2.0 The Relationship Between Nurse Staffing Levels and the Quality of Nursing Home Care[1]

2.1 Introduction

The policy issue of whether minimum staffing standards should be required in nursing homes depends in part on whether it can be demonstrated that low staffing levels are associated with quality of care problems. If such associations between staffing and quality exist, the next step is to identify staffing thresholds below which quality is compromised and above which additional staffing increases are not associated with significant quality improvements. Finally, we need to determine whether identified staffing thresholds vary based on facility case mix, requiring different minimum levels for different facilities. All of these questions can be studied empirically.

Although we might reasonably assume that low staffing will compromise quality of care in nursing homes, demonstrating this relationship and identifying specific thresholds below which quality declines is a more complex task. Studies have shown positive relationships between staffing levels for different types of staff (RN, LPN, nursing assistant) and a range of quality measures (e.g. number of deficiencies, functional improvement, pressure ulcer rates, and discharged home), but these findings cannot be used for national policy for several reasons.[1-5] First, these studies were not aimed at identifying specific thresholds, rather they sought a positive relationship between staffing and quality of care. Second, many of these studies were not large enough to draw inferences about national policy. Third, conflicting evidence was available from other studies in several of these areas.

This chapter presents Phase 2 of an analysis conducted for the congressionally-mandated study of nurse staffing in nursing homes. In Phase 1 of this analysis, we used claims data and Minimum Data Set (MDS) data from Ohio, Texas, and New York to investigate the relationship between staffing thresholds and quality of care measured by rehospitalization for potentially avoidable causes (e.g. congestive heart failure, respiratory infections, sepsis) and functional improvement, pressure sores, resisting care, and weight loss. The Phase 1 analysis resulted in the following findings:[6-9]

1. We were able to demonstrate staffing levels (or thresholds) below which facilities were at substantially greater risk for quality problems.

---

1 Written by Andrew M. Kramer MD and Ron Fish MBA at the Center on Aging, University of Colorado Health Sciences Center under subcontract to Abt Associates. Analytic and design assistance were provided by Michael Lin MPH, and the data files were constructed by Abt Associates and the Data Analysis PRO (DataPRO). Review of analytic strategies was provided by Marvin Feuerberg PhD, Barbara Maynard PhD, and TEP members including Charles Phillips PhD, John Nyman PhD, and Eric Tangalos MD. Editing and document preparation by Lisa Lampinen.

2. These levels were approximately 2 hours per resident day for nurse's aides, .75 to 1.0 hours per resident day for RNs and LPNs combined, of which RNs were required between .20 and .45 hours per resident day. Approximately 54 percent of facilities fell below this nurse's aide standard, 23-56 percent of facilities fell below the RN and LPN standard, and 31-67 percent of facilities fell below the RN standard. Thus, substantial increases in staffing would be required.

3. The minimum staffing levels appeared to be sensitive to case mix, requiring a system to classify facilities into different categories.

The Phase 1 results were presented with considerable caution due to several limitations in the Phase 1 analysis. First, findings were based on facilities in only 3 states from the years of 1996 and 1997 (total number of facilities 1,786). The MDS information used to measure quality and case mix was from the case mix demonstration states, restricting the analyses of quality measures based on MDS data to the states of New York and Ohio (about 1,200 facilities) because Texas did not have an adequate number of facilities collecting MDS data. Even in these two states, several versions of the MDS were used over this period and we could not reliably match MDS to claims data. Finally, there were some inconsistencies among the results from the different states making it hard to generalize the findings about specific staffing levels. While these limitations do not negate the Phase 1 findings, they argue for confirming the results on a larger sample for the purposes of national policy considerations.

This chapter provides results on a larger sample including 10 states using data from 1999. Data quality and data file construction are improved and the analysis refined. While we are confident about the results reported here, additional analyses are continuing to further investigate selected issues (e.g. case mix systems).

2.2 Methods

2.2.1 Design

The study is designed to examine associations between nursing home staffing levels, wages, turnover and retention measured at the facility level, and quality measures that are aggregated from the resident level to the facility level. Because staffing measures are only available at the facility level, the unit of analysis is the facility and quality measures represent facility rates. Recognizing that staffing measures probably do not have a linear relationship with quality, the design included use of continuous quality and staffing measures, as well as quality measures categorized into deciles and staffing measures categorized into dichotomous splits (e.g. < 2.0; ≥ 2.0 nursing assistant hrs per resident day). Staffing measure splits were tested systematically for each quality measure and each available staffing type to identify the split with the strongest association between staffing measures and quality (referred to as a threshold). Controlling for facility case mix was essential in elucidating the association between staffing measures and quality measures. Without adequate control for case mix,

facilities that staff most heavily score worse on the quality measures merely because their residents have the greatest care needs and are at greatest risk for poor outcomes. Thus, risk adjustment for each quality measure was emphasized in these analyses.

2.2.2 Sample

Although a national probability sample would have been preferred, the staffing data for this study were obtained from state Medicaid cost reports, requiring us to select the sample from representative states where we could access Medicaid staffing data. We used Medicaid cost report data for staffing measures because in the Phase 1 study these data were found to be considerably more valid at the facility level particularly at the low end of staffing.[10] We tried to select facilities in all 10 CMS regions but in Region 3 and 4 we could not identify states that could provide Medicaid cost report information in an automated format for the study. Thus, the following states were selected: Massachusetts and Maine (Region 1), New York (Region 2), Illinois and Ohio (Region 5), Texas (Region 6), Iowa (Region 7) Colorado (Region 8), California (Region 9), and Washington (Region 10). For these analyses, we pooled data from all these states, which represent different types of nursing home reimbursement systems for Medicaid, a range in number of beds per thousand elderly (405 in California to 873 in Iowa), a range in facility ownership (47 percent for-profits in New York to 81 percent in Texas), and a range in urban facilities (26 percent in Iowa to 98 percent in Massachusetts), as well as a range in occupancy rates (72 percent in Texas to 96 percent in New York). Quality measures and staffing levels by staffing types were available for all 10 states. Staffing wages were not available in Iowa or Maine. Retention and turnover staffing measures were available only for California and only for nursing assistant staffing type.

Within these states, we selected two different samples of nursing home residents. To study the relationship between staffing and quality of care for short-stay Medicare admissions to nursing homes, we selected a Medicare admission sample. This sample included only Medicare patients and therefore only included facilities that admitted Medicare patients. Because staffing data were obtained from the Medicaid cost report, we could only include facilities that also completed Medicaid cost reports. The second sample is targeted at examining the associations between staffing and quality of care for long-stay residents in nursing homes many of whom are funded by Medicaid or private pay. Care issues differ for the long-stay population and they are often treated on different units than the short-stay population making it imperative to look at quality measures related to both types of care.

