Predicting Nursing Facility Residents' Quality of Life Using External Indicators

Howard B. Degenholtz, Rosalie A. Kane, Robert L. Kane, Boris Bershadsky, and Kristen C. Kling

Purpose. A newly developed brief measure of nursing facility (NF) resident self-reported quality of life (QOL) has been proposed for inclusion in a modified version of the minimum data set (MDS). There is considerable interest in determining whether it is possible to develop indicators of QOL that are more convenient and less expensive than direct, in-person interviews with residents.

Design and Methods. QOL interview data from 2,829 residents living in 101 NFs using a 14-item version of a longer instrument were merged with data from the MDS and the Online Survey and Certification Automated Record (OSCAR). Bivariate and multivariate hierarchical linear modeling were used to assess the association of QOL with potential resident and facility level indicators.

Results. Resident and facility level indicators were associated with self-reported QOL in the expected direction. At the individual resident level, QOL is negatively associated with physical function, visual acuity, continence, being bedfast, depression, conflict in relationships, and positively associated with social engagement. At the facility level, QOL is negatively associated with citations for failing to accommodate resident needs or providing a clean, safe environment. The ratio of activities staff to residents is positively associated with QOL. This study did not find an association between QOL and either use of restraints or nurse staff levels. Approximately 9 percent of the total variance in self-reported QOL can be attributed to differences among facilities; 91 percent can be attributed to differences among residents. Resident level indicators explained about 4 percent of the variance attributable to differences among residents, and facility factors explained 49 percent of the variance attributable to differences among NFs. However, the different variables explained only 10 percent of the variance in self-reported QOL.

Implications. A brief self-report measure of NF resident QOL is consistently associated with measures that can be constructed from extant data sources. However, the level of prediction possible from these data sources does not justify reliance on external indicators of resident QOL for policy purposes.

Key Words. Nursing facilities, long-term care, elderly, health policy, hierarchical modeling, quality of life

Federal and state officials have considerable interest in determining whether it is possible to develop indicators of the quality of life (QOL) of nursing facility (NF) residents that are more convenient and less expensive than direct, in-person interviews with residents, and even to develop indicators that do not require visits to each facility. Such indicators would be analogous to widely used indicators for quality of care (Zimmerman et al. 1995; General Accounting Office 2002; Mor et al. 2003a, 2003b; Mukamel and Spector 2003), which do not directly measure clinical processes or outcomes, but are widely used for quality improvement, benchmarking, and public reporting. Under a contract with Centers for Medicare & Medicaid Services (CMS), we developed a valid and reliable instrument for measuring self-reported QOL among NF residents (Kane et al. 2003) that also can distinguish among nursing homes (Kane et al. 2004b). We next sought to determine if a set of variables extracted from existing data systems on U.S. nursing facilities could be used as “external indicators” of this measure of resident self-reported QOL.

BACKGROUND

The QOL of the approximately 1.5 million NF residents in the U.S. is undoubtedly lower than desired by residents, families, providers, and policy makers. Although a report (Institute of Medicine 1986) and subsequent 1987 legislative reforms placed high priority on QOL, much of the regulatory focus has been geared toward assessing and ensuring quality of care. Because QOL is subjective and personal, it is best collected from nursing home residents, but such data collection is tedious and expensive. There is thus great pressure to use extant measures.

Two main sources of data on NF residents are currently available to regulators: the Online Survey and Certification Automated Record (OSCAR) and the NF Resident Assessment Instrument (RAI). The OSCAR dataset

Address correspondence to Howard B. Degenholtz, Ph.D., Associate Professor, Department of Health Policy and Management, Graduate School of Public Health, Center for Bioethics and Health Law, University of Pittsburgh, 3708 5th Avenue, Suite 300, Pittsburgh, PA 15213. Rosalie A. Kane, Ph.D., Professor, and Robert L. Kane, MD, Professor, are with the Division of Health Services Research and Policy, University of Minnesota School of Public Health, Minneapolis, MN. Boris Bershadsky, Ph.D., Director, is with the Department Orthopaedic Surgery/A41, Orthopaedic Clinical Research Center, Cleveland Clinic Foundation, Cleveland, OH. Kristen C. Kling, Ph.D., Associate Professor, is with the Department of Psychology, St. Cloud State University, St. Cloud, MN.

captures the result of the state survey and certification process. All state licensed and certified NFs submit data on characteristics of their facilities and staffing and are inspected on a 9–15 month cycle as a condition of participation in Medicare and Medicaid. Failure to meet federal requirements in any of the 185 categories, based on observations of residents by state surveyors, can lead to sanctions based on the scope and severity of the problem.

The second source of data comes from the federally mandated RAI (Morris, Murphy, and Nonemaker 1995), which NFs must use to collect data at regular intervals on all residents. These data derived from this instrument, commonly referred to as the minimum data set (MDS), are submitted to state licensure and certification agencies and subsequently to the Centers for Medicare & Medicaid Services. The MDS data, which are based on observations of residents by facility staff, provide a detailed picture of the health status of NF residents, and have been used to construct indicators of the quality of the residents' care (Zimmerman et al. 1995).

The OSCAR and the current MDS systems do not address resident QOL in depth or breadth. The inspection process (generally known as the survey process) focuses on the experience of a small number of residents. Although surveyors use a structured interview tool that addresses aspects of QOL, determination of deficiencies is left to the surveyor's judgment. Also, the survey process does not lead to a quantitative score that can be used to compare facilities. The MDS was designed as a tool for care planning and mainly covers physical function, cognitive function, and health care needs. For example, one section of the MDS addresses social engagement (Mor et al. 1995), but this hardly covers the complex multidimensional concept of QOL (Lawton 1983). Furthermore, as with the rest of the MDS, the social engagement score is based on staff assessment rather than resident self-report.

The recognition that existing data sources on NF residents fail to address QOL with adequate depth, breadth, or validity has led to the development of new measures for use with residents capable of providing self-report (Kane et al. 2003). No instrument has been adopted as yet for national use, but a subset of items from this newly developed measure is currently being considered for the next version of the NF MDS (Centers for Medicare & Medicaid Services 2003).

