. *Intl J Mental Health*. 1994;3:5–27.
3. Agich GJ. *Autonomy and Long-Term Care*. New York: Oxford University Press; 1993.

4. Kane RA, Caplan AL, eds. *Everyday Ethics: Solving Dilemmas in Nursing Home Life*. New York: Springer; 1990.
5. Lidz CW, Fischer L, Anrold RM. *The Erosion of Autonomy in Long-Term Care*. New York: Oxford University Press; 1992.
6. Tobin SS. *Personhood in Advanced Old Age*. New York: Springer; 1991.
7. Institute of Medicine. *Improving the Quality of Care in Nursing Homes*. Washington, DC: National Academy Press; 1986.
8. Morris JN, Fries BE, Mehr DR, et al. MDS cognitive performance scale. *J Gerontol Med Sci*. 1994;49:M174–M182.
9. Zimmerman DR, Karon SL, Arling G, et al. Development and testing of nursing home quality indicators. *Health Care Financ Rev*. 1995;16:107–127.
10. Abt Associates. *Preliminary Report: Pilot Field Data Collection Efforts to Validate Nursing Home Quality Indicators (Performance Measures)*. Report prepared for Centers for Medicare and Medicaid Services, 2001. CMS Contract 500-95-0062. Revised Sept. 26, 2001 by Hebrew Rehabilitation Center for Aged Training Institute, Roslindale, MA and Brown University, Center for Gerontology and Health Care Research, Providence, RI.
11. Harrington C, Carrillo H, Thollaug SC, Summers PR. *Nursing Facilities, Staffing, Residents, and Facility Deficiencies, 1991–95*. San Francisco: University of California, San Francisco, 1996.
12. Zimmerman DR, Bowers BJ. Integrating satisfaction surveys and other sources of information on quality of long-term care. In: Cohen-Mansfield J, Ejaz FK, Werner P, eds. *Satisfaction Surveys in Long-Term Care*. New York: Springer; 2000: 224, 243.
13. Harper GJ. Assessing older adults who cannot communicate. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000: 483–515.
14. Logsdon R, Gibbons L, McCurry S, Terri L. Quality of life in Alzheimer's disease: patient and caregiver reports. *J Mental Health Aging*. 1999;5:21–32.
15. Lawton MP, Van Haitsa K, Perkinson M, Ruckdeschel K. Observed affect and quality of life: further affirmations and problems. *J Mental Health Aging*. 1999;5:69–82.
16. Rubinstein RL. Resident satisfaction, quality of life, and "lived experiences" as domains to be assessed in long-term care. In: Cohen-Mansfield J, Ejaz FK, Werner P, eds. *Satisfaction Surveys in Long-term Care*. New York: Springer; 2000:13–28.
17. Lawton MP. A multidimensional view of quality of life in frail elders. In: Birren JE, Lubben JE, Rowe JC, Deutchman DE, eds. *Concept and Measurement of Quality of Life in the Frail Elderly*. San Diego, CA: Academic Press; 1991:3–27.
18. Grann JD. Assessment of emotions in older adults: mood disorders, anxiety, psychological well-being and hope. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000:129–169.
19. Kane RL. Physiological well-being and health. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000:49–64.
20. Langley LK. Cognitive assessment of older adults. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000:65–128.
21. Pearson VI. Assessment of function in older adults. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000:17–48.
22. Kane RA. Long-term care and a good quality of life: bringing them closer together. *Gerontologist*. 2001;41:293–304.
23. Stewart AL, King AC. Conceptualizing and measuring quality of life in older populations. In: Abeles RP, Gift HC, Ory MG, eds. *Aging and Quality of Life*. New York: Springer; 1994:27–54.
24. Maslow AH. *Towards a Psychology of Being*. Princeton, NJ: Van Nostrand; 1968.
25. Schulz R. Effects of control and predictability on the physical and psychological well-being of the institutionalized aged. *J Personal Soc Psychol*. 1976;33:563–573.
26. Levin CA. Social functioning. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000:170–199.
27. Avorn J, Langer E. Induced disability in nursing home patients: a controlled trial. *J Am Geriatr Assoc*. 1982;30:397–400.
28. Langer E, Rodin J. The effects of choice and enhanced personal responsibility for the aged. A field experiment in an institutional setting. *J Personal Soc Psychol*. 1976;34:191–198.
29. Rodin J. Aging and health: effects of the sense of control. *Science*. 1986;233:1271–1276.
30. Westin A. *Privacy and Freedom*. New York: Atheneum; 1967.
31. Moody HR. Why dignity in old age matters. In: Disch R, Dobrof R, Moody HR, eds. *Dignity and Old Age*. New York: Haworth Press; 1998.
32. Post SG. *The moral challenge of Alzheimer's disease*. Baltimore, MD: Johns Hopkins University Press; 1995.
33. Kimble M, McFadden S, Ellor J, Seeber J, eds. *Aging, Spirituality and Religion: A Handbook*. Minneapolis, MN: Fortress Press; 1995.
34. Fetzer Institute. *Multidimensional Measurement of Religiosity/Spirituality for Use in Health Research*. Kalamazoo, MI: John E. Fetzer Institute; 1999.
35. Lawton PM, Casten R, Parmalee PA, Van Haitsma K, Corn J, Kleban MH. Psychometric characteristics of the Minimum Data Set II: validity. *J Am Geriatr Soc*. 1998;46:736–744.
36. Morris JN, Hawes C, Fries BE, et al. Designing the national resident assessment instrument for nursing homes. *Gerontologist*. 1990;30: 293–307.
37. Brod M, Stewart AL, Sands L, Walton P. Conceptualization and measurement of quality of life in dementia; The Dementia Quality of Life Instrument (DQoL). *Gerontologist*. 1999;39:25–35.
38. Stewart AL, Sherbourne CD, Brod M. Measuring health-related quality of life in older and demented populations. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia: Lippincott-Raven; 1996:819–830.
39. Feldt LS, Woodruff DJ, Salih FA. Statistical inference for coefficient alpha. *Appl Psychol Meas*. 1987;11:93–103.
40. Degenholtz HB, Arnold RA, Meisel A, Lave JR. Persistence of racial disparities in advance care plans among nursing home residents. *J Am Geriatr Soc*. 2002;50:378–381.

Received July 11, 2002

Accepted October 4, 2002