Chapter 4

Family and Line Staff Proxies for Resident Reported Quality of Life

Quality of life (QOL) is a personal and subjective phenomenon. In previous chapters, therefore, we argued that the person living the life is by far the best source of information about his or her QOL. A person cannot be expected to understand how another is experiencing his or her life unless that information is directly communicated. Nonetheless, studying the QOL experienced by nursing home residents using self-report alone would mean that a substantial number of potential respondents who cannot communicate directly because of severe cognitive impairment or some other impairment would be excluded. If these residents are to be included in any efforts to measure QOL, some sort of proxy is needed.

Background

Proxies are regularly used in health and functional surveys, including standard surveys like the National Health Interview Survey. The underlying assumption is that a responsible party has knowledge of facts about another person's health status and use of services. The MDS data relies almost exclusively on a variant of proxy responses to obtain information on nursing home residents. Staff members were expected to observe the residents and infer various items of information from the behaviors noted. Depending on the sections, the person completing the MDS was expected to gather information from a number of staff informants across shifts, from observing and talking to the resident, and/or consulting records.

When factual questions are at issue, accurate proxy data requires selection of a proxy with a high likelihood of knowing the information sought for the relevant time period. When the proxy is asked to respond about the moods, attitudes, and satisfaction of another person, error could occur because the proxy respondent brings his or her own preconceptions or vested

interests to the interpretation of the resident's views.

Proxy information can be used as a direct surrogate for data from a particular resident. Alternatively, proxy data can be aggregated to create a summary measure for a facility. In the latter case, the appropriate test relies not on direct correspondence but on whether the average of the proxy responses compares to the average of he resident responses. The second criterion is easier to meet than the first one.

Proxies are frequently used but their validity is rarely tested. In the National Hospice Study, for example, a substantial amount of the reports on pain control came from proxies because the patients were either already dead or too sick to respond; however, no evidence was presented to show that a third party could describe another person's pain (Greer et al., 1986). Systematic literature review of 24 clinical studies from 1990 to 1999 showed variation in agreement between proxies and older subjects according to the nature of the inquiry. These studies concluded that there was often good comparability around levels of functioning, although proxies tended to identify more impairment; agreement was also good with regard to subjective assessments of overall health, chronic physical conditions and physical symptoms. The investigators report that the limited evidence on agreement around preferences for type of setting was high but that agreement for health state preferences was low. Agreement levels for depressive symptoms and psychosocial well-being was low to moderate (Neumann, Araki, & Gutterman, 2000).

QOL measures have been created for specific conditions (e.g., cancer, hip fractures) or specific aspects of life (health-related QOL) while others are more general and strive to measure global QOL. The sixteen articles presented in Table 4.1 reflect this diversity. In six, the subjects of study are people with some type of cancer. In five, the population is specifically older adults

(although people over 65 are represented in the other categories as well) and in three of these, the population is people with Alzheimer's disease. Two articles examine proxies for people with chronic conditions (stroke and epilepsy) while two more focus on acute care (ICU patients) or treatment (coronary artery bypass surgery) and one additional article assessed proxies for adults with intellectual disabilities.

The nature of the proxy respondent varied in these studies. In eight, the proxy was someone who knew the study subject well in a non-professional capacity, such as a family member, next of kin, friend or informal caregiver. The study subject often nominated these proxies. In two studies the proxies were a health care provider (nursing assistants in one, group home staff in the other), whereas two others used either a family member or provider interchangeably as the proxy, taking whomever was available. Four studies included both a family member and a health care provider.

Despite this diversity in study subjects and proxies, the approach to validity or the criteria used to assess the proxy information was very similar. In fifteen of the sixteen studies, proxy responses were compared to study subject's self-reports.   The authors varied on how they conceptualized this comparison. Most portrayed the study subject as a gold standard, with the subject's responses being considered correct and proxy responses that differed being viewed as wrong. A minority of the authors suggested that self-reports could also be wrong; they lamented the lack of gold standard or proposed that some other source, such as a clinical assessment or administrative records should be used as the criteria for either self or proxy responses.