In this study, we sought to determine how well a set of variables theoretically related to QOL that are already collected and archived by CMS as part of the existing NF data infrastructure (i.e., MDS and OSCAR) are correlated with a measure of resident self-reported QOL (Kane et al. 2003). An indicator of QOL based on these "external" variables could potentially

provide a relatively inexpensive way to screen facilities for quality improvement or other policy purposes.

METHODS

We drew on the work of Lawton (1983) to identify relevant explanatory variables for the predictive model. As noted by Lawton, residents' perceptions of their own QOL can be considered a primary outcome of the clinical care, housing, and other services that are provided in a NF. We therefore sought to identify measures of residents' physical and psycho-social well-being and measures of the NF in which they live from extant data sources. We first examined bivariate relationships among QOL and candidate explanatory variables, then we examined the contribution of each set of variables toward explaining self-reported QOL using hierarchical linear modeling (HLM).

Data

The main data for this study came from two waves of interviews conducted with a multistate sample of NF residents using a new QOL tool (Kane et al. 2003). Further details on wave 1 data collection process and the instruments are presented elsewhere (Kane et al. 2003). The first wave was collected in 1999 and 2000 in 40 NFs in five states (California, New York, New Jersey, Minnesota, and Florida). The second wave was collected in 2001 in a new sample of 61 facilities in five states (Maryland was used in wave 2 instead of New Jersey). For wave 1, NFs were selected to represent a range of sizes and both urban and rural areas in each state. We oversampled facilities with a high proportion of single rooms, and we selected facilities in each state that had been nominated as having "good" QOL. A sample of 50 residents was selected at random from each facility and an interview was attempted with each regardless of cognitive ability. Residents who refused or were ineligible (i.e., under 65, comatose, did not speak English, were in the hospital, or were away from the facility for an extended period during the data collection window) were replaced. Only 2 percent of those approached refused to participate. In two smaller facilities with less than full occupancy it was not possible to reach 50; thus the wave 1 sample was composed of 1988 residents. There was an association between cognitive function and ability to complete the survey. Among residents with poor cognitive function (score of 4–5 on a 0–5 scale [details below]), about 38 percent were able to complete three-fourths of the QOL instrument. Among those with better cognitive function (score of 0, 1, or

2), about 82 percent were able to complete three-fourths of the QOL instrument. The dependent variable for this study uses only 14 items from the overall instrument, thus we were able to compute a QOL score for 1,162 (59 percent) of residents sampled (about 29 per facility).

In wave 2 a somewhat different sampling approach was used to generate a sample of facilities with both high and low expected levels of QOL. As we hypothesized that better QOL is associated with higher staffing ratios and fewer citations, we identified facilities with a broad range of both staffing ratios and numbers of citations. A sample frame of metropolitan regions within each of the study states was identified. A total of 89 facilities were approached to reach our planned sample size of 60 facilities. Ten facilities were determined to be ineligible for various reasons: too few elderly residents, too few residents who spoke English, or the facility specialized in psychiatric care. There were 18 refusals (23 percent) among the facilities asked to participate. One facility with a small number of elderly was replaced, but the data were retained in the analysis. Thus the wave 2 sample includes 61 facilities.

During wave 2, the goal was to obtain complete interviews for a sample of 28 randomly selected residents within each facility, based on analysis of the minimum number of observations required to estimate a stable facility level score. A total of 3,333 residents were approached. The same exclusion criteria were used as in wave 1. Residents who could not be roused or sustain a simple conversation with any coherence were dropped from the sample without attempting an interview (1,471; 44 percent of those approached). Attempted interviews were aborted if residents were unable to provide usable responses to four out of the first six items on the interview (98; 3 percent of those approached). There were 62 refusals (2 percent of those approached); which was comparable with wave 1. The final wave 2 sample had 1,688 residents (51 percent of those approached); because of missing data, we were able to compute QOL scores for 1,667 (99 percent) residents. The combined sample (waves 1 and 2) yielded 2,829 residents.

The QOL interview data were linked with MDS files provided by CMS. A full MDS assessment of each resident is done at admission and annually, and a partial assessment is done each quarter or upon a significant change of status. For each resident in the sample, the most proximate MDS assessment record to the interview was extracted and linked to the interview data. The MDS data were used to construct independent variables at the resident and facility level. We were able to link MDS data for 1,981 residents in wave 1 (seven were discharged between the time of the QOL interview and when their admission assessment was due), and all wave 2 residents.

QOL interview data were also linked with facility data from the OSCAR file, also provided by CMS. Two types of data were extracted from the OSCAR file. The first type was data on the number of full-time equivalent (FTE) of several categories of NF staff. These data are submitted by facilities as part of the state certification process. The second type of data was citations the facility received on its most recent survey for QOL-related problems.

The basic analytic design for this study examines the relationship between the dependent variable, case-mix adjusted resident self-reported QOL, and several sets of independent variables derived from the MDS and OSCAR data. The independent variables include:

- resident level case-mix adjusters (MDS);
- resident level potential indicators of QOL (MDS);
- aggregate measures of resident characteristics for all residents in sampled facilities (MDS);
- measures of NF staffing (OSCAR); and
- survey sanction data (OSCAR).

The latter three categories were classified as NF characteristics.

Self-Reported QOL

The dependent variable was a shortened version of the multidimensional resident self-report QOL instrument developed specifically for use in NFs (Kane et al. 2003). The instrument was developed through a literature review, pilot testing, and review by stakeholders. The full version contains multiple item scales for 11 dimensions of life. The dimensions address concrete (meaningful activities, comfort, enjoyment, security, privacy), and more abstract (relationships, individuality, spiritual well-being dignity, autonomy, and functional competence) aspects of life in a NF. For this study we were interested in modeling facility level differences in overall QOL across multiple domains, rather than as 11 separate scores. We therefore constructed a unidimensional summary scale by selecting a subset of 14 items with high intraclass correlations (Kane et al. 2004a). Two items with the highest intraclass correlation (i.e., responses from residents in the same NF are correlated) were drawn from each of seven dimensions (meaningful activities, enjoyment, security, privacy, relationships, individuality, and spiritual well-being) of the original instrument (see online Appendix 1). These dimensions were selected because the remaining four dimensions (dignity, autonomy, functional competence, and comfort) did not discriminate among facilities (Kane et al. 2004b). The

resulting 14-item scale (QOL-14) has an α reliability of 0.76, intraclass correlation of 0.06, and is moderately correlated (0.5–0.69) with the original 11 subscales. A confirmatory factor analysis supported a single dimension structure (RMSEA = 0.028, 95 percent CI: 0.02–0.036; further details available upon request).