We used the Medicare admission sample to evaluate hospitalization rates for selected conditions within 30 days of admission to the facility. The sample included all SNF stays from the study states in 1999. We selected the sample from the National DataPRO File, which included claims data (Medicare Standard Analytic Files Part A) corresponding to the qualifying hospital stay, SNF stay, and any rehospitalizations linked to MDS data for each SNF admission. This stay-level file was then aggregated to the facility level to assess hospitalization rates for selected conditions. Facilities with fewer than 25 Medicare admissions over the calendar year were excluded from the sample because hospitalization

rates for most conditions were unstable with denominators smaller than 25. The total number of facilities dropped from 5,825 to 4,688 when we restricted the analyses to facilities with at least 25 Medicare admissions. The total number of facilities dropped to 3,632 when we matched the claims data to the Medicaid Cost Report data that provided the staffing information. The total number of facilities dropped to 3,478 for the wage analysis. The total number of facilities used in the staffing turnover and retention analysis was 631. Unmatched facilities were more likely to be hospital-based, have fewer beds, and be non-profit vs. for-profit.

The long-stay resident sample included all residents in the MDS 2.0 database with at least two assessments completed 90 days apart during 1999. By definition, these were long-stay nursing home residents for whom values on the quality measures could be computed over a 90-day interval. Each resident was included in this file only once during the one-year interval, even if he or she remained in the facility for the full year (consecutive 90-day intervals for the same patient would not be mutually exclusive). Only facilities with at least 25 individuals were included in this sample. This restriction reduced our sample size from 6,483 to 6,141. The number of facilities dropped to 5,294 when we matched the quality measure data with Medicaid Cost Report staffing data. The wage analysis was conducted with 4,870 facilities while the staffing turnover and retention analysis was performed with 954 facilities.

2.2.3 Quality Measures

Selection of Quality Measures

The criteria used for selecting quality measures included the following:

1. The quality construct was likely to be affected by nurse staffing;
2. A sufficiently high incidence rate was found such that the measure was stable;
3. Identifiable risk factors were identified for which we could adjust; and
4. We expected secondary data to be accurate based on available information.

We selected the final set of quality measures included in this report by considering and testing potential measures and narrowing the list based on these criteria. For example, we initially considered total hospitalization rates as a quality measure but were concerned that this construct might not be affected by nurse staffing because many hospitalizations are not potentially avoidable, such as strokes, and acute myocardial infarction, or are elective (Criterion 1). Furthermore, risk factors for hospitalization due to any cause are more difficult to identify than for disease-specific events (Criterion 3). We also considered hospitalization for fractures and falls in the last 30 days from the MDS as measures of falls in nursing homes, but the incidence rates for these measures were so low that the measures were not stable (Criterion 2). We relied on incidence rather than prevalence measures (e.g., new pressure ulcers rather than presence of a pressure ulcer or changes in status), because incidence rates are more directly attributable to facility care than prevalence rates. Incident measures for some potentially interesting quality constructs were too low to be stable (e.g.

new indwelling catheters, new restraints). We could not identify risk factors that predicted whether individuals would decline in physical function and care resistance so these measures were not included (Criterion 3).

Hospital transfer measures

The denominator for the hospital transfer quality measures was the number of Medicare admissions to the nursing home during the calendar year. The numerators for the five hospital transfer measures were the number of nursing home admissions who were admitted to the hospital within 30 days for CHF, electrolyte imbalance, respiratory infection, UTI, and sepsis. These diagnoses could be listed as either the primary or secondary diagnosis for the hospitalization.

Hospitalization for CHF includes heart failure regardless of the underlying cause, which is generally damage to the heart from prior heart attack or valvular heart disease. Congestive heart failure is a chronic illness that is the leading cause of hospital admission for elderly persons. The role of nursing home staff in treating CHF involves preventive measures to avoid declining health and early identification of signs and symptoms of CHF that should be brought to the attention of a physician to avoid hospitalization. The prevention side consists largely of proper administration of medications, which would be the responsibility of an LPN or RN. Nurse's aides might help to avoid hospitalization for CHF by making certain that any fluid and dietary restrictions (e.g., low salt) are followed, and by early recognition of increased shortness of breath or increased edema. Nurse's aides and LPNs who see residents frequently could observe breathing difficulties and increased fluid accumulation. The most influential role of RNs might entail the supervision of the nurse's aides and the follow-up on their concerns about particular resident's conditions.

Electrolyte imbalance includes any disorders of the body's fluids or electrolytes (e.g., salt and potassium). Many of the hospitalizations for electrolyte imbalance result from dehydration (fluid depletion) or fluid overload in individuals with CHF. Less common medical conditions relating to kidney disease or acid-base status can also affect electrolyte imbalance. Nurse's aides play a major role in preventing hospitalization for electrolyte imbalance by proper hydration and assistance with eating, while LPNs may play a role in proper medication administration and early recognition that an individual's physical and mental status is declining. RNs are essential for oversight and training of nurse's aides, as well as following up on any potential problems. Furthermore, treatment of electrolyte disorders in the nursing home may be possible through administration of IV fluids if sufficient licensed staff are available in the facility.

Respiratory infections include pneumonia that may be either bacterial or viral, and upper airway infections like bronchitis. While respiratory infection is an acute illness, it occurs with great frequency in individuals with chronic pulmonary disease. In frail, elderly persons who have difficulty swallowing, pneumonia often occurs as a result of aspiration in which food is regurgitated and brought into the lungs, particularly if the individual is not positioned properly. Thus, once again nurse's aides play a major role in helping to prevent aspiration

pneumonia through proper positioning and feeding and reducing spread of contagious respiratory infections through proper infection precautions. LPNs and RNs play a valuable role, not only in supervising aides, but in assuring that all individuals receive both the pneumovax and influenza vaccination, and in enforcing appropriate precautions so that infections do not spread throughout the facility. In addition, early recognition of respiratory infection symptoms, contacting the physician, and initiation of antibiotics are critical to successful treatment of pneumonia.

Urinary tract infections (UTI) include infections of the bladder, kidney, prostate, urethra, or any other part of the urinary tract. These are bacterial infections, which often occur chronically in individuals at high risk, such as persons with urinary catheters or urinary obstruction. While bacteria in the urine may occur without an infection requiring treatment, any time there is an associated fever, discomfort, incontinence, or acute confusion, the UTI requires immediate treatment. Prevention of UTIs involves proper hydration and careful hygiene, including regular bathing, which are the responsibilities of the nurse's aide. Sterile procedures for urinary catheter care are essential responsibilities of LPNs and RNs. Early recognition of the signs and symptoms of UTI can avoid hospitalization by prompt physician contact and initiation of treatment. This requires attention from all staff and sufficient LPN and RN staff to supervise aides and promptly follow up on any atypical resident behavior (e.g., confusion) which might indicate an unrecognized UTI.

Sepsis includes infection of the bloodstream from any bacteria. The source of bloodstream infections is often a UTI, wound infection, or respiratory infection; however, any infection can result in sepsis if not promptly treated. Sepsis can be avoided if infections are identified and treated before bacteria become blood-borne, requiring nurse's aides to assist in preventing primary infections and to recognize any of the initial signs and symptoms of infection. RNs and LPNs must respond promptly when any symptoms of an infection are identified. This requires supervision of the nurse's aides, and attentiveness to the condition of the residents in the nursing home. Once sepsis occurs, the nursing home must hospitalize the patient for treatment, but the mortality rate even after hospitalization is extremely high.