Table 4.1: Studies Evaluating Proxies for QOL Measures

Author (cite)	Title	Topic/Goal	Population	Proxy	Approach to Validity	Method of Analysis
	Results					
Albert et al., 1996.	*Quality of life in patients with Alzheimer's disease as reported by patient proxies*	To measure QOL for persons with AD in terms of affect and engagement in activities	130 people diagnosed with AD and assessed for QOL	Family or institutional care givers. Selection procedure not specified	State the gold standard of a clinical observation was not available. Relied on comparison of types of caregivers, test-retest and relationship between QOL measures and dementia severity.	T-test and Chi-squares for differences between proxies; Pearson correlations between QOL and dementia severity; regression models for the impact of cognitive and functional status on QOL. Kappa used for test-retest.
	Family and institutional caregivers report in similar ways (this is for different subjects). Variation in QOL was captured by measures. The QOL measures and cognitive status were inversely related.					
Andresen, Patrick, Carter, & Malmgren, 1995	*Proxy reliability: health-related quality of life (HRQOL) measures for people with disability*	Compare self-reports and proxy responses on BRSS, SF-36 and ADL/IADL items	131 people from NHs, assisted/ independent living, and spinal cord injury centers.	Someone who knew target well (family, care providers, friends)	Self-report is considered the standard as HRQOL is described as intended to represent self-perception	Kappa, Interclass correlation coefficient, % agreement, direction and magnitude of disagreement
	Proxy reports are biased and the direction and magnitude depends on the domain. There is more variation in individual questions than summary or scales. Relatives agreed with subjects more than friends or health care providers.					
Capuzzo, Grasselli, Carrer, Gritti, & Alvisi, 2000	*Quality of life before intensive care admission: agreement between patient and relative assessment*	Compare self-report and proxy responses	172 adult ICU patients	Relative considered to be the next of kin	Concordance between patient and relative report	Kappas weighted and unweighted; Bland-Altman level of agreement; logistic regression with agreement as DV; IV proxy characteristics
	High kappas (global .78 and .82) items better for physical function than emotional. Relative (proxy) characteristics did not affect results.					

Author (cite)	Title	Topic/Goal	Population	Proxy	Approach to Validity	Method of Analysis
	Results					
Curtis & Fernsler, 1989	*Quality of Life of Oncology hospice patients: a comparison of 90 patient & primary caregiver reports*	Compare self-report and proxy	23 hospice patients with cancer	Primary caregiver (mostly spouses, others include daughters, friends, nephews)	No statement of which is viewed as correct, rather that disagreement indicates poor communication/ misunderstanding and may affect care.	Comparison of patient and caregiver means for each item and a total score. T-tests of differences between the means
	No significant difference on overall rating or 13 items. Pain ratings differed (p-.02) with patients reporting lower levels of pain.					
Dorman & Waddell, 1997	*Are proxy assessments of health status after stroke with the EuroQOl questionnaire feasible, accurate, and unbiased?*	To assess the utility of proxies for stroke patients	122 patients from a stroke registry	Relative or friend selected by patient and available at time of interview	Comparison to patient as the correct answer as it is stated that the patient's view is more likely to be valid given QOL is about perceptions; however. finds no accepted gold standard	Kappas, Bland & Altman plot and Intraclass correlation coefficient. Subgroup analysis by whether patient was able to complete questionnaire independently or was interviewed
	Better agreement with proxies for patients who were able to complete questionnaire than those requiring interview. Agreement best for self-care and worst for psychological outcomes.					
Epstein, Hall, Tognetti, Son, & Conant, 1989	*Using proxies to evaluate quality of life: can they provide valid information about patients= health status and satisfaction with medical care.*	Compare patient and proxy responses, determine characteristics associated with difference	60 clients over 65 years old from one health service clinic	Close family member or friend	Implied that patient response is the standard	Pearson correlation with z-stats to test for differences across subgroups. Mean differences and these regressed on proxy and subject characteristics
	Correlations ranging from .62 - .73 for 4 health status measures. Lower (.43) for satisfaction. Means were significantly different for satisfaction and emotional health. Higher number of hours helping was related to disagreement on functional status with proxy reporting more disability					

Author (cite)	Title	Topic/goal	Population	Proxy	Approach to validity	Method of analysis
	Results					
Hays et al., 1995	*Agreement between self reports and proxy reports of quality of life in epilepsy patients*	Compare patient and proxy reports of QOL in order to understand the implications of using proxies.	292 patients with epilepsy	Person who had contact with the patient at least once a week	Comparison to patient response, discriminant validity (higher correlations patient-proxy than among scales), lack of scatter bias (relation between patient- proxy and level of QOL)	Average absolute value of the difference, mean bias, standard deviation of difference. Product-moment and intra-class correlations. Multitrait-multimethod analysis of product-moment correlations. Forward stepwise regression using difference and patient, proxy characteristics.
	Agreement was moderate. Means closer than individual scores. Higher education was only characteristic associated with patient-proxy difference. Agreement better for observable measures (working, driving, social& physical function) than subjective (emotions, discouragement, social support)					
Hickey & Bourgeois, 2000	*Health-related quality of life (HR-QOL) in nursing home residents with dementia*	To test several measures with nursing home residents assess the utility of proxy data	107 nursing home residents with dementia	90 nursing assistants assigned to care for residents selected for the study	Comparison of resident and staff reports, and whether the comparisons were consistent over time.	Correlations on only one of the measures tested (The Geriatric Depression Scale)
	Resident and NA reports were not related and NAs reported almost twice as many symptoms as residents					
Moinpour & Lyons, 2000	*Substituting proxy ratings for patient ratings in cancer clinical trials: an analysis based on a Southwest Oncology Group trial in patients with brain metastases*	To explore how bias due to the use of proxies may affect the assessment of a treatment in a longitudinal clinical trial	51 patients in a clinical trial of post-surgical care for brain metastases	Selection method not specified. Primarily spouses and daughters, though relationship was missing for 22%	Patient report considered gold standard. Validity of proxy determined in part by whether its use as a surrogate for patient report alters conclusions.	Lin's concordance statistic (similar to ICC), graphical comparisons of difference versus the average of the two, and weighted Kappas. Double repeated measures analyses of impact of proxy on longitudinal estimates of treatment effect.
	Statistics and plots that show poor to moderate agreement between patient and proxy. Modeling results suggest that use of proxy data would change estimate of treatment effect.					