Resident Case-Mix Adjustment

We used a number of resident variables to adjust for differences between facilities on resident characteristics that may be related to QOL. We measured resident age in years, gender, and length of stay. Length of stay was measured with a dummy variable for stays greater than 90 days. We note that residents may experience deterioration or rehabilitation and improvement in the period shortly after admission. Although a dummy variable is a crude way to capture these changes, we explored other cut points and continuous measures (e.g., days from admission) and found no difference in the pattern of inferences.

An index of physical function based on level of independence in eating, dressing, toileting, transferring, and walking was computed using magnitude estimation weights (Finch, Kane, and Philp 1995). As a further measure of physical function, an indicator of whether the resident is bedfast most of the time was also included. In addition, poor physical health is expected to be associated with worse self-reported QOL. We therefore also included measures of visual and hearing capacity, and urinary and fecal continence from the MDS.

Cognitive function, while central to our understanding of the overall health of NF residents, is not clearly linked in a positive or negative direction to self-reported QOL. While the argument may be made that people with poor cognitive function have objectively lower ability to appreciate and take pleasure from the external world, from the perspective of the person with cognitive impairment, the awareness of that impairment may not be present. Thus, their perception of their environment and condition may be generally positive. A cognitive function scale was computed based on short- and long-term memory and cognitive skills for daily decision-making items. This scale correlates very highly with the cognitive performance scale (Morris et al. 1994); however, it does not confound physical function (i.e., eating) with cognitive function.

Resident Level Potential Indicators of QOL

Data from the MDS were used to construct several measures of psycho-social well-being hypothesized to be associated with QOL. The first was the social

engagement scale described by Mor et al. (1995), Schroll et al. (1997), Casten et al. (1998), and Lawton et al. (1998). This scale measures social involvement, including participation in activities. The behavioral symptoms scale is based on items such as wandering and verbal or physical abuse (Mor et al. 1995; Casten et al. 1998; Lawton et al. 1998; Snowden et al. 1999). Conflict in relationships measures interpersonal conflict with staff, other residents and family or friends. (A similar scale was labeled “social quality” [Casten et al. 1998; Lawton et al. 1998].) Depressed mood was measured using the depression rating scale (Burrows et al. 2000). This scale measures seven symptoms of depression and is correlated well with other measures of depressive symptoms. Finally, the use of physical restraints was also included as a potential indicator of poor QOL. Residents were identified who had daily trunk or limb restraints or a chair that prevents rising.

NF Level Potential Indicators of QOL

We constructed three sets of measures at the NF level: aggregate measures of resident health and psycho-social well-being, care staff inputs, and evidence of poor QOL in that facility in the past. First, to test the hypothesis that residents in facilities with better aggregate clinical and psycho-social outcomes will have better personal QOL, we therefore constructed measures of the average physical health and psycho-social well-being of all residents within each facility in the study. Physical health was measured by computing the average physical function and cognitive function (using the same scoring as at the individual level) of all residents living in the facility at the time the QOL interviews were conducted. In addition, we computed the prevalence of certain commonly used indicators of clinical quality of care that are hypothesized to also reflect a poor QOL (Zimmerman et al. 1995): depression, bladder or bowel incontinence, indwelling catheter, weight loss, tube feeding, bedfast residents, daily restraints, and pressure ulcers (among low-risk residents). Psycho-social well-being was measured by computing facility average scores for social engagement, distressed mood, conflict in relationships, and behavioral problems (using the same scoring as at the individual level) of all residents living in the facility at the time the QOL interviews were conducted.

Second, to test whether a higher ratio of staff to residents results in more personal attention and opportunity for personalized care, thus increasing the likelihood of a resident experiencing a good QOL, we computed the ratio of staff to residents for several key categories thought to be related to QOL: certified nursing aides, registered nurses, licensed practical nurses, occupational

therapists, physical therapists, social workers, dietary, housekeeping, and administrative staff. Activity staff and recreational therapists were combined because facilities use these personnel categories interchangeably. Staffing ratios were based on the sum of all full time employees, part time employees, and contract FTEs reported on the most recent OSCAR report for each facility. The number of FTEs per 100 residents was computed by dividing by the sum of FTEs by number of residents living in the facility, and multiplying by 100.

Staffing data reported by facilities are notoriously error prone, containing both implausibly high and low (zero) values (Harrington et al. 2000). Like others, we removed any values that implied a staff to resident ratio of 1–1 or higher and also eliminated the top 2 percent. In order to retain all 101 facilities in our sample we had to replace these values with plausible figures. Where possible, we used data from the previous or subsequent OSCAR record. If this was not available, we used the median value for facilities in the same state, stratified by whether the facility is for profit or nonprofit and whether it is certified as a Medicare Skilled Nursing Facility (SNF). Data were replaced for a total of eight cases. The mean and range was not affected by this procedure. All analyses were conducted on both the complete dataset and a restricted dataset without imputed data. Given that there were no significant differences in the magnitude or pattern of inferences between these two sets of analysis, results are based on the full dataset. We found no significant differences between the staffing levels of the sampled facilities and the median for all other facilities in the 6 study states, stratified by profit status and Medicare SNF certification.

Third, recognizing that some facilities have persistent problems with quality (Castle 2002; Grabowski and Castle 2004), and that those who have faced sanctions from state inspectors in the past are at increased likelihood of having repeated problems in related areas, we developed measures of prior regulatory performance. We identified 54 categories of citations in the Resident Rights section of the regulations that we hypothesized would be related to QOL (see online Appendix 2). Only 20 of these were sufficiently populated in the OSCAR system to be useful. (We used a minimum threshold of 400 observations nationally out of 16,481 facilities [prevalence of ~ 2.4 percent] as the cutoff.) Records of facility citations were extracted from the OSCAR system for these 20 categories for all facilities in the six study states. Up to four surveys for each facility were used. Five out of the 20 citations were dropped from further consideration because they lacked a statistically significant intraclass correlation. To adjust for state differences in how deficiencies are assigned, we used a two level hierarchical model to estimate the predicted probability of having a given citation using facility as the higher level and each

survey as the lower level. The only independent variables in this analytic step were state dummies. All facilities in each of the six study states were used for this model to take advantage of the full distribution of citations in each state. The predicted probability of receiving each citation (for each facility, adjusted for state) was used in the final models as an independent variable, rather than whether a given facility had received a citation on its most recent survey. (The prevalence of each citation category on the most recent survey is shown in Table 2 for ease of interpretation.)