Thus, all of these quality measures meet the first and most important criterion of a potential association with staffing. They all represent incident events in the nursing home of reasonably high prevalence. Claims data are a reasonably good information source because they are used for payment purposes and are audited. While there may be some ambiguities (e.g., CHF episodes coded as electrolyte imbalance), use of either the primary or secondary diagnosis helps to minimize the effects of coding practices with respect to the primary diagnosis.

Long-stay quality measures

The long-stay quality measures were derived from MDS data. They mostly represented changes in health or behavioral status occurring over a 90-day interval where literature or clear hypotheses support relationships to staffing. At the resident level, these were dichotomous variables denoting either improvement or decline. Improvement was denoted

when health or behavioral status was better at the second time-point compared to baseline, whereas staying the same or declining represented no improvement. Decline was denoted when health or behavioral status was worse at the second time-point compared to baseline, while staying the same or improving represented no decline.

The resident-level measures were then aggregated to the facility level. The denominator for these measures was the number of individuals in the facility for whom we had two MDS assessments at 90-day intervals. However, exclusions were necessary for specific quality indicators that were impossible for certain individuals. For example, a resident who was at the best possible status could not improve, and thus was excluded from that improvement quality measure. Similarly, a resident who was at the worst level could not decline and so was excluded from that decline measure. These selectively reduced the denominator for individual measures. The numerator was the number of residents who improved or declined.

Change in ability to perform basic ADLs is an important measure of nursing home quality. Though decline in ability to perform basic ADLs has not been shown to be related to nurse staffing, improvement in ability to perform basic activities of living has been repeatedly shown to be related to skilled (or licensed) nurse staffing.[3,5] Various standard measures have been validated for assessing functional change. Of these, we chose the Barthel Index[11,12] because it has breadth, has been validated and used widely, and correlates with the ability to live independently in the community. This index includes the following activities: Eating, Dressing, Transfer, Grooming, Toileting, Bathing, Walking, Bowel Continence, and Bladder Continence (because the MDS does not assess the ability to climb stairs, a 90-point version of the Barthel was used). We converted the relevant portions of the MDS to a Barthel Index. This conversion has been validated against an independent assessment of function performed by a research nurse. We considered a change of 10 or more points clinically meaningful.

Pressure ulcers are wounds caused by excessive and prolonged pressure on skin. They are such an important cause of morbidity and mortality that we included pressure ulcer incidence as a quality measure in spite of evidence that prevalence has been refractory to multiple interventions.[13,14] Pressure ulcers are graded by the depth of the wound. Stage 1 ulcers are persistent redness of the skin over a pressure point; stage 2 involves a break in the skin; stage 3 is defined by penetration of the wound below the skin; and stage 4 is damage to underlying muscle and/or bone. We defined incident pressure ulcer as a stage 2 or greater pressure ulcer which was not present on the baseline MDS. Prevention of pressure ulcers requires mobilizing patients to relieve pressure every 2 hours and keeping skin clean and dry - both are labor-intensive in immobilized residents. Results have been mixed on whether staffing ratios influence incident pressure sore rates.[3]

Because incidence of pressure ulcer is so resistant to intervention, we also assessed incidence of skin trauma. Skin trauma includes abrasions, bruises, burns, cuts, and skin tears. Nursing homes where nurse's aides and nurses feel rushed to provide personal care and assistance are

likely to have higher incidence of these occurrences. More involvement of licensed personnel may prevent problems in this area.

Weight loss is a marker of declining nutritional status, particularly when weight falls below acceptable levels. A body mass index (the ratio of weight in kilograms to height in meters squared, BMI) of less than 21 is associated with increased morbidity and mortality. In the Phase 1 study, significant weight loss was found to be associated with nurse's aide staffing and RN plus LPN staffing using a small sample of facilities for which primary data were available.[8] Eating is also a major element in quality of life. The effect of inadequate staffing on assistance with eating and overall quality of life has been confirmed by many observations of nursing home residents.[15,16] Thus, a decline in weight to a level below a BMI of 21 is an important quality measure influenced by staffing ratios.

Resisting assistance with ADLs is a marker of the personal relationship between residents and staff. According to Bowers[17] and Kayser-Jones[18], patients and nursing staff regard the relationship that develops between a vulnerable adult and her caregiver to be of paramount importance in determining the quality of a resident's life. Residents describe the importance of gentleness, personal engagement, not being rushed and feeling respected. Nursing assistants report that they value having time to promote physical comfort, not make residents wait or rush, and sharing treats or personal stories. We reasoned that over time residents who initially resist assistance with ADLs out of fear or confusion should gradually become more accepting of care if well-trained and supervised staff are available to permit development of personal rapport. Improvement is defined as not resisting assistance with ADLs at the second assessment if resistance had been noted at the first.

2.2.4 Covariates

For the hospital transfer quality measures, we obtained covariate data from two sources. First, we obtained covariate data from the DataPRO files using diagnostic information from institutional admissions occurring before the hospitalization. We chose the covariates based on clinical considerations and available literature and specified the appropriate ICD-9-CM code. If the diagnosis was listed for any stay in the prior six months (either SNF stay or hospital) as either a primary or secondary diagnosis, the case mix covariate was denoted as present for the individual. The reason for using this relatively liberal criterion is that the covariates are all chronic conditions that would persist over time but are frequently under-reported during episodes with different primary diagnoses.

For both the short-stay analysis and long-stay analysis, we obtained covariate data from the Minimum Data Set (MDS). For the short-stay file, we used the five-day MDS that was matched to the SNF claim. For the long-stay file, we used the first of the two MDSs in the 90-day interval as the source of covariate information. A range of covariates were chosen based on the quality measures that were used. These included variables related to demographics, function, resuscitation orders, the cognitive performance scale for

cognition,[19,20] behaviors, and clinical conditions. The set of candidate covariates were reduced using systematic methods described in section 2.2.6.

2.2.5 Staffing Measures

We used four staffing level measures for these analyses: nursing assistant staff hours per resident day, LPN hours per resident day, RN hours per resident day, and the total licensed staff - sum of RN and LPN hours per resident day. We report staffing level analyses for nursing assistants , total licensed staff (RN+LPN), and the RN staffing level, because for some functions there was widespread substitution between RN and LPN staff. However, LPN qualifications are not equivalent to RNs, and some functions require RNs. Thus, RN staffing levels were also assessed to determine the thresholds within the licensed staff total that require RNs. We used one staffing wage measure, TOTAL (nursing assistant, LPN, and RN added together) wages per resident day. Staffing turnover and retention measures were limited to the nurses assistants because such data were not provided for any other individual staff type and nurses assistants dominated total staffing measures of retention and turnover. Turnover is the percent of total nursing assistants that left the facility during the entire cost report period (usually 12 months). Retention is the percent of total nursing assistants that remained with the facility for the entire cost report period. Turnover could and often did exceed 100% while retention was at or below 100%. More information on data and data sources for the turnover and retention measures is provided in Chapter 4. The staffing measures did not include administrative FTEs such as Directors of Nursing.