Author (cite)	Title	Topic/goal	Population	Proxy	Approach to validity	Method of analysis
	Results					
Novella & Jochum, 2001	*Agreement between patients' and proxies' reports of quality of life in Alzheimer's disease*	Study agreement between patients with AD; determine what factors influence agreement and if proxies systematically under or over estimate QOL	76 Patients with AD	Two types of proxies: 1. family members; 2. care providers	Patient report considered the standard and difference between patient and proxy defined as bias.	Percent of exact agreement across items in each subscale of the measure. ICC for subscale scores. Mean difference and effect size for subscale scores.
	Poor to moderate agreement overall. Better agreement for observable measures of function than subjective components. Spouses and nurses agreed with subject more than children or nurse's aides					
Page & Verhoef, 1995	*Quality of life, bypass surgery and the elderly*	Determine if QOL is affected by coronary artery bypass surgery (CABS) and compare patient and proxy ratings both pre and post-op	18 patients over 65 who had CABS from one primary care facility	Selection method not specified. 15 were spouses, others were brother, daughter and friend	States that agreement was the criteria used, but that it is unknown whether the patient of proxy response is more valid	Wilcoxon's match-pair signed-ranks test
	Generally good agreement. Pre-op there were differences in depression and anxiety and in the overall function score with patients reporting more negatively. Post-op no differences were found.					
Sigurdardottir & Brandberg, 1996	*Criterion-based validation of the EORTC QLQ-C36 in advanced melanoma: the CIPS questionnaire and proxy raters*	Validate scales in a QOL measure and determine the feasibility of using nurses and/or relatives as proxies	52 melanoma patients with metastases receiving chemotherapy	Nurses who worked on ward where patients were treated and next of kin selected by patient	Used patient responses as reference. Never compared nurse and family responses	Correlations between subscale scores for patient-nurse and patient-relative pairs and mean difference in scores
	Very low correlations between nurse and patient measures with nurses underrating subjective symptoms and overrating well being. Family correlations were moderate to high with no significant mean differences					

Author (cite)	Title	Topic/goal	Population	Proxy	Approach to validity	Method of analysis
	Results					
Sneeuw et al., 1997	*The use of significant others as proxy raters for the quality of life of patients with brain cancer*	Evaluate the level of agreement between patient and proxy ratings	103 patients with recently diagnosed or recurrent brain cancer	Significant other named by subject	Agreement of proxy with subject defines the quality of the proxy measure	ICC for test-retest reliability of both subject & proxy; % exact & approximate agreement & ICC for proxy subject agreement, ANOVA for factors affecting agreement; Repeated ANOVA and ICC comparisons for change in agreement over time.
	Exact agreement greater than 50%; approximate over 90%. Agreement was lower for patients with worse mental and physical function and for patients whose function deteriorated. Response bias was consistent (proxies lower) but small.					
Sneeuw et al., 1998	*Evaluating the quality of life of cancer patients: assessments by patients, significant others, physicians and nurses*	Assess the usefulness of different types of proxies	90 cancer patients with various diagnoses treated with inpatient chemotherapy at one hospital	Significant other named by patient (90); ward physicians (15) and nurses (35)	Patient as reference for each proxy; agreement across all four raters was used; suspect responses could come from any of the raters, including the patient	t-tests for differences in mean scores, Intra class correlation coefficients, exact and proximate category agreement and simultaneous comparisons of the four raters
	Proxy responses are reasonably accurate and one type of proxy is not better. There is a U-shaped relationship between agreement & patient function with more agreement at low and high levels of function					
Stancliffe, 1999	*Proxy respondents and the reliability of the Quality of Life Questionnaire Empowerment factor*	To assess the reliability of proxy responses on a QOL subscale on empowerment	63 adults with intellectual disabilities in group homes or supportive housing	Community living staff who had extensive contact with study subject	Uses both comparison of proxy-subject and the mean of two proxies	Percent agreement and Spearman correlation coefficients
	Generally good agreement between subject and staff; however a small number of pairs had extreme differences. Agreement was best on concrete items and worse on items about decisions made in the past.					