Analysis

The analysis was conducted in two main steps. First, bivariate correlations between QOL and candidate explanatory variables were calculated. Data were examined separately at the resident and facility levels to ensure that the appropriate unit of analysis was used for each set of explanatory variables. The correlations between QOL and resident characteristics were done at the resident level. Resident health status and psycho-social well-being were also computed at the facility average level to compute the correlation of those variables with facility average QOL. Other NF characteristics (i.e., staffing levels and citations) were inherently measured at the facility level.

We distinguish between resident case-mix variables, which are characteristics of the sampled residents that may affect their QOL, and measures that we consider potential indicators of QOL. Case-mix variables adjust for differences between facilities in the type of residents they serve and sampling variability. Measures of resident psycho-social well-being and the use of restraints are included in the model because they are potential indicators of QOL.

Second, we used multivariate methods to assess the relative amount of variation accounted for by NF characteristics after adjusting for resident factors. Two methodological issues needed to be addressed. It was necessary to allow for the potential that data from residents sampled from within the same NF were correlated. An intraclass correlation of 0.05 can significantly bias the standard errors (Murray and Hannan 1990; de Leeuw and Kreft 1995). QOL was measured at a different level of aggregation than the objective environment, creating a potential ecological fallacy. To address both of these problems, we used HLM, which takes into account the nested nature of the data, avoids the ecological fallacy, and provides correct estimates of the standard errors (Bryk and Raudenbush 1992; Goldstein 1995). Analyses were conducted using SAS PROC MIXED (Singer 1998) and *MLWiN* (Goldstein

1998). All continuous variables were standardized (using z scores) for analysis; descriptive statistics were calculated using raw scores. In the interests of parsimony, only variables with bivariate correlations with $p < .10$ were included in the multivariate analysis.

FINDINGS

Descriptive Statistics—Residents

Residents who had a QOL interview ranged in age from 65 to 109; the mean was 84 (SD 8.1) (Table 1). The majority of residents were female (74 percent) and white (89 percent). Most residents (82 percent) had lived in the NF for

Table 1: Resident Variables

<i>Variable</i>	<i>Mean (SD) or %</i>
Self-reported QOL*	3.07 (0.49)
Demographics	
Age	84 (8.07)
Gender (female)	74%
Race (white)	89%
Length of stay (> 90 days)	82%
Health status	
Physical function [†]	0.98 (1.52)
Cognitive function [‡]	2.2 (1.64)
Bladder incontinence (H1a: frequently or completely incontinent)	45.5%
Bowel incontinence (H1b: frequently or completely incontinent)	31%
Vision (D1: moderately, highly, or severely impaired)	14%
Hearing (C1 moderately, highly, or severely impaired)	8%
Restraints [§]	4.7%
Bedfast (G6a)	2.4%
Psycho-social well-being	
Social engagement	2.6 (1.6)
Distressed mood	0.83 (1.5)
Conflict in relationships	0.10 (0.40)
Behavioral problems	0.58 (1.4)

Notes: $n = 2,829$.

*Fourteen-item scale based on seven domains of QOL; higher score indicates better QOL; range from 1 to 4.

[†]Magnitude Estimation Scale based on independence in: eating, toileting, transfer, dressing, and walking; higher indicates greater disability; range from 0 to 4.86.

[‡]Based on short-term memory, long-term memory, and judgment items; higher score indicates greater disability; range 0–5.

[§]Daily use of trunk or limb restraints, or chair that prevents rising.
QOL, quality of life.

over 90 days. Cognitive function was measured on a six-point scale, ranging from 0 (to no memory or decision-making problems; 23.8 percent), to 5 (severe impairment; 5.2 percent). The mean level of cognitive impairment was 2.2 (SD 1.6). Physical function was measured on a scale that ranged from 0 (no limitations), to 4.855 (complete assistance required in five ADLs). The average level of physical impairment was 0.99 (SD 1.52); 7 percent had no impairment and 5.5 percent were completely impaired. There were 4.7 percent of residents in physical restraints and 2.4 percent were bedfast.

Nearly half of the sample (45 percent) was frequently or completely incontinent of bladder, and 30 percent were frequently or completely incontinent of bowel. About 14 percent were moderately, highly, or severely impaired in vision and 8 percent were moderately, highly, or severely impaired in hearing. Social engagement, which ranges from 0 to 6, had a mean of 2.6 (SD 1.6). Distressed mood had a mean of 0.84 (SD 1.5); 64 percent had a score of 0 on this 14-point scale. The conflict in relationships scale averaged 0.10 (SD 0.40), and behavioral problems averaged 0.58 (SD 1.4).

The dependent variable, QOL, ranged from 1 to 4; mean 3.1 (SD 0.49). A score of 4 represents a response of “often” to all 14 items, indicating good QOL, and a score of 1 represents a response of “never” to all 14 items, indicating poor QOL. The mean response implies that most residents rated their QOL as “sometimes” having their expectations met with regard to the items covered in the scale.

Descriptive Statistics—Facilities

Facilities ranged from 50 to 668 residents; mean 156 (SD 87). The total number of residents at the 101 facilities in the study was 15,715. About 17 percent of facilities were for-profit. At the facility level, the average cognitive function score was 3.2 (SD 0.5) (Table 2). The average physical function score was 3.3 (SD 0.35), and the average age was 87 (SD 3).

Comparison of Interviewed Residents with Population

Compared with the overall facility population, the sample of residents who participated in the QOL interview was less physically disabled ($t = 16.19$; $p = .000$); less cognitively disabled ($t = 16.95$; $p = .000$); and slightly younger ($t = 9.56$; $p = .000$). There were fewer long-stay residents ($t = 33.92$; $p = .000$) and slightly fewer residents in the sample had moderate or severe urinary incontinence ($t = 2.06$; $p = .039$). There was no difference in the gender, race, or level of visual impairment.