Staffing data from the Medicaid cost reports were used rather than OSCAR data because they were found to be more valid in a comparison with payroll data collected for a sample of facilities in Ohio.[10] The correlation between Medicaid cost report data and payroll data was 0.73 for RN staffing, 0.64 for LPN staffing, and 0.39 for nurse's aide staffing. The Medicaid cost report data also tended to report higher RN, LPN and nurse's aide hours per resident day than payroll data in the 20 percent of facilities with the lowest staffing levels. Thus, nurse's aide staffing data are less accurate than data for RNs and LPNs, and staffing for the lowest-staffed facilities is probably overstated in these analyses. We eliminated extreme outliers (total hours per resident day < .5 or > 12) which comprised only 1 percent of the samples of facilities with Medicaid staffing data.

2.2.6 Analysis

Estimating facility risk scores

To estimate a facility risk score related to each quality measure, reflecting the average risk of poor outcomes based on the characteristics of the residents in the facility, we started the analysis at the resident level. For each quality measure, we generated a resident-level risk score, which reflected the probability of a bad outcome, using relevant covariates in the data set. The covariates for the risk scores were selected based on a combination of identifying known risk factors from the literature and clinical experience, an examination of correlations among the covariates and the outcome measures, and stepwise logistic regression procedures.

A set of candidate covariates was selected first, and then correlations among these candidate variables and stepwise regression models were used to select final variables for each risk model. These models were estimated at the resident level across all facilities including samples of approximately 594,000 residents for the short-stay sample and 134,000-461,000 residents for the long-stay sample (varying slightly depending upon the quality measure of interest and the selected covariates). These resident risk score models are contained in Appendix A for each quality measure.

Once the resident risk score models were finalized, the facility average risk score for each measure (or expected rate for the quality measure) was calculated as the mean risk score for all residents in the facility. As expected, the aggregated mean risk score was correlated with the actual facility rate for each outcome. This method differs from the risk adjustment method used in the Phase 1 analysis because it uses resident-level data that were not available in the prior analysis. The advantage of first developing resident-level risk scores is that the relationship between risk factors and outcomes was much stronger at the resident level than they were when the individual covariates were aggregated to the facility level first and used in a model to predict facility-level quality measures. Thus, these models were more effective at accounting for the variance in quality explained by facility case mix.

Determining associations between staffing levels and quality measures

Because facility case mix confounds simple associations between staffing and quality, we used the facility mean risk score to take case mix into consideration while analyzing associations between staffing levels and quality. All of these analyses were conducted separately for the individual quality measures.

We began by stratifying facilities into risk categories, including quintiles and thirds, and using CART® (a robust decision-tree application) to identify split points for each category of facilities. As a dependent variable we used both the continuous quality measure rate for each facility and whether the facility appeared in the worst decile or quintile of facilities. Across the board, models were much stronger for predicting the worst decile. Using the range of values that were generated by the decision-tree analysis, we used logistic regression to test the association between these staffing levels and quality while controlling for the facility mean risk score for the quality measure. With the continuous quality measure rate as a dependent variable, we used ordinary least squares regression, whereas logistic regression was used to investigate associations between staffing and whether the facility was in the worst decile with respect to the quality measure. Once again, the worst decile models were considerably stronger.

To identify the minimum staffing level for each quality measure, we tested a range of staffing levels in each of these models. Nursing assistant staffing levels were tested at increments of 0.1 hours per patient day (e.g. 1.9, 2.0, 2.1, etc) and both RN and licensed staff levels were tested using increments of 0.05 (e.g. 0.4, 0.45, 0.5, etc.). A consecutive range of staffing levels was usually significantly associated with whether a facility was in the worst decile. This range of staffing levels with significant odds ratios represented staffing levels where

quality improved incrementally. Thus, these were all points on the steep part of the curve in the relationship between staffing and quality, where gains in quality were achieved as staffing levels were increased. At the top of these ranges, when there was no longer a significant increase in quality associated with additional staffing, we reasoned that no further benefits of increased staffing were apparent at least with respect to preventing quality problems. Thus, these represented staffing thresholds.

To combine the staffing thresholds across quality measures, we weighted the measures based on the prevalence of the quality measures across facilities. We reasoned that the prevalence of the outcome problems across facilities reflected the proportion of individuals at risk for these problems and the proportion of staffing resources that should be targeted on these problems. While a problem might be very important and require a high-level of staffing to avoid it, if it occurs only rarely, facilities should not be staffed to address that problem in all residents; only in the proportion of residents at risk. This weighting was conducted separately for the short-stay quality measures and the long-stay quality measures because the relative prevalence of these events in facilities depends on the mix of short-stay and long-stay patients in a facility. Thus, we developed weighted average thresholds pertinent to short-stay patients, and weighted average thresholds for long-stay patients.

In addition to these models, we tested whether these staffing levels were associated with quality when included in the same model, i.e. controlling for whether staffing thresholds for other types of staff were met. We also conducted some analyses to determine whether different staffing levels could be identified that were associated with facilities in the best decile or quintile.

Analysis of thresholds by case mix categories

To evaluate possible threshold variation by facility case mix categories, we used the facility mean risk score to create strata based on risk for each outcome. We tried classifying facilities into 3, 5, and 10 categories. Within these categories, we determined risk-adjusted thresholds for the various quality measures to determine whether different thresholds would be identified for facilities with different risk for poor outcomes.

Determining associations between staffing wages and quality measures

As with the staffing level analysis, we used the facility mean risk score to take case mix into consideration while analyzing associations between staffing wages and quality. Total wages was used rather than separate wages for different staffing types (i.e. nursing assistant, RN, LPN). All of these analyses were conducted separately for the individual quality measures.

Facilities were stratified into worst decile vs. all others for each quality measure. As in the previous analysis with staffing levels, using a dichotomous quality measure rather than a continuous measure resulted in a more robust analysis. Total wages per resident day, a continuous variable, was transformed into nine dichotomous variables based on decile split points. Continuous total wages as well as the decile levels were tested individually as independent variables against the dependent quality measures of the worst decile using

logistic regression in order to identify any overall relationships as well as possible minimum wage levels for each quality measure.

Determining associations between staffing turnover and retention and quality measures
As with the staffing level and wage analyses, we used the facility mean risk score to take case mix into consideration while analyzing associations between staffing turnover and quality as well as retention and quality. All of these analyses were conducted separately for the individual quality measures.