Author (cite)	Title	Topic/goal	Population	Proxy	Approach to validity	Method of analysis
	Results					
Wilson, Dowling, Abdolell, & Tannock, 2000	*Perception of QOL by patients, partners and treating physicians*	3 way comparison of patient, partner, & MD assessments of QOL	71 patients with metastatic breast cancer; 29 with metastatic prostate cancer	spouse partner or 1st degree relative for breast CA, spouse/partner for prostate CA) and oncologist	Patient self-report is the standard for comparison	ANOVAs with Mean absolute scores and differences; proportion of exact agreement and intra-class correlations between patient and proxy
	Mean agreements were high, but individual agreements were low. Differences were greater on the more subjective domains (social, relationships, role and sexual function).					

Although comparison to self-reports is the most frequent approach used to evaluate proxies, others were occasionally used as well. Three studies included characteristics of the measure (e.g., test-retest agreement/consistency or discriminate validity) or confirmation of a predicted relationship between the proxy reports and some other variables such as dementia severity or level of QOL in their assessments. The approach central to this research, the comparison of proxies, was used in two studies (one used this approach exclusively), while one additional study collected data from two proxies and compared each to the study subjects but never analyzed agreement across the two proxies.

The analytic methods used to quantify the agreement between self-reports by subjects and proxies or to identify factors that might affect agreement varied. In some cases studies used multiple analysis methods because they posed multiple questions. In others, the authors used different methods and explained that there is no consensus on the best measures as each measure has different problems. The most frequently used was the kappa statistic or the intra-class correlation coefficient (ten articles), both of which are measures of agreement adjusted for the probability of chance agreement. Other measures frequently used included correlations (six articles) and percent agreement (six articles). Another category of approaches focuses on the size and/or direction of the differences in responses.   These were used in six articles and ranged from raw differences to graphical representations of standardized differences. Mean values for group responses were also compared using t-tests in four cases. When the research sought to determine how variables such as proxy demographics or frequency of contact between proxy and study subject affect agreement in responses different regression models and ANOVAs were used. Given the differences in populations, proxies, and analysis methods it is not surprising that the results and conclusions often differ as well. For the most part, investigators seemed to rely

on statistical significance to declare proxies as successful alternatives, but the actual correlation coefficients rarely exceeded 0.3, with an occasional 0.4.

With this background, the CMS nursing home QOL study examined the relationship between proxy information on QOL from two sources (family and direct care staff) and information provided directly by nursing home residents.

Methods

Sample

Family sample and data collection. For each resident in Wave 1, one or more family members identified by the resident and/or the nursing home as being close to the respondent received a mailed questionnaire. If no questionnaire was returned for a resident, we contacted the family member by phone, and offered the option of telephone completion. We ranked family members as to current closeness to the resident. If we received two questionnaires for a family member, we used the higher-ranking one for primary analyses (see Appendix D for family questionnaire and contact form).

Staff sample and data collection. For each resident in the Wave 1 sample, we also completed a staff contact form to identify line staff members on the day or evening shift who were most likely to be knowledgeable about that resident. To be eligible for inclusion, the staff member must have cared for the resident for at least two weeks. When primary CNA assignments were used, we identified the resident's primary caregivers.   We, then, assigned a staff member who would be queried about each resident. When choices were available, we allocated the assignments in such a way that no staff member would provide responses related to more than five residents. Staff members received a ten-dollar Wal-Mart gift certificate for each proxy interview they completed. (See Appendix C for the questionnaire and staff contact form.)

<u>Instruments</u>

The questionnaires for family and for staff were developed to parallel the questions posed to residents on QOL and on resident affect. Not all domains were included, however, because they were felt to be unknowable by a third party. Neither family nor staff members were asked the relationship or spiritual well-being items since they entail high subjectivity; and we omitted the dignity items in the staff questionnaire since they seemed to call for an unreasonable amount of self-criticism (e.g. do staff treat Mrs. Jones roughly while giving care). Accordingly, the staff data addressed eight domains and the family data addressed nine domains that were comparable to the resident questionnaire. Staff were interviewed in person and given a $10 gift certificate for each interview. Both staff and family were asked questions about their familiarity and contact with the resident, which were used as control variables in the analysis.

<u>Statistical Analysis</u>

Comparison across domains was done with simple correlations of scale values. We created comparable scales based on matching items for each of the proxy respondent groups and the residents. The scales for both staff and family closely approximate the resident scales in length, being either identical, 1 question shorter, or (in 2 instances for staff and 1 for family) 2 questions shorter. We compared the mean values across groups using ANOVA for independent samples and we compared matched samples of each proxy and the corresponding residents with t tests. To test the correlations, we used three different correlation statistics: Pearson, Kendall and Spearman, as well as calculating Kappa statistics and inter-class correlations (ICC). To obtain the Kappas, we dichotomized the continuous scores at both the 20th and 25th percentiles.

Results

Table 4.2 shows the mean QOL scales for each of the three groups. For each domain the mean vales are significantly different across the groups. Table 4.3 examines the correlations between resident and staff proxy ratings. Three different forms of the analyses are shown: the Pearson correlation coefficients, the kappa statistic based on using the 20th percentile, and the ICC. In all cases, the results are highly statistically significant but the actual values do not show high levels of correlation. Only four of the Pearson coefficients exceed 0.2. The Spearman coefficients were slightly higher, and the Kendall coefficients were lower. The kappa values were all below .2 and improved if the 25th percentile was used. Only four ICC values exceeded 0.2. A regression model was used to adjust proxy responses to account for case mix differences. The independent variables describing residents used in the model were derived from MDS data. We used the same MDS cognition scale reported in Chapter 2.