Table 2: Nursing Facility Factors

<i>Variable</i>	<i>Mean (SD) or %</i>	<i>Min</i>	<i>Max</i>
Health status*			
Average physical disability [†]	3.3 (0.35)	2.5	4.0
Average cognitive disability [†]	3.2 (0.45)	2.0	4.4
Prevalence of depression (QI-4) (%)	3.3	0	25.3
Prevalence of bladder or bowel incontinence (QI-8) (%)	54.5	33	84.4
Prevalence of indwelling catheter (QI-10) (%)	7.6	0	24.1
Prevalence of weight loss (QI-13) (%)	13.0	2.5	24.8
Prevalence of tube feeding (QI-14) (%)	9.2	0	29.9
Prevalence of bedfast residents (QI-16) (%)	5.8	0	35.8
Prevalence of restraints (QI-22) (%)	10.9	0	41.0
Prevalence of stage 1–4 pressure ulcers among low-risk residents (QI-24) (%)	4.0	0	20.8
Psycho-social well-being			
Average social engagement	1.95 (0.7)	0.15	3.99
Average distressed mood	0.45 (0.36)	0.02	1.8
Average conflict in relationships	0.10 (0.12)	0	0.9
Behavioral problems	0.96 (0.79)	0.11	4.87
Staff levels [§] (FTE/100 resident days)			
Certified nurses aides	39 (9.68)	8.45	69.72
Registered nurses [¶]	6.55 (3.74)	1.43	28.83
Licensed practical nurses	12.2 (4.47)	1.69	28.78
Occupational therapists	0.9 (0.8)	0	5.01
Physical therapists	1.75 (1.42)	0	7.18
Social work	1.16 (0.74)	0	3.3
Dietary staff	0.52 (0.56)	0	4.11
Housekeeping	10.07 (3.78)	1.46	19.95
Activities/recreation staff	3.49 (1.66)	0.31	10.47
Administrative staff	10.48 (4.68)	2.48	29.97
QOL-related deficiencies			
F157 notice of changes	9.9%		
F159 management of personal funds	3.0%		
F166 resolve grievances	6.9%		
F176 self-administer drugs	5.9%		
F221 physical restraints	19.8%		
F223 abuse	3.0%		
F224 staff treatment of residents	5.9%		
F241 dignity	24.8%		
F242 self-determination	5.0%		
F246 accommodate needs	16.8%		
F248 activities program	9.9%		
F250 social services	16.8%		
F252 clean, safe, comfortable environment	8.9%		
F253 housekeeping	22.8%		
F280 comprehensive care plan	8.9%		

Notes: $n = 101$.

*Facility average score based on all residents in facility at time of data collection.

[†]Magnitude Estimation Scale based on independence in: eating, toileting, transfer, dressing, and walking; higher indicates greater disability.

[‡]Zero to five scale based on short-term memory, long-term memory, and judgment items from the MDS; higher score indicates greater disability.

[§]FTE per 100 residents, including recreation therapist (full time, part time, contract)

[¶]Excluding nurses with administrative duties.

^{||}Percent of facilities receiving citation on most recent survey.

QOL, quality of life; FTE, full-time equivalent; MDS, minimum dataset.

Multivariate Modeling

Variables statistically significant in the bivariate analysis were entered into the hierarchical linear model (Table 3). Resident case-mix factors that were significant in the multivariate model were physical function, visual impairment, being bedfast, and length of stay. Resident level indicators of QOL were social engagement, depression, and conflict in relationships. NF characteris-

Table 3: Hierarchical Linear Model of Resident and Facility Indicators of QOL

	<i>B</i>	<i>p-value</i>
Resident case-mix adjustment		
Gender (female)	0.092	.028
Long stay (over 90 days)	0.164	.001
Physical disability	− 0.073	.004
Bowel incontinence (H1a)	− 0.021	.160
Urinary incontinence (H1b)	− 0.003	.853
Visual impairment (D1)	− 0.038	.049
Bedfast (G6a)	− 0.247	.044
Resident-level indicators		
Restraints	− 0.042	.633
Social engagement	0.155	<.0001
Distressed mood	− 0.049	.000
Conflict in relationships	− 0.055	.003
Facility-level indicators		
Average physical disability	0.012	.892
Prevalence of depression (QI-4)	0.014	.064
Prevalence of indwelling catheter (QI-10)	0.000	.951
Prevalence of weight loss (QI-13)	− 0.099	.035
Prevalence of tube feeding (QI-14)	− 0.002	.549
Prevalence of bedfast residents (QI-16)	− 0.010	.061
Prevalence of restraints (QI-22)	0.000	.853
Average social engagement	− 0.034	.204
Average distressed mood	0.000	1.000
Average behavioral problems	0.038	.176
Social worker	0.020	.394
Dietary	− 0.043	.075
Activities/recreation	0.052	.030
Housekeeping	− 0.069	.002
F224 staff treatment of residents	− 0.028	.222
F246 accommodate needs	− 0.060	.014
F250 social services	− 0.020	.386
F252 clean, safe, comfortable environment	− 0.050	.022
F253 housekeeping	0.010	.663
− 2 log likelihood		7,649.8
Bayesian information criterion (BIC)		7,802.1

tics significant in the final multivariate model were the prevalence of depression, indwelling catheter, weight loss, and bedfast residents. The ratios of activities/recreation staff and housekeeping staff were significant, as were two categories of QOL citation: accommodating resident needs and providing a clean, safe comfortable environment.

The total variance in resident self-reported QOL was partitioned into the portion attributable to residents (91 percent) and the portion attributable to facilities (9 percent). The resident case-mix adjustment part of the model explained about 2 percent of the total variance in self-reported QOL; resident level indicators of QOL explained about 4 percent. The variance attributable to differences between facilities was reduced by 49 percent after adding NF characteristics to a model containing only resident factors (not tabled); however this represents only about 5 percent of the overall variance in resident self-reported QOL.

DISCUSSION

Previous research established the validity of a new measure of resident self-reported QOL (Kane et al. 2003). This paper extends that work, examining whether data in the MDS and OSCAR systems can be used to develop a set of "external indicators" of NF residents' QOL. Although we found that NF characteristics explained nearly half of the between facility differences in resident QOL, only about 10 percent of the total variance was explained by the full model. This suggests that a predictive or screening model that relied only on external indicators would do a poor job of identifying facilities with high or low QOL. Thus, we argue that the goal of improving QOL is best served by direct measurement.

