The Gerontologist
Vol. 47, No. 5, 672–682

Nursing Effort and Quality of Care for Nursing Home Residents

Greg Arling, PhD,[1] Robert L. Kane, MD,[2] Christine Mueller, RN, PhD,[3] Julie Bershadsky,[2] and Howard B. Degenholtz, PhD[4]

Purpose: The purpose of this study was to determine the relationship between nursing home staffing level, care received by individual residents, and resident quality-related care processes and functional outcomes. ***Design and Methods:*** Nurses recorded resident care time for 5,314 residents on 156 units in 105 facilities in four states (Colorado, Indiana, Minnesota, and Mississippi). We linked residents' care times to their measures of health and functioning from Minimum Data Set assessments. Major variables were unit- and resident-specific minutes of care per day, process measures (physical restraints, range of motion, toileting program, and training in activities of daily living [ADLs]), outcome measures (ADL decline, mobility decline, and worsening behavior between the time study and 90-day follow-up), and covariates such as unit type and resident health status. We used multilevel analysis to examine staffing and quality relationships. ***Results:*** Residents with toileting programs, range of motion or ADL training, and restraints received significantly more care from unlicensed but not from licensed staff. However, functional outcomes were not significantly related to care received from licensed or unlicensed staff, except for ADL decline, which was greatest for residents receiving more unlicensed minutes of care. Unit staffing level (licensed and unlicensed) was unrelated to any of the care processes or outcome measures, although higher overall staffing was associated with more time devoted to direct resident care. ***Implications:*** Future research into nursing home quality should focus on organization and delivery rather than simply the amount of care available.

Key Words: Nursing home, Quality, Nurse staffing, Outcomes, Multilevel model

Poor-quality nursing home care is a persistent concern. Efforts to address this concern have included stronger regulation, stricter educational standards, and expanded consumer information. Another proposed strategy is to increase nursing home direct care staffing levels under the assumption that many nursing homes have insufficient staff resources to deliver good-quality care.

The evidence to support the relationship between staffing and quality deserves closer examination. Although the majority of studies (described here) have shown a positive relationship between at least one staffing variable and one quality variable, these same studies have found that there are no significant relationships between other staffing and quality variables.

Most studies examining the relationship between nurse staffing and quality have used facility-level staffing data from the Centers for Medicare & Medicaid Services (CMS) Online Survey, Certification and Reporting (OSCAR) database (Akinci & Krolikowski, 2005; Bostick, 2004; Castle, 2002; Cherry, 1991; Harrington, Zimmerman, Karon, Robinson, & Beutel, 2000; Jette, Warren, & Wirtalla, 2004; Mosely & Jones, 2003; Spector & Takada, 1991; Stevenson, 2005; Weech-Maldanado, Meret-Hanke, Neff, & Mor, 2004; Zhang & Grabowski, 2004; S. Zimmerman, Gruber-Baldini, Hebel, Sloane, & Magaziner, 2002). The inconsistency in facility reporting of OSCAR staffing data has raised questions

This study was supported by Grant 1 R01 AG021985 from the National Institute on Aging and by earlier contracts from the Medicaid programs in Colorado, Indiana, Minnesota, and Mississippi. Myers and Stauffer LC conducted the nursing home time studies, and CareTrack, Inc., prepared the data for analysis. Teresa Lewis assisted in analysis and manuscript preparation. We thank the staff and residents of nursing homes in the study who gave generously of their time. We are solely responsible for any errors or omissions as well as the opinions expressed herein.

Address correspondence to Greg Arling, PhD, Indiana University Center for Aging Research, Regenstrief Institute, HITS Building Room 2000, 410 West 10th Street, Indianapolis, IN 46202-3012. E-mail: GArling@iupui.edu

[1]Indiana University Center for Aging Research and Regenstrief Institute, Indiana University School of Medicine, Indianapolis.

[2]Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis.

[3]School of Nursing, University of Minnesota, Minneapolis.

[4]Department of Health Policy and Management, Graduate School of Public Health, and Center for Bioethics and Health Law, University of Pittsburgh, PA.

about their reliability (CMS, 2001a). Other studies have used administrative data such as Medicaid cost reports (Anderson, Hsieh, & Su, 1998; Bates-Jensen, Schnelle, Alessi, Al-Samarrai, & Levy-Storms, 2004; CMS, 2001b; Munroe, 1990; Nyman, 1988; Rantz et al., 2004; Schnelle et al., 2004), which typically cover a long time period (e.g., 12 months) and, thus, may be insensitive to variation in staffing over time. Also, none of the previous studies have had unit-level measures of staffing.

Quality measures in previous studies have varied greatly and have included such variables as care deficiencies cited by nursing home surveyors (Akinci & Krolikowski, 2005; Castle, 2002; Harrington et al., 2000; Mosely & Jones, 2003; Munroe, 1990), care processes (e.g., use of restraints and catheters; Graber & Sloane, 1995; Phillips et al., 1996; Schnelle et al., 2004; Sullivan-Marx, Strumpf, Evans, Baumgarten, & Maislin, 1999; Svarstad & Mount, 2001; Weech-Maldanado et al., 2004), and health or functional outcomes (e.g., weight loss, pressure ulcers, or functional decline; Anderson et al., 1998; Bliesmer, Smayling, Kane, & Shannon, 1998; Bostick, 2004; CMS, 2001b; Cohen & Spector, 1996; Horn, Buerhaus, Bergstrom, & Smout, 2005; Jette et al., 2004; Weech-Maldanado et al., 2004; S. Zimmerman et al., 2002). The validity of survey data as quality indicators has been challenged because of differences in the way surveyors investigate and cite deficiencies within and between states (CMS, 2001c). Studies using Minimum Data Set (MDS) or other resident-level data have lacked common definitions of processes or outcomes (Bostick, 2004; CMS, 2001b; Horn et al., 2005; Weech-Maldanado et al., 2004; Zhang & Grabowski, 2004). In addition, endogeneity between measures of staffing and quality makes causal inference problematic, particularly for process measures for which staffing level both influences and is influenced by the types of care being provided to residents. Another threat to causal inference is spuriousness, whereby higher staffing and higher quality, although correlated, may result from a third factor. For example, facilities committed to better care may hire more staff and manage them more effectively.