Table 4.2: Comparison of Mean QOL Values by Respondent Group

Domain	Residents		Staff		Families		F test	
	Mean	SD	Mean	SD	Mean	SD	F value	Sig
Comfort	3.02	0.62	3.00	0.57	2.89	0.57	21.424	0.0001
Functional Competency	3.25	0.74	3.09	0.96	2.85	0.99	57.500	0.0001
Privacy	3.34	0.63	3.53	0.68	3.44	0.66	29.252	0.0001
Dignity	3.67	0.43			3.81	0.36	-9.193	0.0001
Meaningful Activity	2.69	0.72	2.64	0.79	2.51	0.70	18.134	0.0001
Enjoyment	3.22	0.73	3.37	0.58	3.20	0.69	29.918	0.0001
Individuality	2.82	0.73	2.74	0.83	3.08	0.71	69.428	0.0001
Security	3.40	0.56	3.56	0.58	3.33	0.54	69.622	0.0001

Table 4.3: Correlations Between Resident and Staff Proxies

Domain	Pearson Correlation		Kappa (20th percentile)		ICC	
	Coefficient	Sig.	Value	Sig.	Value	Sig
Comfort	0.251	0.0001	0.056	0.0001	0.2498	0.0001
Functional Competency	0.365	0.0001	0.139	0.0001	0.3649	0.0001
Privacy	0.151	0.0001	0.060	0.0001	0.1470	0.0001
Meaningful Activity	0.262	0.0001	0.142	0.0001	0.2613	0.0001
Enjoyment	0.198	0.0001	0.174	0.0001	0.1891	0.0001
Individuality	0.212	0.0001	0.067	0.0001	0.2100	0.0001
Security	0.131	0.0001	0.037	0.0001	0.1314	0.0001
Autonomy	0.169	0.0001	0.063	0.0001	0.1655	0.0001

A similar analysis for the correlations between family proxies and residents is shown in Table 4.4. The level of agreement is a little better than for the staff correlations. Three Pearson correlations are 0.3 or higher. Neither the Kendall nor the Spearman coefficients perform better. The kappa values are all low. One kappa improves greatly if the 25th percentile is used; comfort has a value then of 0.3, still well below any acceptable level. All but one of the ICC values is greater than 0.2 but only one is greater than 0.4.

We performed similar analyses using the single summary rating response for each domain. These results are shown in Table 4.5. None of the Pearson coefficients for staff proxies was equal to 0.2. Nor were any of the ICC values. The family proxies performed slightly better; two of the Pearson coefficients were greater than 0.2, as were the ICC values for these same domains.

Table 4.4: Correlations Between Resident and Family Proxies

Domain	Pearson Correlation		Kappa (20th percentile)		ICC	
	Coefficient	Sig.	Value	Sig.	Value	Sig
Comfort	0.260	0.0001	0.068	0.0001	0.2588	0.0001
Functional Competency	0.458	0.0001	0.111	0.0001	0.4475	0.0001
Privacy	0.226	0.0001	0.088	0.0001	0.2260	0.0001
Dignity	0.142	0.0001	0.083	0.0001	0.1408	0.0001
Meaningful Activity	0.364	0.0001	0.074	0.0001	0.3637	0.0001
Enjoyment	0.354	0.0001	0.136	0.0001	0.3540	0.0001
Individuality	0.259	0.0001	0.090	0.0001	0.2585	0.0001
Security	0.250	0.0001	0.061	0.0001	0.2503	0.0001
Autonomy	0.214	0.0001	0.106	0.0001	0.2695	0.0001

Table 4.5: Correlations Between Proxies and Residents for Summary Items

Domain	Residents vs. Staff				Residents vs. Family			
	Pearson	Sig.	ICC	Sig.	Pearson	Sig.	ICC	Sig.
Comfort	0.126	0.001	0.1244	0.0001	0.151	0.0001	0.1244	0.0001
Functional Competency	0.180	0.001	0.1795	0.0001	0.181	0.0001	0.1795	0.0001
Privacy	0.079	0.004	0.0774	0.0021	0.147	0.0001	0.0774	0.0021
Control-choice	0.073	0.009	0.0728	0.0045	0.199	0.0001	0.0728	0.0045
Dignity	0.110	0.001	0.1081	0.0001	0.132	0.0001	0.1081	0.0001
Interesting things to see & do	0.142	0.001	0.1396	0.0001	0.227	0.0001	0.1396	0.0001
Food enjoyment	0.120	0.001	0.1182	0.0001	0.248	0.0001	0.1182	0.0001
Following preferences	0.026	0.364	0.0256	0.1829	0.159	0.0001	0.0256	0.1829
Relationships	0.110	0.001	0.1097	0.0001	0.160	0.0001	0.1097	0.0001
Security	0.095	0.001	0.0949	0.0002	0.145	0.0001	0.0949	0.0002
Spiritual well-being	0.150	0.001	0.1501	0.0001	0.122	0.0001	0.1501	0.0001
Life as a whole	0.094	0.001	0.0936	0.0003	0.176	0.0001	0.0936	0.0003