The pattern of findings is consistent with the theory that self-reported QOL is related to, but conceptually distinct from other measurable factors of the person and their environment. Our findings are in keeping with research on quality of care in nursing facilities. For example, a number of studies have found that different quality of care outcomes are typically poorly correlated with one another (i.e., on the order of 0.10–0.15) (Mukamel and Brower 1998; Porell and Caro 1998; Mukamel and Spector 2000; Mor et al. 2003c), which suggests that different outcomes reflect distinct aspects of clinical care and practice. "Good" nursing facilities do not necessarily perform well across different domains (Mukamel and Spector 2003). Similarly, risk adjustment models for quality outcomes in the NF setting (Mukamel et al. 2003), and in other situations such as Medicare expenditures (Newhouse, Buntin, and Chapman 1997), typically explain no more than 20–25 percent of the variance. Thus,

while low, the variance explained by our model is not unusually so, and reinforces the conclusion that QOL is a distinct aspect of NF performance.

Two important conclusions flow from the observation that resident self-reported QOL cannot be predicted with much confidence from objective data about either residents themselves or the facilities in which they live that have important implications for policy and regulation. First, it likely is not possible to develop a simple screening formula derived from extant data sources that will perform well enough for policy or regulatory purposes. The data for this study come from a database developed to test the new QOL instruments, thus restricting the generalizability of our model. However, we argue that even if resident QOL data were collected from a new nationally representative sample of nursing facilities the results would be similar.

Second, we note that even after selecting items for our dependent variable that maximize between facility discriminatory power, the majority of variation remains within facilities. Much of the variance at the individual level was explained by demographic and health status factors that are not modifiable. A pessimistic conclusion may be that there is not much facilities can do to improve resident QOL. However, even after adjusting for resident factors, we still find differences among facilities. One interpretation of this finding is that even though facility average QOL may be relatively low in a particular facility, some residents in that facility may have high QOL and vice versa. This observation suggests that facilities can potentially affect QOL, even after adjusting for case mix. That is, the substantial amount of unexplained variance at the resident level is potentially amenable to improvement. However, any successful intervention will need to be "close" to the individual. For example, in addition to facility-wide programs or policies, an assessment and care planning model could be developed to promote individualized attention to QOL. Admittedly, the evidence base for improving resident QOL is lacking, in part because of the absence of a valid, reliable measure. Additional research is needed to determine that our new measures of QOL are responsive to theoretically driven interventions.

Researchers and policy makers tend to use facility characteristics as a signal of the QOL in that facility. For example, the prevalence of physical restraint use is often taken to be an indicator of QOL. Although we found that individuals who are regularly restrained report lower QOL, in the multivariate analysis, we found no association between resident QOL and the facility-level prevalence of restraints. The movement to reduce the use of physical restraints is clearly in the service of improving QOL (Institute of Medicine 1986) and reducing injuries (Neufeld et al. 1999), and this finding does not diminish the importance of that goal. However, our findings suggest that one cannot con-

clude that a facility with a low level of restraints has good QOL, or that the reverse is true. It is important to note that in our study, residents who were capable of providing self-report data were less likely to be in restraints than those who were not (4.7 percent of respondents were in restraints compared with 14 percent among nonrespondents). Just as self-reported QOL data cannot be used to extrapolate to those incapable of self-report, the prevalence of restraints (which differentially affects those incapable of self-report) should not be taken as an indicator of the QOL of those who are capable of providing insight into their own experience. The policy and regulatory process needs to go further and dig deeper to assure that NF residents are treated with respect and dignity and are given the opportunity for a meaningful existence.

In our previous work to develop measures of resident QOL, we presented a multidimensional instrument that addresses a broad range of issues relevant to the lived experience. For this study we identified a subset of items that form a unidimensional scale (the QOL-14) that discriminates between facilities and is highly correlated with the parent measure. Our results using this short-form were consistent with separate analyses (not reported) conducted on each of the original scales. The QOL-14 is similar to a set of items being considered for inclusion in the next version of the MDS (Centers for Medicare & Medicaid Services 2003; Kane et al. 2004a). Thus, our findings suggest the level of prediction that might be possible using facility factors to predict resident QOL in the MDS 3.0. Inclusion of QOL on the MDS is an important step toward making the subjective experience of residents a focus of individual care planning and facility quality improvement. Thus, our findings using this restricted set of items further underscore the importance of direct measurement rather than reliance on facility characteristics as a signal. It is important to recognize, however, that the QOL-14 addresses only a subset of QOL domains with only two items per domain, and thus cannot serve as an individual diagnostic or care planning tool. Providers need to take a low score (either at the individual or facility level) as an indication that further investigation of underlying factors is warranted. Nevertheless, availability of a brief instrument will make it possible for providers and policy makers to have access to a measure of resident QOL. This has the potential to place needs of NF residents for individual meaning, enjoyable activities, stimulating social contact, respect, and security higher on the agenda.

Several limitations to this study should be noted. First, there is no “gold standard” for self-reported QOL in the NF setting. We have relied on our previous work to develop and test self-report instruments, and have used theory to identify potential indicators from “external” data sources. The

observed pattern of observations generally supports the validity of the measure, however the strength of the underlying associations was low. This may be because of differences between data collection for the MDS and OSCAR, which are documented by facility staff, and QOL, which was collected by trained research interviewers. The MDS is conducted every 90 days (or after a significant change in status); the average gap between the QOL interview and the MDS assessment was about 45 days. Thus, resident characteristics may have changed between the MDS assessment and the interview.

Second, the data come from two cross-sectional samples of NFs selected with somewhat different sampling criteria. Both samples were designed to identify facilities demonstrating a range of resident QOL in order to test the discriminatory ability of the new instrument. Our goal was not to generate a nationally representative sample, but to ensure that we had selected fairly typical facilities within the states from which we sampled with a range of characteristics. Thus, although our sample had a relatively low prevalence of for-profit facilities, the staffing levels were comparable with the levels in the states from which they were selected (within for-profit and not-for-profit sectors). Although our estimate of the level of QOL in NFs cannot be extrapolated beyond our sample, the estimated associations between resident and facility factors and QOL should not be biased. Indeed, we found that the direction of bivariate associations was consistent when stratifying by state and by ownership type.