Researchers have found staff skill mix (i.e., proportion registered nurse [RN], licensed practical nurse [LPN], nursing assistant, or other staff types) to be an important factor in quality. The majority of studies have pointed to the importance of licensed nurse time, particularly RN time (Akinci & Krolikowski, 2005; Anderson et al., 1998; Bliesmer et al., 1998; Bostick, 2004; Castle, 2002; Cherry, 1991; Cohen & Spector, 1996; Harrington et al., 2000; Horn et al., 2005; Munroe, 1990; Spector & Takada, 1991; Weech-Maldanado et al., 2004; S. Zimmerman et al., 2002). Fewer studies have found unlicensed staff time to be significant (Akinci & Krolikowski, 2005; Bostick, 2004; Graber & Sloane, 1995; Harrington et al., 2000; Schnelle et al., 2004). A number of studies have found that licensed practical/vocational nurse time either was not significantly related to quality (Castle, 2002; Harrington et al., 2000) or had a negative relationship (Bostick, 2004; Horn et al., 2005; Sullivan-Marx et al., 1999; Zhang & Grabowski, 2004; S. Zimmerman et al., 2002).

All previous studies have examined quality and staffing relationships at the facility or nursing unit level. Yet effects of staffing are relevant at two levels—the overall staffing resources available on the unit (RN, LPN, or aide hours per resident day [HPRD]) and the amount of care provided to each resident. Available staffing constrains the amount of care a resident can receive, but the allocation of care across residents on a unit is also influenced by clinical decisions, management practices, physical environment, and other factors that are difficult to observe. In the end, it is the time spent by staff with an individual resident that should have the greatest effect on care quality. Ours is the first study to measure the time spent by different staff types with individual residents and to employ multilevel modeling to examine simultaneously the effects of unit staffing and care received by individual residents on quality measures.

Multilevel Framework

We approached issues of nursing effort and quality from a multilevel perspective (Raudenbush & Bryk, 2002), which recognizes that residents are nested within nursing units and that relationships between staffing and quality relationships need to be examined at the nursing unit and resident levels simultaneously. We addressed two interrelated questions: Is the amount and mix of direct care on a nursing unit overall related to a resident's care processes and outcomes? And, given a unit's overall care level, how does the amount of care received by a resident relate to that resident's care processes and outcomes? Overall levels of staffing assigned to a unit may reflect its case mix; the effort directed toward an individual resident is limited by unit staffing levels. It is thus appropriate to employ multilevel modeling to account for the proportionate influence of each level of staffing. In the multilevel framework, unit staffing is a contextual variable that influences but does not predetermine the amount of care received by individual residents. We could not directly observe care allocation decisions being made by nursing staff. Care time may be directed to residents with the greatest need (e.g., higher acuity or risk), to those who have the greatest potential for restoration or maintenance of health and functioning, or on some other basis that has little relationship to outcomes (e.g., the most pleasant residents or those with the most insistent families). Nonetheless, by including individual- and facility-level measures of acuity or risk we could assess the role of these potentially confounding or interacting factors. It was important, however, to choose risk factors that were independent of clinical interventions

to avoid circular effects of poor care practices and outcomes (Arling, Karon, Sainfort, Zimmerman, & Ross, 1997; D. R. Zimmerman et al., 1995).

The multilevel framework provides a much stronger test of staffing and care quality relationships than earlier facility-level (single-level) studies that either aggregated data to the facility level and thus lost information about care provided to individual residents, or relied on resident-level models that treated staffing as a characteristic of the individual, thereby running the risk of an ecological fallacy and biased standard errors. We assessed the independent additive effects of overall care available on the unit and amount of care provided to each resident.

Hypotheses

Our study focused on process and outcome quality indicators related to resident functioning: (a) care processes that could inhibit functioning (physical restraints) or have a beneficial impact (range of motion exercises, toileting program, or activities of daily living [ADLs] training); and (b) outcomes consisting of functional decline in ADLs, continence, and behavioral problems. Care time measures were: (a) unit staffing level (or the average number of HPRD spent by nursing and other direct care staff on the unit during the study) and (b) the amount of care received individually by each resident. We hypothesized that residents on units with higher staffing levels would be more likely to have a toileting program, ADL training, range of motion programs, and fewer restraints. In addition, individual residents with ADL training and range of motion programs would likely receive more care than residents without these services. We also hypothesized that residents who were restrained would receive greater care because of the additional time required for restraint monitoring and documentation. With regard to outcome quality indicators, we hypothesized that residents on units with higher staffing levels and those receiving more care at baseline would be less likely to have a decline in ADL, worsening incontinence, and worsening behavioral problems between baseline and follow-up. In general, one might expect that process measures of quality would be more sensitive to staffing than outcomes, because the latter can be influenced by many other factors as well. We patterned our measures after nursing home quality indicators/quality measures. Other quality indicators covering untoward health events, such as pressure sores, weight loss, and so on, did not have a high enough incidence rate (>5%) to be adequately modeled.

Methods

Sample

The combined four-state sample from studies in Colorado (1998), Indiana (1999), Mississippi (2001), and Minnesota (2004) totaled 105 facilities, 156 nursing units, and 5,314 residents. The samples were drawn independently in each state with similar procedures. Alzheimer's special care units in Indiana, Minnesota, and Mississippi were oversampled for separate analysis. About 10% of facilities in each state were excluded from the sample frame due to serious nursing home survey deficiencies, low staffing, or recent ownership change. A randomly chosen sample of facilities consisting of twice as many facilities as was estimated to be needed for the analysis was invited to participate. Approximately one half of the invited facilities in each state chose to participate. Study facilities were similar to nonstudy facilities in the four states in percentage hospital based (10%); percentage Medicaid (62%) and Medicare (9%); and licensed, unlicensed, and total nursing HPRD. However, study facilities were more likely to be nonprofit (50%), nonchain affiliated (44%), and larger (97 residents) and to have Alzheimer's special care units (40%). According to OSCAR staffing data for the approximate study period, study facilities were close to the national average of 3.6 total nursing HPRD; they averaged 3.40 HPRD, with a high of 3.68 HPRD (Colorado) and a low of 3.39 HPRD (Indiana). Also, study facilities did not differ significantly in average HPRD compared to other facilities in their states. See Arling, Kane, Mueller, and Lewis (2007) for details about sampling and data collection.