Again, facilities were stratified into worst decile vs. all others for each quality measure for reasons previously discussed. Turnover and retention were tested as independent continuous variables against individual dichotomous and continuous quality measures using logistic regression and linear regression. Turnover and retention were also transformed into dichotomous split variables and tested against the same dependent worst decile quality measure using logistic regression. Initially, decile splits were used to create the independent dichotomous variables. Based on results from decile splits, further dichotomous variables were created to optimize the analyses. These analyses were done to identify any overall relationships as well as possible thresholds for % turnover and retention where quality was adversely impacted.

Determining the percent of facilities below the highest significant threshold by staff and stay type
Using the weighted staffing level threshold values developed out of these studies in conjunction with OSCAR staffing levels for all fifty states, the percent of facilities staffing at or below these thresholds was calculated for short and long stay by staffing type. The percent of facilities falling at or below these thresholds was also calculated when all three staffing types were at or below and also when any one of the three staffing types fell at or below for short and long stay samples.

2.3 Results

2.3.1 Staffing Levels

Descriptive statistics for the short-stay facility sample (n=3,632) and long-stay facility sample (n=5,294) are provided in Tables 2.1 and 2.2. In addition to means and standard deviations, the percentile values are given for the top and bottom 1 percent, the quartiles, and the highest decile. These percentiles show the nature of the distribution of the variables. For example, in Table 2.1 the variable for nurse's aide hours per resident day was close to normally distributed increasing from about 1 hour per resident day to a median of about 2 hours per resident day and reaching about 3 hours per resident day between the 90th and 99th percentile, with the 99th percentile at about 3.7 hours per resident day. Licensed staff (RN+LPN), however, was skewed ranging from about .4 at 1 percent up to about 1.0 at the median value, with the 99th percentile about three and a half times the median. The hospitalization transfer quality measures were even more skewed with very little change between 1 percent and the median, but the 90th percentile was between two to five times the

value of the median, and the 99th percentile was about four to ten times greater than the median. Similar distributions to these were found for the long-stay quality measures (Table 2.2).

Table 2.1
Staffing, Quality Measure, and Covariate Descriptive Statistics for the Short-Stay Facility Sample (n=3,632)

	Mean	S.D.	Percentiles					
			1%	25%	50%	75%	90%	99%
Staffing Type (Hours per Resident Day)								
Aide	2.07	.521	0.96	1.76	2.02	2.32	2.67	3.73
LPN	0.65	.293	0.04	0.46	0.63	0.80	0.96	1.51
RN	0.47	.422	0.01	0.23	0.38	0.59	0.83	2.41
Licensed Staff (RN plus LPN)	1.11	.526	0.40	0.83	1.02	1.26	1.53	3.58
Hospital Transfer Quality Measures*								
Congestive Heart Failure[1,2,5,6, 8,9,12]	.057	.037	0.00	0.03	0.05	0.08	0.11	0.17
Electrolyte Imbalance[1,2,3,4,5,6,7,8,10,11]	.064	.044	0.00	0.03	0.06	0.09	0.12	0.2
Respiratory Infection[2,3,4,5,6,7,8,11]	.049	.037	0.00	0.02	0.04	0.07	0.10	0.16
Sepsis[2,3,4,6,8,10,11]	.020	.024	0.00	0.00	0.01	0.03	0.05	0.11
Urinary Tract Infection[1,2,4,5,6,7,8,11]	.044	.037	0.00	0.02	0.04	0.06	0.09	0.16
Covariates								
1 Age	80.6	3.39	69.7	79.1	81.0	82.8	84.3	86.9
2 Barthel ADL score	33.1	8.80	14.8	26.7	32.9	39.2	44.5	52.8
3 Bedfast	.260	.154	0.02	0.14	0.23	0.35	0.48	0.69
4 Cognitive Performance Scale	1.63	.625	0.51	1.19	1.55	1.99	2.45	3.43
5 Congestive Heart Failure	.322	.090	0.11	0.26	0.32	0.38	0.47	0.56
6 Do not resuscitate	.443	.216	0.03	0.27	0.44	0.60	0.74	0.91
7 Dysphagia	.085	.080	0.00	0.03	0.06	0.11	0.20	0.38
8 Feeding tube present	.125	.102	0.00	0.05	0.10	0.17	0.27	0.46
9 Hypertension with complications	.059	.053	0.00	0.02	0.05	0.08	0.13	0.24
10 Renal failure	.089	.059	0.00	0.05	0.08	0.12	0.16	0.28
11 Requires assistance to eat	.388	.156	0.10	0.27	0.37	0.49	0.60	0.81
12 Respiratory disease	.262	.086	0.08	0.20	0.26	0.32	0.38	0.49

*Numbers correspond to the numbered covariates that were used in the resident risk score models.

These distributions show that the top (or worst) decile of facilities for the quality measures were generally outliers relative to the rest of the distribution. Thus, the worst decile for each quality measure, after controlling for risk, is a reasonable indicator of poor quality. Facilities in the worst decile all had quality measure rates that were at least double and up to ten times the median value for the quality measures. Furthermore, the continuous quality measures, which demonstrated less variation across much of the continuum, may not be useful to assess

facility performance. For risk adjustment, ordinary lease squares regression assumes the dependent variable is normally distributed, which is not the case with these continuous quality measures, further supporting the use of a dichotomous dependent variable (i.e. the worst decile) and logistic regression for risk adjustment.

The overall rehospitalization rate within 30 days was about 16 percent. Because we included diagnoses that were listed as the primary or secondary hospital discharge diagnosis, the sum of the percentages for these five diagnoses exceeds the percent hospitalizations for all diagnoses (Table 2.1). For each diagnosis, the covariates that were used in the resident risk score model are enumerated, referencing the covariate numbers provided below in the table. The logistic regression models are provided in Appendix A.

Older persons were less likely to be hospitalized for UTI and electrolyte imbalance and more likely to be hospitalized for CHF, whereas the Barthel Index, which represents independence in function, was negatively associated with hospitalization for all measures. Bedfast patients, patients with cognitive impairment based on the Cognitive Performance Scale (CPS), persons with chronic underlying CHF, difficulty swallowing (dysphagia), feeding tubes, eating disability, hypertension with complications, renal failure, and respiratory disease all tended to have higher rates of hospitalization for one or more of the conditions. Across the board, persons with do not resuscitate orders were less likely to be hospitalized. While some may argue that controlling for presence of a feeding tube is not appropriate because prevalence of feeding tubes suggest the facility has quality problems, these values were taken from the initial 5-day MDS and the entire window of time that we are examining is within 30 days of admission. In sub-acute care, it is not appropriate to hold facilities accountable for chronic or acute patient health problems upon admission. If residents are admitted with swallowing impairments or impairments in their ability to eat or take in nutrients, feeding tubes may be necessary for assuring adequate intake to avoid dehydration and declining nutrition. Thus, in this population these may be reasonable markers for patient acuity, although they may reflect quality of care for long-stay patients where alternative care strategies may be more feasible.