Third, although the staffing data in OSCAR are reported by NF staff and may be subject to some bias, in general the system is considered accurate and reliable (Harrington et al. 2000) and has been used by researchers as well as the General Accounting Office in numerous studies. The reliability and validity of the MDS are generally considered acceptable (Hawes et al. 1995; Casten et al. 1998; Lawton et al. 1998) and these data are increasingly being used for research and policy. However, evidence of interfacility variation in the reliability of MDS data suggests that care be taken when drawing conclusions (Mor et al. 2003b). Although the QOL data were collected by trained research interviewers monitored for quality, reliability, and "drift," the strength of the associations between QOL and measures derived from the MDS may be attenuated by measurement error in the MDS and by the fact that the data from the MDS may have been captured up to 45 days from the date of the QOL interview. The proposed inclusion of direct collection of QOL data in MDS 3.0 represents a shift in the way the MDS is conducted, and will require training for staff to achieve a level of proficiency. As with the current MDS program, any modifications to it would benefit from some form of ongoing reliability and validity checking.

Fourth, although our sample of facilities and residents is quite large compared with previous research on NF QOL, we recognize that with 101 facilities, there is a limit to the number of explanatory variables that can be included in a multivariate analysis. Nevertheless, the sample size was more than adequate for bivariate analysis and to establish that the majority of variation in QOL is at the individual rather than the facility level. Our analytic approach was selected to assess the percentage of variance explained by resident and facility factors. For this reason we refrained from exploring "random slopes" models that tested the hypothesis that the impact of individual resident characteristics on QOL (e.g., impaired vision) is attenuated in certain types of facilities. This is a fruitful avenue for future research.

Lastly, we note that our findings are limited to the 60 percent of residents capable of providing responses to a self-report interview. There are several ways to interpret these findings. One approach is to consider these residents as key informants whose experience is indicative of the climate in the entire facility. If those residents who are cognitively intact enough to reflect on their own QOL are unable to achieve a modicum of respect and dignity, then it is unlikely that demented residents will be effective advocates for their own self-interest in such an environment. On the other hand, the interactions and relationships between staff and cognitively intact and cognitively impaired residents are likely qualitatively different because of the different demands that they place on the facility. Thus we do not assume that data on the experience of residents who are capable of self-report is sufficient to characterize the entire resident population. Nevertheless, this study, and the underlying data collection instruments, represents a dramatic step forward in knowledge about the lives of nursing home residents. There is a continued need for additional theoretical and empirical work to define and measure the QOL of the large proportion of the nursing home population who are capable of an in-person interview.

CONCLUSION

Efforts to measure NF resident QOL are of interest to policy makers, residents, and families. Data on the QOL at a particular facility provides information about the lives of the people who live there that is distinct from the clinical aspect of the care they receive. This information may be used by residents and family members to select a facility, or by policy makers and regulators to target enforcement activities and incentives. Additional emphasis on the QOL of NF residents would be a positive augmentation to the current NF quality initiative (Zimmerman et al. 1995; Berg et al. 2001; General Accounting Office 2002). The very act

of measuring QOL and reporting it will make it more salient; ultimately we may want to identify ways to build incentives for achieving better QOL.

ACKNOWLEDGMENTS

The work described in this paper was supported by a 1998 contract from the Centers for Medicare & Medicaid Services (under Master Contract # 500-96-0008). Dr. Degenholtz received supported from a NIH career development award (K01AG20516). The content and opinions are solely those of the authors and should not be construed as representing policy of CMS, the University of Minnesota, or the University of Pittsburgh. The authors thank their CMS project officers, Mary Pratt and Karen Schoeneman. They also thank M. Powell Lawton, an investigator on this project until his death in 2001.

SUPPLEMENT MATERIAL

REFERENCES

- Berg, K., B. E. Fries, R. N. Jones, S. Mattke, T. Moore, V. Mor, J. N. Morris, K. Murphy, and S. Nonemaker. 2001. *Identification and Evaluation of Existing Quality Indicators That Are Appropriate for Use in Long-Term Care Settings*, p. 163. Cambridge, MA: Abt Associates Inc.
- Bryk, A., and S. Raudenbush. 1992. *Hierarchical Linear Models*. Newbury Park: Sage.
- Burrows, A. B., J. N. Morris, S. E. Simon, J. P. Hirdes, and C. Phillips. 2000. "Development of a Minimum Data Set-Based Depression Rating Scale for Use in Nursing Homes." *Age and Ageing* 29 (2): 165-72.
- Casten, R., L. M. Powell, P. A. Parmelee, and M. Kleban. 1998. "Psychometric Characteristics of the Minimum Data Set I: Confirmatory Factor Analysis." *Journal of the American Geriatrics Society* 46 (6): 726-35.
- Castle, N. G. 2002. "Nursing Homes with Persistent Deficiency Citations for Physical Restraint Use." *Medical Care* 40 (10): 868-78.
- Centers for Medicare & Medicaid Services. 2003. "Draft Version of MDS 3.0" [accessed on May 27, 2003]. Available at <http://www.cms.hhs.gov/quality/mds30/DraftMDS30.pdf>
- de Leeuw, J., and I. G. G. Kreft. 1995. "Questioning Multilevel Models." *Journal of Educational and Behavioral Statistics* 20 (2): 171-89.
- Finch, M., R. L. Kane, and I. Philp. 1995. "Developing a New Metric for ADLs." *Journal of the American Geriatrics Society* 43 (8): 877-84.
- General Accounting Office. 2002. *Nursing Homes: Public Reporting of Quality Measures Has Merit but National Implementation Is Premature*, p. 35. Washington, DC: General Accounting Office.
- Goldstein, H. 1995. *Multilevel Statistical Models*. New York: John Wiley & Sons.