Data Collection

Nurse staff time and resident data were collected at the unit level. One to three units were selected randomly from each participating facility. A project team spent a week on site in each nursing home training nursing home staff, monitoring data collection, and ensuring data accuracy. All direct care staff on sampled units participated in the study: RNs, LPNs, and nursing assistants; restorative aides; activity directors and aides; social workers and social services aides; and physical, occupational, and speech therapists. Staff members entered time into hand-held computers while delivering care to each resident. Nursing staff recorded time over a 48-hr period and ancillary staff over 7 days. All residents on sampled units were included in the study, and all staff time (including breaks and meals) was accounted for. As staff members went about their tasks, they recorded their time for specific residents if possible. All hands-on care and some administrative tasks such as charting could be associated with specific residents and was termed *resident-specific time* (RST), whereas time spent with other tasks involving general administration or unit maintenance could not be associated with specific residents and was termed *non-RST* (NRST). Each resident had a unique number of RST minutes; the NRST minutes were

Table 1. Care Process and Outcomes Quality Measures and Covariates

Quality Measure	Definition	Covariates
Physical restraints	Daily use of trunk, limb, or chair restraint in past week. [Exclusions: None]	Dementia, resident age, and CPS score
Toileting program	Any scheduled toileting plan or bladder retraining program used in past 2 weeks. [Exclusions: None]	Resident gender, age, CPS score, RUG Special Care/Extensive category, LOS <45 days
Active or passive range of motion training	Active or passive range of motion restorative care provided at least 15 min/day for at least 5 days in past week. [Exclusions: None]	Resident age, CPS score, sum of range of motion limitations, and LOS <45 days
ADL training	Any rehabilitative ADL training (bed mobility, transfer, walking, or dressing/grooming) provided at least 15 min/day for at least 5 days in past week. [Exclusions: None]	Resident age, CPS score, RUG Special Care/Extensive category, LOS <45 days
ADL decline	Decline of in ADL items (bed mobility, transfer, eating, toileting) between STM and follow-up assessments. [Exclusions: Totally ADL dependent at STM assessment, comatose, end-state disease, or hospice]	Resident hemiplegia/hemiparesis diagnosis, CPS score, ADL Long score, extensive/total eating assistance, Personal Severity Index score, age, gap between time study and assessment.
Worsening bowel or bladder incontinence	Increase in bowel incontinence or bladder incontinence between STM and follow-up assessments. [Exclusions: Totally bowel or bladder incontinent at STM assessment, ostomy use in past 2 weeks, comatose, end-state disease, or hospice]	Resident CPS score, ADL Long score, RUG Special Care/Extensive Category, gap between time study and assessment, self-locomotion on unit score
Worsening behavior problems	Greater count of any occurrence of behavior problems (wandering, verbally abusive, physically abusive, or socially inappropriate/disruptive) between STM and follow-up assessments. [Exclusions: Any occurrence of all four behavior problems at STM or comatose]	Resident dementia diagnosis, anxiety disorder, gender, ability to make self understood, moderately/severely impaired decision making, motor agitation, RUG Rehabilitation category, behavior
All quality indicators		Unit level: State: Colorado, Indiana, Minnesota, or Mississippi; Alzheimer's special care unit; unit percentage RUG-III Rehabilitation, Special Care/Extensive, or Clinically Complex categories; unit mean ADL Long and CPS scores

Note: ADL = activity of daily living; STM = Staff Time Measurement Study; CPS = Cognitive Performance Scale; RUG = Resource Utilization Group III; LOS = length of stay.

averaged at the unit level because there was no basis for assigning them to individual residents. The staff time data were subjected to extensive error testing after being entered into the study database. Residents' health and functional status information came from MDS assessments that were successfully matched to the time study data for 98% of the residents.

Variables

Both resident- and unit-level resource use measures were based on care times recorded by staff during the study. Resident-level variables were operationalized as RST per resident day received by each resident from RN, LPN, aide, or activity/social work staff. Unit staffing variables were operationalized as the unit average RST per resident day from RN, LPN, and a combined unlicensed staff category (aide, activity, and social work staff); and unit average NRST per resident day from licensed (RN and LPN) and unlicensed (aide, activity, and social work) staff. Nonproductive time (e.g., breaks and meals, and time spent by staff off unit) were excluded from the analysis. Licensed therapy time (physical, occupational, and speech) was also excluded from the resource use measures because only 6% of residents received therapies, most therapies were provided off unit, and addition of therapy time to other direct care times would have resulted in highly skewed resource use variables. The unit staffing variables (i.e., number of licensed and unlicensed HPRD) were based on recorded time during the time study; total HPRD was equal to RST HPRD + NRST HPRD.

Measures of care quality, defined in Table 1, were based primarily on quality indicators or measures from national sources (Morris et al., 2003; D. R. Zimmerman et al., 1995). Process quality measures

and covariates were derived from the MDS assessment closest to the time study (STM assessment), whereas the outcome quality measures were measured as changes in status between the STM assessment and the resident's next MDS assessment (follow-up assessment) approximately 90 days later. The mean gap between STM assessment and study date was 0.2 days (SD = 24.2). The mean gap between STM and follow-up assessment dates was 90.4 days (SD = 19.3). Some residents were excluded from quality indicator calculations if they were inappropriate or did not meet risk criteria (Morris et al., 2003; D. R. Zimmerman et al., 1995).