A longer list of covariates is provided for the long-stay models because the measures are more varied and different sets of covariates are required for risk adjustment (Table 2.2). Older persons were less likely to have good outcomes and more likely to have worse outcomes. The Barthel Index was used in two different ways. Greater independence in the Barthel Index as a continuous measure was associated with less skin trauma. The relationship between baseline functional status and functional improvement, however, was not linear, so a dichotomous split in the Barthel was used in these models. Residents with a score between 25 and 70 were denoted by 1 and those at either end of the distribution (i.e. extremely functionally impaired or very independent) were denoted by 0. This dichotomous variable was positively associated with functional improvement because residents in this middle range were more likely to improve in function than those at the extremes. As we might expect, severe cognitive impairment was also associated with quality problems or lack of improvement for a number of measures. The other covariates were more related to the specific indicators.

Table 2.2
Staffing, Quality Measure, and Covariate Descriptive Statistics for the Long-Stay Facility Sample (n=5,294)

	Mean	S.D.	Percentiles					
			1%	25%	50%	75%	90%	99%
Staffing Type (Hours per Resident Day)								
Aide	2.02	0.58	0.85	1.71	1.99	2.30	2.66	3.66
LPN	0.63	0.29	0.02	0.45	0.61	0.79	0.95	1.37
RN	0.40	0.32	0.00	0.19	0.34	0.55	0.77	1.33
Licensed Staff (RN plus LPN)	1.03	0.41	0.28	0.78	0.97	1.20	1.44	2.34
Quality Measures*								
Functional Improvement[3,6,7,9,10,13,16]	.075	.045	0.00	0.04	0.07	0.10	0.14	0.21
Incident Pressure Ulcer[3,4,7,8,9,15]	.031	.025	0.00	0.01	0.03	0.04	0.06	0.11
Resisting Care Improvement[1,2,9,12]	.301	.186	0.00	0.16	0.29	0.43	0.57	0.71
Skin Trauma[3,5,7,8,9,10,14,15,17]	.100	.095	0.00	0.02	0.08	0.15	0.24	0.40
Weight Loss[3,4,11,13]	.055	0.54	0.00	0.00	0.05	0.08	0.12	0.25
Covariates								
1 Abusive Behavior, Physical	.082	.064	0.00	0.04	0.07	0.11	0.16	0.29
2 Abusive Behavior, Verbal	.127	.093	0.00	0.06	0.11	0.17	0.24	0.45
3 Age	81.6	5.62	57.2	79.9	82.9	85.0	86.4	89
4 Ambulation Dependent	.552	.153	0.04	0.46	0.56	0.65	0.74	0.88
5 Barthel Index	37.7	10.7	17.5	30.9	36.9	43.5	49.7	79.7
6 Barthel Index between 25 and 70	.386	.112	0.09	0.31	0.38	0.46	0.53	0.65
7 Bed Mobility (≥3, Extensive Asst.)	.363	.158	0.00	0.26	0.35	0.46	0.56	0.81
8 Body Mass Index < 21	.382	.108	0.13	0.31	0.38	0.44	0.51	0.71
9 Cognitive Performance Scale ≥ 4	.295	.125	0.04	0.21	0.28	0.36	0.45	0.67
10 Do not resuscitate	.532	.261	0.00	0.33	0.56	0.75	0.86	0.99
11 Eating (≥ 3, Extensive Asst.)	.243	.103	0.01	0.18	0.24	0.30	0.37	0.52
12 Incontinence, Bladder	.581	.131	0.14	0.51	0.59	0.67	0.74	0.87
13 Incontinence, Bowel	.467	.155	0.04	0.37	0.47	0.57	0.66	0.81
14 Resists Care	.210	.129	0.02	0.12	0.19	0.28	0.38	0.62
15 Transfer Assistance Needed	.657	.167	0.05	0.56	0.67	0.77	0.86	0.96
16 Visual Impairment	.178	.115	0.00	0.10	0.16	0.24	0.33	0.53
17 White	.849	.215	0.06	0.80	0.95	0.99	1.00	1.00

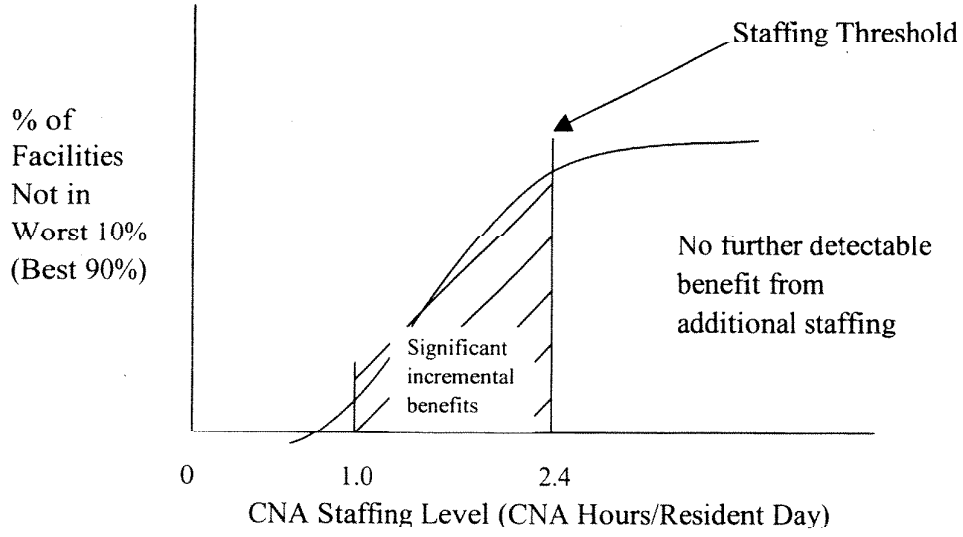
* Numbers correspond to the numbered covariates that were used in the resident risk score models

The facility profile in the samples differed from the universe of facilities in ways that we would expect given the data sources. Only about 4 percent of the short-stay sample and 3 percent of the long-stay sample were hospital-based facilities. Non-profit facilities were slightly underrepresented at about 23 percent of the total which is consistent with fewer hospital-based providers (about 74 percent were for-profits and government facilities were about 3 percent of the sample). The short-stay sample was about 81 percent urban and the long-stay sample was about 75 percent urban, which is slightly higher than national rates

because we excluded facilities with fewer than 25 cases. We cannot compare staffing levels in our sample with national averages because national average staffing levels are only obtainable from OSCAR, which is provider reported and lacks comparability to Medicaid cost report data.

Although significant associations were found between staffing and quality measure rates across the staffing continuum, for the reasons previously discussed we report the findings related to the logistic regression models for the likelihood of being in the worst 10 percent of facilities. For each quality measure and each type of staff, we present statistics for the levels representing thresholds below which facilities were significantly more likely to be in the worst 10 percent and above which staffing increases were not associated with continued reductions in quality of care problems (Figure 2.1). Although the slope of the curves and threshold levels differed among quality measures, they generally followed a similar pattern to the depiction in Figure 2.1. Staffing levels below these thresholds were significantly associated with incremental improvements in quality, demonstrating the positive association between staffing and quality until the threshold was reached and no further benefits were detected. We report the results of the logistic regression models for the threshold staffing levels in subsequent tables.