- . 1998. *MLwiN* release software. London: Institute for Education.
- Grabowski, D. C., and N. G. Castle. 2004. "Nursing Homes with Persistent High and Low Quality." *Medical Care Research Review* 61 (1): 89–115.
- Harrington, C., D. Zimmerman, S. L. Karon, J. Robinson, and P. Beutel. 2000. "Nursing Home Staffing and Its Relationship to Deficiencies." *Journal of Gerontology B Psychological Science and Social Science* 55 (5): S278–87.
- Hawes, C., J. N. Morris, C. D. Phillips, V. Mor, B. E. Fries, and S. Nonemaker. 1995. "Reliability Estimates for the Minimum Data Set for Nursing Home Resident Assessment and Care Screening (MDS)." *Gerontologist* 35 (2): 172–8.
- Institute of Medicine. 1986. *Improving the Quality of Care in Nursing Homes*. Washington, DC: National Academy Press.
- Kane, R. A., R. L. Kane, B. Bershadsky, H. Degenholtz, and K. C. Kling. 2004a. "Chapter 7: Development of Short Quality of Life Screeners." In *Measures, Indicators and Improvement of Quality of Life in Nursing Homes: Final Report Submitted to the Centers for Medicare and Medicaid Services*. Minneapolis: University of Minnesota (Accessed on May 13, 2005, available at <http://www.cms.hhs.gov/quality/nhqi/QualityOfLife.asp>).
- Kane, R. A., K. C. Kling, B. Bershadsky, R. L. Kane, K. Giles, H. B. Degenholtz, J. Liu, and L. J. Cutler. 2003. "Quality of Life Measures for Nursing Home Residents." *Journals of Gerontology: Biological Science and Medical Science* 58 (3): 240–8.
- Kane, R. L., B. Bershadsky, R. A. Kane, H. Degenholtz, J. Liu, and K. C. Kling. 2004b. "Using Resident Reports of Quality of Life to Distinguish among Nursing Homes." *Gerontologist* 44 (5): 624–32.
- Lawton, M. P. 1983. "Environment and Other Determinants of Well-Being in Older People." *Gerontologist* 23 (4): 349–57.
- Lawton, P. M., R. Casten, P. A. Parmalee, K. VanHaitsma, J. Corn, and M. H. Kleban. 1998. "Psychometric Characteristics of the Minimum Data Set II: Validity." *Journal of the American Geriatrics Society* 46 (6): 736–44.
- Mor, V., J. Angelelli, D. Gifford, J. Morris, and T. Moore. 2003a. "Benchmarking and Quality in Residential and Nursing Homes: Lessons from the US." *International Journal of Geriatrics Psychiatry* 18 (3): 258–66.
- Mor, V., J. Angelelli, R. Jones, J. Roy, T. Moore, and J. Morris. 2003b. "Inter-Rater Reliability of Nursing Home Quality Indicators in the U.S." *BMC Health Service Research* 3 (1): 20.
- Mor, V., K. Berg, J. Angelelli, D. Gifford, J. Morris, and T. Moore. 2003c. "The Quality of Quality Measurement in U.S. Nursing Homes." *Gerontologist* 43 (2): 37–46.
- Mor, V., K. Branco, J. Fleishman, C. Hawes, C. Phillips, J. Morris, and B. Fries. 1995. "The Structure of Social Engagement among Nursing Home Residents." *Journals of Gerontology: Psychological Sciences* 50B (1): P1–8.
- Morris, J. N., B. E. Fries, D. R. Mehr, C. Hawes, C. Phillips, V. Mor, and L. A. Lipsitz. 1994. "MDS Cognitive Performance Scale." *Journal of Gerontology: Medical Sciences* 49 (4): M174–82.

- Morris, J. N., K. Murphy, and S. Nonemaker. 1995. *Long Term Care Facility Resident Assessment Instrument (RAI) User's Manual for Version 2.0*. Baltimore: Health Care Financing Administration.
- Mukamel, D. B., and C. A. Brower. 1998. "The Influence of Risk Adjustment Methods on Conclusions about Quality of Care in Nursing Homes Based on Outcome Measures." *Gerontologist* 38 (6): 695–703.
- Mukamel, D. B., and W. D. Spector. 2000. "Nursing Home Costs and Risk-Adjusted Outcome Measures of Quality." *Medical Care* 38 (1): 78–89.
- . 2003. "Quality Report Cards and Nursing Home Quality." *Gerontologist* 43 (2): 58–66.
- Mukamel, D. B., N. M. Watson, H. Meng, and W. D. Spector. 2003. "Development of a Risk-Adjusted Urinary Incontinence Outcome Measure of Quality for Nursing Homes." *Medical Care* 41 (4): 467–78.
- Murray, D., and P. Hannan. 1990. "Planning for the Appropriate Analysis in School-Based Drug-Use Prevention Studies." *Journal of Consulting and Clinical Psychology* 58 (4): 458–68.
- Neufeld, R. R., L. S. Libow, W. J. Foley, J. M. Dunbar, C. Cohen, and B. Breuer. 1999. "Restraint Reduction Reduces Serious Injuries among Nursing Home Residents." *Journal of the American Geriatrics Society* 47 (10): 1202–7.
- Newhouse, J. P., M. B. Buntin, and J. D. Chapman. 1997. "Risk Adjustment and Medicare: Taking a Closer Look." *Health Affairs* 16 (5): 26–43.
- Porell, F., and F. G. Caro. 1998. "Facility-Level Outcome Performance Measures for Nursing Homes." *Gerontologist* 38 (6): 665–83.
- Schroll, M., P. V. Jonsson, V. Mor, K. Berg, and S. Sherwood. 1997. "An International Study of Social Engagement among Nursing Home Residents." *Age and Ageing* 26 (S2): 55–9.
- Singer, J. D. 1998. "Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models." *Journal of Educational and Behavioral Statistics* 24 (4): 323–55.
- Snowden, M., W. McCormick, J. Russo, D. Srebnik, K. Comtois, J. Bowen, L. Teri, and E. B. Larson. 1999. "Validity and Responsiveness of the Minimum Data Set." *Journal of the American Geriatrics Society* 47 (8): 1000–4.
- Zimmerman, D. R., S. L. Karon, G. Arling, B. R. Clark, T. Collins, R. Ross, and F. Sainfort. 1995. "Development and Testing of Nursing Home Quality Indicators." *HCFA Review* 16 (4): 107–27.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:

APPENDIX S1. QOL-14 Items and Correlation with Original Domain Scales from which Items were Drawn.

APPENDIX S2. Survey Sanction Related to QOL