All quality indicators were adjusted for resident characteristics proposed by developers of the quality measures (Morris et al., 2003; D. R. Zimmerman et al., 1995) or by the project research team (Kane, Flood, Bershadsky, & Keckhafer, 2004; Kane et al., 2005) and by unit-level measures of many of the same variables. To the extent possible, risk factors should be predictive of a quality indicator yet outside provider control and not themselves a result of poor quality. A project advisory panel recommended the final adjuster covariates for each quality measure. These covariates proved statistically significant and were judged to be conceptually independent of facility actions. Major categories from the Resource Utilization Group III (RUG-III) served as indicators of resident and unit acuity (Fries et al., 1994). Covariates scored as indices or scales were the long form ADL dependency index (Morris, Fries, & Morris, 1999), with a range of 4 (independent) to 18 (totally dependent); Cognitive Performance Scale, with a range of 0 (intact) to 6 (very severely impaired; Morris et al., 1994); and the Personal Severity Index, with scores ranging from 0 to 18 on a series of diagnoses, functional or cognitive deficits, or other conditions (Morris et al., 2003).

Table 2. Resident- and Unit-Level Care Times per Resident Day

	Minutes		Hours	
Unit Care Time (N = 157)	M	SD	M	SD
RST				
Unit avg. RN	12.6	16.7	0.21	0.28
Unit avg. LPN	20.2	11.6	0.34	0.19
Unit avg. unlicensed	77.5	28.6	1.29	0.48
Unit avg. total RST (RN, LPN, and unlicensed)	110.3	33.2	1.84	0.55
NRST				
Unit avg. licensed staff (RN and LPN)	27.9	13.9	0.47	0.23
Unit avg. unlicensed	48.7	22.2	0.81	0.37
Unit avg. total (licensed and unlicensed)	76.6	28.7	1.28	0.48
Total time (RST + NRST)				
Unit avg. licensed staff (RN and LPN)	60.8	28.4	1.01	0.47
Unit avg. unlicensed	126.2	45.6	2.10	0.76
Unit avg. total (licensed and unlicensed)	186.9	52.2	3.12	0.87
Lowest quartile total time (n = 39)	133.0	16.6	2.22	0.28
2nd quartile total time (n = 39)	167.6	5.9	2.79	0.10
3rd quartile total time (n = 40)	189.7	9.2	3.16	0.15
Highest quartile total time (n = 39)	257.7	47.0	4.29	0.78
Resident RST care times (N = 5,242)				
RN RST	11.5	19.6	0.19	0.33
LPN RST	20.1	19.3	0.34	0.32
Aide RST	77.4	49.1	1.29	0.82
Activity staff RST	6.0	5.1	0.10	0.09
Total RST (all staff)	115.0	59.3	1.92	0.99

Notes: RST was for hands-on care and other tasks that could be associated with specific residents; NRST represents the remainder of staff time for tasks such administration or unit maintenance that could not be associated with specific residents. RST = resident-specific time; Avg. = average; RN = registered nurse; LPN = licensed practical nurse; NRST = non-RST; SD = standard deviation.

Descriptive Statistics

Table 2 presents unit- and resident-level care times for the sample. The average unit staffing level consisted of 1.84 HPRD of RST, 1.28 HPRD of NRST, and 3.12 HPRD of combined RST and NRST. This was divided into 12.6 RN min, 20.2 LPN min, and 77.5 unlicensed staff (nursing assistant and activity) min of RST per resident day; 27.9 licensed staff min (RN and LPN) and 48.7 unlicensed staff min of NRST per resident day; and a total of 60.8 min of licensed and 126.2 min of unlicensed minutes per resident day. Units ranged from 1.53 to 6.18 total HPRD. The staffing levels for the top quartile of nursing units averaged a total of 4.29 HPRD, whereas the bottom quartile of units averaged a total of 2.22 HPRD. The mean RST minutes across all residents were close to the unit averages: 11.5 RN min, 20.1 LPN min, 77.4 aide min, and 6.0 activity staff min. In a separate analysis, not reported in the table, we found modest correlations between unit RST and NRST for licensed (r = .60) and unlicensed (r = .40) staff. Also, about 70% of units falling into the top quartile in total HPRD also were in the top quartile in RST HPRD, whereas the same percentage in the bottom total HPRD quartile was also in the bottom RST HPRD quartile. Although RST HPRD and total HPRD were correlated, units varied in the proportion of total staffing devoted to direct resident care.

Table 3 shows resident quality measures. Among the process quality measures, 10% of residents were physically restrained, 39% had a toileting program, 12% received active or passive range of motion training, and 9% had ADL training. For outcome quality measures, 20% of residents had a decline in

Table 3. Quality Measures and Covariates

Measure	%	*M* (*SD*)
Resident-level quality measures		
Physical restraints ($N = 5{,}209$)	10	
Toileting program ($N = 5{,}242$)	39	
Active or passive range of motion training ($N = 5{,}205$)	12	
ADL training ($N = 5{,}205$)	9	
ADL decline ($N = 4{,}076$)	20	
Worsening bowel or bladder incontinence ($N = 2{,}929$)	31	
Worsening behavior problems ($N = 4{,}533$)	12	
Resident-level covariates ($N = 5{,}242$)		
Ability to make self understood		0.98 (1.07)
ADL score at time of study		15.62 (8.80)
Age		83.05 (11.79)
Anxiety disorder	15	
Behavior score at time of study		1.21 (2.15)
Cognitive Performance Scale score		3.07 (1.82)
Dementia	56	
Extensive/total eating assistance	28	
Female	76	
Day gap between STM study and STM assessment		−0.2 (24.2)
Day gap between STM assessment and follow-up assessment		90.4 (19.3)
Hemiplegia/hemiparesis	8	
Length of stay <45 days	8	
Moderately/severely impaired decision making	81	
Motor agitation	22	
Personal Severity Index score		4.94 (1.77)
RUG Clinically Complex category	16	
RUG Rehabilitation category	6	
RUG Special Care/Extensive category	12	
Self-locomotion on unit score		1.98 (1.66)
Sum of range of motion limitations		1.98 (2.87)
Unit-level covariates ($N = 157$)		
State – Colorado	11	
State – Indiana	32	
State – Mississippi	17	
State – Minnesota	40	
Alzheimer's special care unit	29	
Mean Cognitive Performance Scale score		3.08 (0.93)
Mean ADL score		15.26 (3.74)
% RUG Clinically Complex	16	(10.0)
% RUG Special Care/Extensive	11	(12.1)
% RUG Rehabilitation	9	(15.3)

Note: ADL = activity of daily living; STM = Staff Time Measurement Study; RUG = Resource Utilization Group III.