Another test of agreement, using the measures of emotions derived from Brod and Stewart. (Brod, Stewart, Sands, & Walton, 1999), is shown in Table 4.6. The correlations between staff and residents were generally low; none of the Pearson coefficients or the ICC values reached 0.2. The correlations with family proxies were somewhat better. Three Pearson coefficients and two ICC values were greater than 0.2.

Table 4.6: Correlations Between Proxies and Residents for Emotion Items

Emotion	Residents vs. Staff				Residents vs. Family			
	Pearson	Sig.	ICC	Sig.	Pearson	Sig.	ICC	Sig.
Lonely	0.085	0.001	0.0831	0.0008	0.159	0.0001	0.1526	0.0001
Happy	0.149	0.0001	0.1405	0.0001	0.219	0.0001	0.2151	0.0001
Bored	0.071	0.009	0.0690	0.0052	0.206	0.0001	0.1969	0.0001
Angry	0.152	0.0001	0.1520	0.0001	0.188	0.0001	0.1860	0.0001
Contented	0.122	0.0001	0.1176	0.0001	0.191	0.0001	0.1887	0.0001
Worried	0.113	0.0001	0.1119	0.0001	0.160	0.0001	0.1567	0.0001
Interested	0.176	0.0001	0.1744	0.0001	0.228	0.0001	0.2273	0.0001
Sad	0.087	0.001	0.0850	0.0007	0.169	0.0001	0.1632	0.0001
Afraid	0.117	0.0001	0.1164	0.0001	0.101	0.002	0.1006	0.0008
Future					0.156	0.0001	0.1553	0.0001

To explore whether the characteristics of the residents affected the concordance of proxy and resident reports, we regressed the resident ratings against the proxy ratings incorporating measures of residents' age, gender, disability level, cognitive status, and length of stay. Table 4.7 shows that adjusting for resident characteristics did not affect the relationship of the proxy report. Nor did the full model account for much of the variance in the resident QOL reports. For staff reports, the best models could account for only 7-8% of the variance.

Table 4.7: Regression Coefficients for Resident Ratings by Domain: Residents vs. Staff

	Domain							
	Comfort	Functional Competence	Privacy	Meaningful Activity	Enjoy-ment	Individuality	Security	Autonomy
(Constant)	1.840	2.592	3.042	2.145	1.754	2.951	2.652	2.793
Proxy	$0.262^{\dagger}$	$0.292^{\dagger}$	$0.133^{\dagger}$	$0.235^{\dagger}$	$0.268^{\dagger}$	$0.158^{\dagger}$	$0.122^{\dagger}$	$0.128^{\dagger}$
LOS	0.073	0.037	0.000	0.075	0.009	0.108	-.140**	-0.037
Gender	0.009	0.037	-0.010	0.053	-0.030	$-0.185^{\dagger}$	-0.008	-0.026
Age	0.003	-0.004	0.000	-0.002	0.006*	-0.006*	0.005**	0.002
ADL	-0.017	$-0.086^{\dagger}$	-0.047**	-0.041*	-0.020	-0.032	-0.024	$-0.079^{\dagger}$
MDS	0.032**	0.037**	-0.059	0.023	$0.051^{\dagger}$	$-0.052^{\dagger}$	-0.005	-0.011
R^2	0.072	0.157	0.059	0.079	0.058	0.080	0.036	0.056

* means statistically significant at the level of $p < .05$.
** means statistically significant at the level of $p < .01$.
† means statistically significant at the level of $p < .001$.

For families, the best model (functional competence) accounted for 23% of the variance because of the strong role played by the ADL adjuster. The models for two other domains accounted for 14% (enjoyment) and 15% (meaningful activity) of the variance in resident reports. The effects of the cognitive variable were mixed. Resident cognitive status was sometimes positively and other times negatively related to the concordance with the resident's QOL rating. By contrast, the ADL dependency measure score was consistently negatively associated with the resident's QOL rating.