Figure 2.1
Illustration of Threshold in the Association between Staffing and Likelihood that Facilities are Not in the Worst 10% of Facilities



For the short-stay sample, nurse's aides staffing thresholds ranged between 2.3 for electrolyte imbalance and 2.4 for all others (Table 2.3). Licensed staff levels ranged between 1.05 hours per resident day for respiratory infection up to 1.30 hours per resident day for sepsis. RN hours per resident day thresholds were at 0.55 hours for all measures. These RN hours would be a portion of the licensed staff hours not in addition to the licensed staff hours. The

weighted averages of these staffing thresholds for the short-stay population totaled 2.37 hours for nursing assistants, 1.14 hours for licensed staff, and .55 hours for RNs.

Table 2.3
Staffing Thresholds Below Which Facilities Were at Increased Likelihood of Being in the Worst 10% for Short-Stay Quality Measures and Above Which There Were No Additional Improvements in Quality*

Staff Type	Staffing Hours per Resident Day	Adjusted Odds Ratio (95% CI)†
Aide		
Congestive Heart Failure	2.40	1.47 (1.02 - 2.11)
Electrolyte Imbalance	2.30	1.45 (1.02 - 2.04)
Sepsis	2.40	2.43 (1.51 - 3.92)
Urinary Tract Infection	2.40	1.53 (1.01 - 2.30)
Weighted Average	2.37	
Licensed (RN+LPN)		
Electrolyte Imbalance	1.15	1.40 (1.04 - 1.89)
Respiratory Infection	1.05	1.31 (1.01 - 1.71)
Sepsis	1.30	1.49 (1.02 - 2.18)
Urinary Tract Infection	1.15	1.60 (1.17 - 2.18)
Weighted Average	1.14	
RN		
Electrolyte Imbalance	0.55	1.41 (1.01 - 1.99)
Sepsis	0.55	1.44 (1.02 - 2.02)
Urinary Tract Infection	0.55	1.46 (1.03 - 2.06)
Weighted Average	0.55	

* Includes 3,632 facilities from short-stay facility sample
† All odds rations are significant at $p < .05$ or lower

Nurse's aide staffing hours for the long-stay quality measures ranged from 2.4 hours per resident day for functional improvement to 3.1 hours per resident day for weight loss. Licenses staff levels ranged from 0.95 hours per resident day for weight loss up to 1.55 hours per resident day for functional improvement. RN thresholds ranged from 0.6 hours per resident day for pressure ulcers to 0.8 hours per resident day for functional improvement and resisting care improvement (Table 2.4). The weighted threshold for long-stay quality measures were 2.8 hours for aides, 1.3 hours for licensed staff, and .75 hours for RNs.

Table 2.4
Staffing Thresholds Below Which Facilities Were at Increased Likelihood of Being in the Worst 10% for Long-Stay Quality Measures and Above Which There Were No Additional Improvements in Quality.*

Quality Measure	Staffing Hours per Resident Day	Adjusted Odds Ratio (95% CI)†
Aide		
Functional Improvement	2.40	1.34 (1.03- 1.73)
Incident Pressure Ulcer	2.80	1.60 (1.08 - 2.36)
Resisting Care Improvement	2.80	1.54 (1.01 - 2.36)
Skin Trauma	2.80	1.59 (1.04 - 2.42)
Weight Loss	3.10	2.25 (1.04 - 4.83)
Weighted Average	2.78	
Licensed (RN+LPN)		
Functional Improvement	1.55	1.79 (1.08 - 2.73)
Resisting Care Improvement	1.35	1.53 (1.14 - 2.05)
Skin Trauma	1.15	1.31 (1.06 - 1.63)
Weight Loss	0.95	1.23 (1.02 - 1.49)
Weighted Average	1.30	
RN		
Functional Improvement	0.80	1.54 (1.04 - 2.27)
Incident Pressure Ulcer	0.60	1.33 (1.06 - 1.69)
Resisting Care Improvement	0.75	1.68 (1.18 - 2.39)
Weighted Average	0.75	

* Includes 5,294 facilities from long-stay facility sample
† All odds ratios are significant at $p < .05$ or lower

Using the facility mean risk score to stratify facilities into case mix categories did not yield a staffing threshold that differed significantly by stratum. The proportion of facilities in the worst decile varied by strata, as expected, with the facilities in the highest risk strata more likely to be in the worst decile. However, staffing thresholds above which there were no further quality improvements were similar for all strata, whether facilities were divided into deciles, quintiles, or thirds.

The staffing level thresholds developed in these analysis for both the short and long stay samples were applied to the 1999 national OSCAR database. Decision rules to exclude facilities with unreliable staffing measures were applied which dropped the number of facilities from 16,290 to 15,844. Table 2.5 provides the percentage of facilities at or below these thresholds for staffing type individually and in two summary forms.

Table 2.5
Percent of Facilities At or Below Staffing Thresholds for the National Sample (n=15,844)

Sample Type	All Three Staffing Types*	At least One Staffing Type†	Nursing Assistant	RN	RN/LPN
Short Stay	52%	92%	80%	77%	63%
Long Stay	72%	97%	92%	87%	76%

* Those facilities where all three staffing types are at or below their respective thresholds.
† Those facilities where at least one of the staffing types is at or below that staffing type's threshold

2.3.2 Total Wages

Total wages in dollars per resident day for nursing assistant, RN, and LPN was initially stratified into 9 dichotomous variables based on deciles. There were significant associations between dichotomous total wage variables and the worst decile dichotomous Quality Measures. For five of the quality measures, a significant association was found between total wages and whether the facility was in the worst 10% of facilities for the quality measures, after risk adjustment (Table 2.6). For the measures in which these associations were found, the associations between total wages and quality occurred all the way up to the top wage rate decile. Thus, there was nearly a continuous relationship between quality and total staff wages per resident day over the wage range. These findings are not as strong as the staffing level findings because the relationship existed for fewer quality measures.

Table 2.6
Thresholds for Total Wages Below Which Facilities Were at Increased Likelihood of Being in the Worst 10% and Above Which There Were No Additional Improvements in Quality*

Quality Measure	Total Wages ($) per Resident Day	Adjusted Odds Ratio (95% CI)†
Short Stay		
Electrolyte Imbalance	$50	1.58 (1.06 - 2.36)
Sepsis	$57	1.88 (1.08 - 3.28)
Urinary Tract Infection	$57	1.76 (1.02 - 3.04)
Long Stay		
Functional Improvement	$62	2.09 (1.02 - 4.28)
Resisting Care Improvement	$60	1.92 (1.07 - 3.47)

* Includes 3,632 facilities from short-stay sample or 5,294 facilities from the long-stay sample
† All odds rations are significant at $p < .05$ or lower

2.3.3 Nursing Assistant Retention and Turnover

Nursing assistant retention and turnover were initially stratified into 9 dichotomous variables based on deciles. There were significant associations between dichotomous nursing assistant retention variables at different deciles and the worst decile for five selected Quality Measures (Table 2.7). For these five measures, the associations between retention and quality were found at almost all levels of retention up to the top decile, suggesting that these relationships

were virtually continuous. The ordinary least squares regression analysis, confirmed the existence of a linear relationship between retention and quality. The number of quality measures that were associated and the strength of the associations (odds ratios) were much greater for retention than turnover (Table 2.8). The turnover thresholds occurred up until about the third decile, suggesting a threshold for nursing turnover above which there was no further impact on quality. As with staffing levels, the thresholds in the tables were identified based on additional splits beyond the decile level to identify the threshold for a particular Quality Measure.