ADL, 31% had worsening bowel or bladder incontinence, and 12% experienced worsening behavioral problems. The residents' average age was 83 years, and 76% were female. Most were longer stay residents; only 8% had been in the nursing home for less than 45 days at the date of the staff time measurement study. About one third of residents were in higher intensity RUG-III categories of Extensive, Rehabilitation, Special Care, or Clinically Complex. The mean Cognitive Performance Scale score was 3.07 (moderately cognitively impaired range), and the mean ADL score was 15.6. In all, 29% of the study units were Alzheimer's special care units. The unit-level average Cognitive Performance Scale score (3.08) and ADL score (15.26) and the percentage of residents in high acuity RUG-III categories were very close to the resident-level averages.

Analysis

We fitted a separate multilevel model for each quality measure. Each model had the resident-level quality measure as the dependent variable and independent variables consisting of resident-level care time (minutes of care per day from RN, LPN, nursing assistant, or activity staff), resident-level covariates, unit-level care time (average RN, LPN, nursing assistant, or activity staff hours per day on the unit), and unit-level covariates. We initially entered all covariates into each model; for ease of interpretation and display, we reduced the final models to contain only those covariates with a p value less than .10. We performed multilevel modeling with HLM 6.02 statistical software (Raudenbush, Bryk, & Congdon, 2002). Because the quality measures were binary outcomes, we used a hierarchical general linear model framework, assuming a Bernoulli sampling distribution with a logit link function and a linear structural model. We fitted intercepts-as-outcomes models. We tested for variation in slopes of the staff time variables and covariates but found insufficient evidence for slopes-as-outcomes models. We performed estimation with the expectation–maximization LaPlace method in HLM 6.02, which produces accurate approximation to maximum likelihood estimates, particularly in correcting bias that arises when variance is large and probability of outcomes is small.

Results

Table 4 presents the relationships between unit staffing, care received by individual residents, and quality-related care processes. Unit- and resident-level care times were reported in hours per resident to facilitate interpretation of the regression coefficients. We regressed each process—physical restraints, toileting program, active or passive range of motion training, and ADL training—on the staffing and resident care time variables and covariates. Therefore, results in the table show the independent effects of staffing or care time variables after controlling for covariates (actual hierarchical general linear model coefficients are shown). We entered the RST and NRST HPRD staffing variables simultaneously in order to ascertain the independent effects of unit RST when controlling for NRST, and vice versa. None of the unit staffing variables (average RN, LPN, or

Table 4. Results From Multilevel Models for Process Measures of Quality (Laplace Estimated Coefficients and *p* Values)

Measure	Physical Restraints ($N = 5{,}208$)	Toileting Program ($N = 5{,}241$)	Range of Motion Training ($N = 5{,}204$)	ADL Training ($N = 5{,}204$)
Unit-level direct care hours/day				
Avg. RN	0.54	−0.58	−1.91	−0.88
Avg. LPN	−0.78	−1.10	0.32	0.24
Avg. unlicensed	−0.73	0.21	0.12	−0.58
Resident-level direct care hours/day				
RN	−0.06	−0.13	−0.05	−0.11
LPN	0.11	−0.11	−0.06	−0.05
Aide	0.78**	0.60***	0.58***	0.42***
Activity staff	−0.28	2.64***	−0.43	1.11
Unit-level risk adjustors				
Alzheimer special care unit		1.28**		
State – Colorado		—	3.27**	1.61**
State – Indiana		—	2.70**	0.74
State – Minnesota	1.27**	−2.52**	—	−1.34*
Unit % RUG Clinically Complex		2.55		
Avg. ADL	0.09**		0.16**	
Avg. CPS score			−0.83**	
Resident-level risk adjustors				
Age	−0.10 *	0.02**	−0.01*	0.01
CPS score	0.48**	0.08**	0.10**	−0.16**
Dementia	−0.37**			
Female		0.15		
Length of stay <45 days		−0.55**	−1.71**	−0.98**
RUG Special Care/ Extensive		−0.37**		−0.75**
Sum of range of motion limitations			0.16**	

Notes: Covariates with $p \geq .10$ were excluded from the final models; cells in the table corresponding to these covariates have been left blank. ADL = activity of daily living; Avg. = average; RN = registered nurse; LPN = licensed practical nurse; RUG = Resource Utilization Group III; CPS = Cognitive Performance Scale.

*$p < .05$; **$p < .01$; ***$p < .001$.

unlicensed staff RST, or licensed or unlicensed NRST) were significantly related to either the process or outcome quality indicators. Minutes of care received by individual residents was significantly related to process quality indicators. Nurse aide time was significantly higher for residents who were physically restrained, participated in a toileting program, or received range of motion or ADL training. Greater activity staff time was also associated with a toileting program. However, neither RN nor LPN minutes received by a resident was significantly related to the process quality indicators.

The process quality indicators also were related to unit- and resident-level covariates. Physical restraint use was most likely (a) on units with greater average ADL dependency scores and (b) for residents who were younger and had greater cognitive impairment (as measured by the Cognitive Performance Scale). A toileting program was most likely (a) on Alzheimer's special care units and (b) for residents who were older, were more cognitively impaired, had a longer stay, and were not in the RUG-III Extensive or Special Care categories. Receipt of range of motion training was most likely (a) on units with more ADL dependent, but less cognitively impaired, residents; and (b) for residents who were younger, were more cognitively impaired, had a longer stay, and had more range of motion limitations. Training in ADLs was most likely for residents who were older, were less cognitively impaired, had a longer stay, and were not in the RUG-III Extensive or Special Care categories. Finally, the quality indicators varied by state, but no state had consistently higher quality indicator scores.