Table 4.8: Regression Coefficients for Resident Ratings by Domain: Residents vs. Family

	Domain								
	Comfort	Functional Competence	Privacy	Dignity	Meaning-ful activity	Enjoy-ment	Individ-uality	Relation-ships	Security
Constant	1.740	2.477	2.719	3.137	1.548	1.663	2.652	2.590	2.281
Proxy	0.276†	0.344†	0.191†	0.154†	0.371†	0.366†	0.230†	0.239†	0.249†
LOS	0.090	0.038	0.056	-0.057	0.205**	0.025	0.249**	0.144*	-0.061
Gender	0.047	0.010	-0.009	-0.038	0.038	-.129*	-0.230†	0.015	-0.033
Age	0.004	-0.004	0.001	0.001	0.000	0.004	-0.006	-0.004	0.005
ADL	-0.021	-0.076†	-0.033	-0.031**	-0.027	-0.019	-0.030	-0.007	-0.020
Cognition	0.025*	0.055†	-0.052†	-0.018	0.045**	0.027	-0.057†	-0.017	-0.005
R^2	0.080	0.233	0.076	0.042	0.152	0.138	0.116	0.081	0.073

* means statistically significant at the level of p $<$.05.
** means statistically significant at the level of p $<$.01.
† means statistically significant at the level of p $<$.001.

The proxy ratings were used in a second way. For each type of respondent (resident, family, and staff), we created a mean score for each facility by averaging the proxy ratings for that facility. When these mean scores were compared by type of respondent, the levels of agreement were much higher than for the correlations at the individual level. Table 4.9 shows the Pearson correlations and ICC values for the mean respondent facility ratings for each applicable domain and the family member and staff means. The results are much stronger. Among the staff five of the possible 8 Pearson correlations and four ICC values are above 0.4 and all but one is statistically significant despite the much smaller sample size (N=40). For family members the pattern is stronger. For all but one domain (meaningful activities) the Pearson correlations and

the ICC values are 0.4 or greater.

Table 4.9.   Correlations of Facility Level Domain Scores Between Residents and Proxies

Domain	Residents vs. Staff				Residents vs. Family			
	Pearson	Sig.	ICC	Sig.	Pearson	Sig.	ICC	Sig.
Comfort	0.6355	0.0000	0.6352	0.0000	0.4136	0.0080	0.4134	0.0036
Functional Competency	0.4296	0.0057	0.3692	0.0088	0.6385	0.0000	0.5203	0.0002
Privacy	0.5863	0.0001	0.5539	0.0001	0.6125	0.0000	0.5488	0.0001
Dignity					0.5378	0.0003	0.5377	0.0001
Meaningful Activity	0.4053	0.0095	0.4024	0.0046	0.1267	0.4359	0.1247	0.2187
Enjoyment	0.2587	0.1069	0.2457	0.0607	0.5029	0.0009	0.5022	0.0004
Individuality	0.4523	0.0034	0.4522	0.0015	0.4951	0.0012	0.4882	0.0006
Relationships					0.5018	0.0010	0.4054	0.0043
Security	0.3317	0.0365	0.3307	0.0174	0.4337	0.0052	0.5504	0.0004
Autonomy	0.3162	0.0468	0.3069	0.0255	0.4501	0.0036	0.4317	0.0024

Discussion

Faced with respondents who cannot respond fully or reliably, researchers routinely turn to proxies. Reliance on proxies must depend on the nature of the data being sought and the way it will be used. If the goal is to reproduce the reports given by the actual respondents, then proxies can perform in only circumscribed areas. If the goal is simply to create an overall aggregated score across many respondents, then the need for specific individual level agreement can be relaxed.

It is one thing to use proxies to provide information about factual events, such as hospital admissions or falls. There the basic question is the opportunity to witness the event and the usual concerns about the accuracy of recall. It is quite another thing to rely on proxies to provide

information about what another person thinks or feels. The poor results reported here speak strongly to the need to be very careful in employing such approaches. Although the choice of correlation statistic should be based on the nature of the data being analyzed, in this study the results were similar regardless of the statistic used.

This low level of agreement between either family or staff proxies and residents has forced us to be extremely cautious in using either family or direct staff as proxies for resident reports when we have any possibility of getting feedback from residents themselves. We still intend additional work with both the family and staff reports to see if any items might be helpful for approximating QOL for residents for whom self-report is impossible, and to see if any family or staff characteristics better predict congruence with residents. We are undertaking some additional work with the family questionnaire to consider its properties in its own right, considering family members as themselves consumers on behalf of their relatives with dementia.

Our results were quite similar to what others have reported in other settings. The difference lies in the fact that what they celebrated, we mourn. Achieving statistical significance with correlation coefficients on the order of 0.3 or less does not mean that one can comfortably substitute a proxy report for that of a resident. When one thinks of these results in terms of variance explained, a correlation of 0.3 means explaining less than 10% of the variance in the residents' report; even a correlation of 0.4 explains only 16% of the variance. We, thus, conclude that one should use proxy reports of nursing home residents' QOL very cautiously.

Comparing aggregated mean values represents a less stringent test of proxies. Here the individual pairs need not agree, as long as one high or low score in one group offsets a similarly high or low score in the other. Thus, for creating an aggregate score for a nursing home, family proxies may work; but even in this more relaxed test, staff proxies did not fare well.

The generally poor performance of proxies who are customarily used as information sources should raise some alarms in other areas as well. The Minimum Data Set (MDS) mandated by CMS (HCFA) relies almost exclusively on observations and inferences from nursing home staff. In some cases measures of pain (Fries, et al. 2001) are created from these observations. Although there is a gross correlation with these generated measures and more typical measures of these phenomena, questions must arise about the ultimate validity of this approach.