Table 5 summarizes the results from multilevel models for functional outcome quality measures—decline in ADL, worsening bowel or bladder continence, and worsening behavior. We based the outcome variables on change in status between the MDS assessment closest to the STM study date (STM assessment) and a follow-up assessment approximately 90 days later. We also entered into these models covariates measured at the first assessment. Of the 21 effects tested (3 outcomes × 7 care time variables), only 1 was significant. The more aide time received by a resident, the more likely the resident's decline in ADLs. None of the unit staffing variables (average RST or NRST HPRD) were significantly related to functional outcomes, nor were any of the other resident-level care time variables.

Among the covariates, ADL decline was most likely (a) on units having a higher percentage of rehabilitation residents and (b) for residents who were less ADL dependent, were older, were more

Table 5. Results From Multilevel Models for Outcome Measures of Quality (Laplace Estimated Coefficients and *p* Values)

Measure	ADL Decline ($N = 4,049$)	Worsening Bowel or Bladder Incontinence ($N = 2,633$)	Worsening Behavior Problems ($N = 3,509$)
Unit-level direct care hours/day			
Avg. RN	−0.27	−0.27	−0.79
Avg. LPN	0.25	0.24	0.45
Avg. unlicensed	−0.27	−0.37	0.17
Resident-level direct care hours/day			
RN	0.09	0.31	0.23
LPN	0.13	0.08	−0.20
Aide	0.42***	0.09	0.02
Activity staff	0.39	−0.51	0.92
Unit-level risk adjustors			
State – Indiana	0.53**		
State – Minnesota			−0.52**
Unit % RUG Clinically Complex		−2.21**	−1.80**
Unit % RUG Special Care/Extensive	1.14		
Unit % RUG Rehabilitation	1.20*		
Resident-level risk adjustors			
Ability to make self understood			−0.26**
ADL at time of study	−0.05**	0.07**	
Age	0.01*		
Anxiety disorder			0.41**
Behavior score at time of study			−0.07*
Cognitive Performance Scale score	0.16**	0.20**	
Dementia			0.35**
Extensive/total eating assistance	−0.66**		
Female			−0.39**
Hemiplegia/hemiparesis	−0.50**		
Moderately/severely impaired decision making			1.12**
Motor agitation			0.36**
Personal Severity Index	−0.07*		
RUG Rehabilitation			0.50*
RUG Special Care/Extensive		0.44**	
Self-locomotion on unit score		−0.11*	
Study–assessment gap in days	−0.004*	−0.003*	

Notes: Covariates with $p \geq .10$ were excluded from the final models; cells in the table corresponding to these covariates have been left blank. ADL = activity of daily living; Avg. = average; RN = registered nurse; LPN = licensed practical nurse; RUG = Resource Utilization Group III.
*$p < .05$; **$p < .01$; ***$p < .001$.

cognitively impaired, did not require extensive eating assistance, did not have hemiplegia, and had lower Personal Severity Index scores. Worsening incontinence was more likely (a) on units with a lower percentage of clinically complex residents and (b) for residents who had greater ADL dependency or cognitive impairment and who were not in the RUG-III Extensive or Special Care categories. Worsening behavior problems were most likely for residents who had fewer behavior problems on their STM assessment; were male; could not make themselves understood; had an anxiety disorder, dementia diagnosis, impaired decision making, or motor agitation; or were in the RUG-III Rehabilitation category. Finally, there were some differences in outcome quality indicators across the states and by day gap between STM and follow-up assessments.

We subjected the multilevel models to considerable sensitivity analysis. For example, we tested resident-level models without unit-level variables (including staffing) and multilevel models with only the staffing and care time variables (excluding covariates). We ran models for long-stay residents (>45 days) under the assumption that they might be most influenced by staff interventions. We also ran separate models for residents having MDS assessments before the STM study and after the study to see if the timing of the assessment might make a difference. In no case did additional statistical relationships emerge between staffing or care time variables and either process or outcome quality indicators.

Finally, we tested for possible threshold effects and nonlinear relationships by assigning nursing units to quartiles according to total staffing HPRD. Units in the top quartile in overall HPRD tended to be in direct resident care (RST), and bottom quartile units in overall HPRD tended to be in the bottom quartile in RST. However, when we entered staffing

quartiles as dummy variables in the multilevel models, being in a very high- (top 25% HPRD) or low- (bottom 25% HPRD) staffed unit had no significant relationship to the process or outcome measures. We should qualify these findings because the number of units in each quartile (39–40) was relatively small, and we may not have had the necessary statistical power to detect an effect.

Discussion

We found only limited support for relationships between the amount of care provided by different types of staff (RN, LPN, nursing assistants, and other) and quality-related care processes and outcomes. Unlicensed staff time was significantly related to receipt of restorative care processes such as toileting, ADL and range of motion training, and use of physical restraints. However, neither licensed nor unlicensed time was related to outcomes. Although more highly staffed units tended to devote more time to direct resident care, unit staffing variables (resident-specific and non-resident-specific HPRD on the unit) had no significant relationship to either the care processes or outcomes quality indicators. Taken together, these findings offer little evidence that more staff time, per se, is associated with better process of care or outcomes in the quality indicators studied. The absence of consistent relationships, especially in the expected directions, suggests that the role of staffing on quality is more complex than previously thought. How staff members are used may be as important as how many are employed. Staffing decisions are administratively made at the unit level. The total staffing on a unit and the amount of time devoted to direct resident care can be a function of resident acuity, management practices, skill mix, and technology. Available staffing is only one factor in determining how staff members assist individual residents. How much care is proactive and how much reactive? What are the skill levels, experience, and dedication of nursing staff? Is care being allocated effectively across residents according to their level of need?

Limitations

Some elements of the study design may have contributed to the lack of significant findings. The first possibility is error in key measures—amount of care received by residents or the indicators of care quality. Any under- or overreporting of care time by nursing staff, or errors in recording MDS assessment data, would have introduced error into our statistical models and made it more difficult to detect relationships between staffing and care quality if they in fact existed. Although the measures of time were proactively collected and specific to residents, we were not able to record at a more detailed level the amount of time spent with specific tasks, such as toileting or ADL training, or to verify that staff were performing these tasks effectively. The higher care times recorded for residents on toileting or range of motion programs may not have been related to provision of these services. The quality indicators relied on the MDS-based quality indicators, the accuracy of which has been challenged (Arling, Kane, Lewis, & Mueller, 2005; Mor et al., 2003). In addition, the timing of our measurement may have affected the results. The STM study took place over a 48-hr period, whereas the quality indicators were measured through MDS assessments spaced over a 45-day period before or after the STM study for process indicators and a 45- to 135-day period for outcome indicators. The care times recorded during the 2-day staff time measurement may not have been representative of the overall amount of care received by residents during the quality indicator (MDS) measurement periods. We took steps to record data as accurately as possible and to select MDS assessments that were as close as possible to the STM study. Nonetheless, measurement error may have played some role.

A second study limitation was the representativeness of the sample. The variation in staffing and care quality among facilities participating in the STM study was likely more restricted than would have been the case in a true random sample. Although nursing homes invited to participate in the study varied widely in staffing levels and general quality indicators, the facilities agreeing to participate may have represented a narrower range of quality and staffing. They might also have had organizational characteristics that could have contributed to care quality. In the Minnesota study, for example, staff attitudes toward work were quite positive in nearly all study facilities, and most facilities had experienced staff with high rates of retention. The majority of residents in our sample did not experience a decline in health status—most remained stable, with some having improved health or functioning. Nor was our study designed to test threshold effects. Some studies have suggested that staffing must be at levels considerably higher than most of the facilities in our study (>4 total HPRD) before having an impact on quality (CMS, 2001c).

Issues for Future Research

Our study raises methodological and theoretical issues that future research needs to address. The first issue is uncertainty about cause and effect in relationships between nursing effort and outcomes. For example, we discovered a positive relationship between amount of nurse aide time and ADL decline. More nursing effort is unlikely to cause residents to decline. More likely, unmeasured resident risk factors accounted for more staff time and declining status. Even with a longitudinal outcome measure, it was

impossible for us to rule out endogeneity between staffing and outcomes. Participants were observed in natural settings where conditions could not be controlled or manipulated. In a true experiment, participants might be randomly assigned to nursing units of various staffing levels, or the researcher might manipulate staffing levels up or down to see what impact this has on care processes and outcomes. In an observational study in which people are observed over time, causality is much more difficult to establish. Does the amount of care provided to a resident determine his or her health outcome, or does the risk of a health outcome determine the amount of care he or she will receive? If a greater amount of care yields better outcomes, then these items should be positively correlated. However, if people at risk for poor outcomes receive more care, then the amount of care would be negatively correlated with good outcomes. Both causal processes might operate at once, allowing the positive effects to cancel out the negative effects, resulting in a nonsignificant relationship. Even with extensive statistical controls for acuity and risk factors, it is difficult to disentangle the complex interactions between resident conditions and care delivery.

A second issue is the intractable nature of chronic health conditions and disease processes for many residents in the nursing home population. Even with the very best of care, residents may fail to improve or may decline in health status over time. Outcomes such as ADL decline, decline in range of motion, worsening incontinence, weight loss, or increasing pain may be difficult to prevent, particularly in the latter stages of Alzheimer's disease, cancer, or other progressively debilitating conditions. Future research should examine effects of staffing on a range of outcomes for resident populations that vary in their responsiveness to care-related interventions.

Third, our findings point to the importance of studying the organizational context of care delivery and not just the resources available. A certain minimum level of staffing (which most if not all facilities in our study may have met) is a necessary condition for good quality. However, after that the most important determinants may be the expertise of direct care staff, staff morale and teamwork, facility or unit management practices, care-related technologies, and so on. These factors likely intervene between the level of staffing on a unit and its impact on health outcomes.

References

Akinci, F., & Krolikowski, D. (2005). Nurse staffing levels and quality of care in northeastern Pennsylvania nursing homes. *Applied Nursing Research, 18*(3), 130–137.

Anderson, R., Hsieh, P., & Su, H. (1998). Resource allocation and resident outcomes in nursing homes: Comparison between the best and worst. *Research in Nursing & Health, 21,* 297–313.

Arling, G., Kane, R. L., Lewis, T., & Mueller, C. (2005). Future development of nursing home quality indicators. *The Gerontologist, 45,* 147–156.

Arling, G., Kane, R. L., Mueller, C., & Lewis, T. (2007). Explaining direct care resource use of nursing home residents: Findings from time studies in four states. *Health Services Research, 42,* 827–846.

Arling, G., Karon, S., Sainfort, F., Zimmerman, D., & Ross, R. (1997). Risk adjustments of nursing home quality indicators. *The Gerontologist, 37,* 757–766.

Bates-Jensen, B., Schnelle, J., Alessi, C., Al-Samarrai, N., & Levy-Storms, L. (2004). The effects of staffing on in-bed times of nursing home residents. *Journal of the American Geriatrics Society, 52,* 931–938.

Bliesmer, M. M., Smayling, M., Kane, R. L., & Shannon, I. (1998). The relationship between nursing staffing levels and nursing home outcomes. *Journal of Aging & Health, 10,* 351–371.

Bostick, J. E. (2004). Relationship of nursing personnel and nursing home care quality. *Journal of Nursing Care Quality, 19*(2), 130–136.

Castle, N. (2002). Nursing homes with persistent deficiency citations for physical restraint use. *Medical Care, 40,* 868–878.

Cherry, R. (1991). Agents of nursing home quality of care: Ombudsmen and staff ratios revisited. *The Gerontologist, 31,* 302–308.

Centers for Medicare & Medicaid Services. (2001a). Data sources of nursing home staffing analysis: Assessment of OSCAR compared to Medicaid cost reports. In *Appropriateness of minimum nurse staffing ratios in nursing homes: Report to Congress: Phase I* (pp. 8.1–8.23). Baltimore: Author.

Centers for Medicare & Medicaid Services. (2001b). The relationship between nurse staffing levels and quality of nursing home care. In *Appropriateness of minimum staffing ratios in nursing homes. Report to Congress: Phase II final* (pp. 2.1–2.26). Baltimore: Author.