References

Albert, S., Del Castillo-Castaneda, C., Sano, M., Jacobs, D., Marder, K., Bell, K., Blysma, F.,
Lafleche, G., Brandt, J., Albert, M., & Stern, Y. (1996). Quality of life in patients with
Alzheimer's disease as reported by patient proxies. *Journal of the American Geriatrics
Society, 44*, 1342-1347.

Andresen, E. M., Patrick, D. L., Carter, W. B., & Malmgren, J. A. (1995). Comparing the
performance of health status measures for older adults. *Journal of the American
Geriatrics Society, 43*(9), 1030-1034.

Brod, M., Stewart, A. L., Sands, L., & Walton, P. (1999). Conceptualization and measurement of
quality of life in dementia; The Dementia Quality of Life Instrument (DQoL). *The
Gerontologist, 39*(1), 25-35.

Capuzzo, M., Grasselli, C., Carrer, S., Gritti, G., & Alvisi, R. (2000). Quality of life before
intensive care admission: agreement between patient and relative assessement. *Intensive
Care Medicine, 26*, 1288-1295.

Curtis, A. E., & Fernsler, J. I. (1989). Quality of life of oncology hospice patients: A
comparison of patient and primary caregiver reports. *Oncology Nursing Forum, 16*(1),
49-53.

Dorman, P. J., & Waddell, F. (1997). Are proxy assessments of health status afters troke with the
EuroQOL Questionnaire feasible, accurate, and unbiased. *Stroke, 28*(10), 1883-1887.

Epstein, A. M., Hall, J. A., Tognetti, J., Son, L. H., & Conant, L. (1989). Using proxies to
evaluate quality of life. *Medical Care, 27*(3), S91-S98.

Fries, B. E., Simon, S. E., Morris, J. N., Flodstrom, C., & Bookstein, F. L. (2001). Pain in U.S.
nursing homes: Validating a pain scale for the Minimum Data Set. T*he Gerontologist,*

41(2), 173-179.

Greer, D. S., Mor, V., Morris, J. N., Sherwood, S., Kidder, D., & Birnbaum, H. (1986). An

 alternative in terminal care: Results of the national hospice study. *Journal of Chronic*

 Disease, 39(1), 9-26.

Hays, R. D., Vickrey, B. G., Hermann, B. P., Perrine, K., Cramer, J., Meador, K., Spritzer, K., &

 Devinsky, O. (1995). Agreement between self reports and proxy reports of quality of life

 in epilepsy patients. *Quality of Life Research, 4*, 159-168.

Hickey, E. M., & Bourgeois, M. S. (2000). "Health related quality of life (HR-QOL) in nursing

 home residents with dementia: Stability and relationships among measures. *Aphasiology,*

 14(5/6), 669-679.

Moinpour, C. M., & Lyons, B., et. al. (2000). Substituting proxy ratings for patient ratings in

 cancer brain metasteses: An analysis of a SOuthwest Oncology Group trial in patients

 with brain metasteses. *Quality of life research, 9*(2), 219-231.

Neumann, P. J., Araki, S. S., & Gutterman, E. M. (2000). The use of proxy respondents in

 studies of older adults: Lessons, challenges, and opportunities. *Journal of the American*

 Geratrics Society, 48, 1646-1654.

Novella, A. M., & Jochum, c. (2001). Agreement between patients' and proxies' reports of

 quality of life in Alzheimer's patients. *Quality of life research, 10*(5), 443-452.

Page, S. A., & Verhoef, A. M. (1995). Quality of life, bypass surgery, and the elderly. *Canadian*

 Journal of Cardiology, 11(9), 777=782.

Sigurdardottir, V., & Brandberg, Y. (1996). Citerion-based validation of the EORTC QLQ-C36

 in advanced melanoma: THe CIPS quesitonnaire and proxy raters. *Quality of life*

research, 9(10), 1117-1126.

Sneeuw, K. C., Aaronson, N. K., Sprangers, M. A., Detmar, S. B., Wever, L. D., & Schornagel, J. H. (1998). Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *Journal of Clinical Epidemiology, 51*(7), 617-631.

Sneeuw, K. C. A., Aaronson, N. K., Osoba, D., Muller, M. J., Hsu, M.-A., Yung, W. K. A., Brada, M., & Newlands, E. S. (1997). The use of significant others as proxy raters of the quality of life of patients with brain cancer. *Medical Care, 35*(5), 490-506.

Stancliffe, R. J. (1999). Proxy respondents and the reliability of the Quality of Life Empowerment factor. *Journal of Intellectual Disability Research, 43*(Pt. 3), 185-193.

Wilson, K. A., Dowling, A. J., Abdolell, M., & Tannock, I. F. (2000). Perception of quality of life by patients, partners and treating physicians. *Quality of Life Research, 9*, 1041-1052.