Centers for Medicare & Medicaid Services. (2001c). Review of selected research on nursing home staffing and resident outcomes. In *Appropriateness of minimum nurse staffing ratios in nursing homes: Report to Congress: Phase I* (pp. 6.1–6.31). Baltimore: Author.

Cohen, J. W., & Spector, W. D. (1996). The effect of Medicaid reimbursement on quality of care in nursing homes. *Journal of Health Economics, 15,* 23–48.

Fries, B., Schneider, D., Foley, W., Gavazzi, M., Burke, R., & Cornelius, E. (1994). Refining a case-mix measure for nursing homes: Resource Utilization Groups (RUG-III). *Medical Care, 32,* 668–685.

Graber, D. R., & Sloane, P. D. (1995). Nursing home survey deficiencies for physical restraint use. *Medical Care, 33,* 1051–1063.

Harrington, C., Zimmerman, D., Karon, S. L., Robinson, J., & Beutel, P. (2000). Nursing home staffing and its relationship to deficiencies. *Journal of Gerontology: Social Sciences, 55B,* S278–S287.

Horn, S. D., Buerhaus, P., Bergstrom, N., & Smout, R. J. (2005). RN staffing time and outcomes of long-stay nursing home residents: Pressure ulcers and other adverse outcomes are less likely as RNs spend more time on direct patient care. *American Journal of Nursing, 105*(11), 58–70; quiz 71.

Jette, D. U., Warren, R. L., & Wirtalla, C. (2004). Rehabilitation in skilled nursing facilities: Effect of nursing staff level and therapy intensity on outcomes. *American Journal of Physical Medicine & Rehabilitation, 83,* 704–712.

Kane, R. L., Flood, S., Bershadsky, B., & Keckhafer, G. (2004). Effect of an innovative Medicare managed care program on the quality of care for nursing home residents. *The Gerontologist, 44,* 95–103.

Kane, R. L., Homyak, P., Bershadsky, B., Lum, T., Flood, S., & Zhang, H. (2005). The quality of care under a managed-care program for dual eligibles. *The Gerontologist, 45,* 496–504.

Mor, V., Berg, K., Angelelli, J., Gifford, D., Morris, J., & Moore, T. (2003). The quality of quality measurement in U.S. nursing homes. *The Gerontologist, 43*(Special Issue II), 37–46.

Morris, J. N., Fries, B. E., Mehr, D. R., Hawes, C., Phillips, C., Mor, V., et al. (1994). MDS Cognitive Performance Scale. *Journal of Gerontology: Medical Sciences, 49A,* M174–M182.

Morris, J. N., Fries, B. E., & Morris, S. A. (1999). Scaling ADLs within the MDS. *Journal of Gerontology: Medical Sciences, 54A,* M546–M553.

Morris, J. N., Moore, T., Jones, R., Mor, V., Angelelli, J., Berg, K., et al. (2003). *Validation of long-term and post-acute care quality indicators. CMS final report.* Retrieved November 30, 2006, from http://www.cms.hhs.gov/NursingHomeQualityInits/35_NHQIArchives.asp

Mosely, C., & Jones, L. (2003). Registered nurse staffing and OBRA deficiencies in Nevada nursing facilities. *Journal of Gerontological Nursing, 29*(3), 44–50.

Munroe, D. J. (1990). The influence of registered nurse staffing on the quality of nursing home care. *Research in Nursing & Health, 13,* 263–270.

Nyman, J. A. (1988). Improving the quality of nursing home outcomes. Are adequacy- or incentive-oriented policies more effective? *Medical Care, 26,* 1158–1171.

Phillips, C. D., Hawes, C., Mor, V., Fries, B. E., Morris, J. N., & Nennstiel, M. E. (1996). Facility and area variation affecting the use of physical restraints in nursing homes. *Medical Care, 34,* 1149–1162.

Rantz, M., Hicks, L., Grando, V., Petroski, G., Madsen, R., Mehr, D., et al. (2004). Nursing home quality, cost, staffing and staff-mix. *The Gerontologist, 44,* 24–38.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
Raudenbush, S. W., Bryk, A. S., & Congdon, R. T. (2002). HLM (Version 5) [computer software]. Lincolnwood, IL: Scientific Software International.
Schnelle, J., Simmons, S., Harrington, C., Cadoan, M., Garcia, E., & Bates-Jensen, B. (2004). Relationship of nursing home staffing to quality of care. *Health Services Research, 39,* 225–250.
Spector, W. D., & Takada, H. A. (1991). Characteristics of nursing homes that affect resident outcomes. *Journal of Aging & Health, 3,* 427–454.
Stevenson, D. G. (2005). Nursing home consumer complaints and their potential role in assessing quality of care. *Medical Care, 43,* 102–111.
Sullivan-Marx, E. M., Strumpf, N. E., Evans, L. K., Baumgarten, M., & Maislin, G. (1999). Initiation of physical restraint in nursing home residents following restraint reduction efforts. *Research in Nursing & Health, 22,* 369–379.
Svarstad, B. L., & Mount, J. K. (2001). Chronic benzodiazepine use in nursing homes: Effects of federal guidelines, resident mix, and nurse staffing. *Journal of the American Geriatrics Society, 49,* 1673–1678.
Weech-Maldanado, R., Meret-Hanke, L., Neff, M., & Mor, V. (2004). Nurse staffing patterns and quality of care in nursing homes. *Health Care Management Review, 29,* 107–116.
Zhang, X., & Grabowski, D. (2004). Nursing home staffing and quality under the nursing home reform act. *The Gerontologist, 44,* 13–23.
Zimmerman, D. R., Karon, S. L., Arling, G., Clark, B. R., Collins, T., Ross, R., et al. (1995). Development and testing of nursing home quality indicators. *Health Care Financing Review, 16*(4), 107–127.
Zimmerman, S., Gruber-Baldini, A. L., Hebel, J. R., Sloane, P. D., & Magaziner, J. (2002). Nursing home facility risk factors for infection and hospitalization: Importance of registered nurse turnover, administration, and social factors. *Journal of the American Geriatrics Society, 50,* 1987–1995.

Received December 3, 2006
Accepted March 30, 2007
Decision Editor: Linda S. Noelker, PhD