Table 2.7
Thresholds for % Retention Below Which Facilities Were at Increased Likelihood of Being in the Worst 10% and Above Which There Were No Additional Improvements in Quality*

Quality Measure	% Nursing Assistant Retention / Resident Day	Adjusted Odds Ratio (95% CI)†
Short Stay		
Electrolyte Imbalance	43%	3.06 (1.16 - 8.12)
Urinary Tract Infection	51%	3.66 (1.04 - 12.9)
Long Stay		
Functional Improvement	40%	2.72 (1.07 – 6.90)
Incident Pressure Ulcer	41%	2.54 (1.07 – 6.02)
Resisting Care Improvement	37%	2.48 (1.06 - 5.80)

* N= 631 (Retention data were available for California facilities only.)
† All odds rations are significant at $p < .05$ or lower

Table 2.8
Thresholds for % Turnover Above Which Facilities were at Increased Likelihood of Being in the Worst 10% and Below Which There Were No Additional Improvements in Quality*

Quality Measure	% Turnover per Resident Day	Adjusted Odds Ratio (95% CI)†
Short Stay		
Urinary Tract Infection	47%	1.90 (1.02 - 3.55)
Long Stay		
Incident Pressure Ulcer	46%	1.61 (1.02 - 2.54)

* N=631 (Turnover data were available for California facilities only.)
† All odds rations are significant at $p < .05$ or lower

2.4 Discussion

These findings demonstrated clear associations between nurse staffing levels and quality measures for short-stay Medicare patients as well as quality measures for long-stay nursing home residents. The findings confirm the Phase 1 report findings that for all types of nursing staff, ratios of staff to residents exist below which residents are at substantially higher risk of quality problems. The sample was nationally representative including 10 states, the risk adjustment was more rigorous, and we tested a wide range of staffing levels for each measure providing increased support for these findings.

The relationship between staffing and quality did not appear linear because regression models using continuous quality measure rates were in the same direction but not as strong as logistic regression models predicting the worst decile for each measure. More importantly, logistic regression models over a range of different staffing levels yielded thresholds above which higher staffing ratios were no longer associated with reductions in the likelihood of compromised quality. Such a relationship between staffing levels and quality was hypothesized in the Phase 1 report,[1] and is consistent with the theory that increased numbers of staff can reduce quality problems up until a point, but other issues relating to staffing practices, management, etc. are the major determinants of quality once these staffing thresholds are reached. Thus, our findings suggest that increased staffing levels reduce quality problems until staffing thresholds are met, above which there are no marginal benefits of additional staffing.

The quality measures chosen for this study are conceptually associated with staffing levels. For short-stay patients, hospital transfers of nursing home residents are frequently avoidable.[21-23] For CHF and electrolyte imbalance, nurse's aides play an essential role in proper attention to fluid intake and dietary issues, as well as early recognition of any changes in a person's breathing or edema. Similarly, licensed staff have a major role in medication administration and might be able to detect changes in physical or mental status that are the early warning signs of problems that could lead to hospitalization. RNs play a role not only in oversight of nurse's aides, but also by enhancing the ability of the nursing home to administer IV fluids and thereby avoid hospitalization for dehydration.

Infections, whether respiratory, urinary tract, or sepsis, are the major cause of hospital transfer for nursing home residents. Early recognition of infection in elderly persons can be difficult if they have an underlying dementia or do not experience elevated temperature. Once again, the role of nurse's aides is critical to identify early signs of infection, such as a change in behavior or activity. Licensed staff play a crucial role in early identification of infections, enforcing precautions to prevent infections from spreading throughout the facility, and making sure that treatment is initiated so that sepsis - a life-threatening blood-borne infection - does not result. Well-trained and supervised nursing home staff are more likely to identify early symptoms such as confusion, agitation, or non-specific complaints.

For the long-stay sample, the range of care areas are conceptually linked to staffing and relationships have been found with these measures in previous studies including the Phase 1 report. Functional improvement requires considerable involvement of all nursing staff that come in contact with the resident to maximize opportunities for functional independence. Pressure ulcer prevention and skin trauma depends on labor intensive measures like turning immobile patients every two hours to relieve pressure and carefully cleaning skin. While this might be considered the domain of nursing assistants, we found relationships with both nursing assistant and RN staff - the latter probably necessary for supervision. Improvement in resisting care was found to be associated with staffing in our Phase 1 study. Observational studies have shown how residents are more likely to resist assistance with eating when hurried or fed by staff who are not particularly attentive to their needs.[15,16] When staff have time to be patient and caring, residents were less likely to resist care. This involves not only nursing assistants, but licensed staff, and RNs for supervision. Weight-loss depends on nursing assistant time to assist with eating and if necessary assure that supplements are given between meals. Licensed staff are required in sufficient numbers to evaluate situations relating to loss of appetite and declining weight.

To summarize the staffing thresholds across measures, the use of a weighted average is most appropriate. These potential quality problems generally apply to different populations (based on risk) and occur with different frequency. Thus, weighting according to the prevalence of the quality measure, which in turn approximates the risk of the quality problem occurring across facilities, is a logical approach to determining an average staffing level across resident care areas. However, these thresholds only pertain to selected aspects of nursing facility care and particularly for the short-stay sample, where only hospital transfers were investigated, staffing thresholds might be altered by examining other aspects of care (e.g., rehabilitation). Because our samples were separate for the short-stay and the long-stay population, the blend of short-stay and long-stay levels could reflect the blend of short-stay Medicare and long-stay non-Medicare patient days in a facility. Thus, for a given facility the thresholds could be weighted by the proportion of Medicare to total patient days.

The specific staffing thresholds identified using these methods were clearly associated with reductions in the likelihood that facilities fall in the worst decile and hence have markedly increased quality problems. Lower staffing levels up until these thresholds were also associated with incremental quality benefits, so staffing minimums could be set at any levels up until these thresholds. Because the dependent variables were defined as the worst facilities, these staffing levels should not be considered those associated with optimizing care. However, they are consistent with the case studies suggesting that above certain levels the ratio of staff to resident days is not the major determinant of whether quality problems will occur, but rather the way in which staff are allocated, trained, and supervised in the facility.

The case mix analyses suggested that regardless of the facilities risk score, these staffing levels to prevent inclusion in the worst 10 percent of facilities were relatively similar. Fewer

facilities in the lowest risk category were in the worst 10 percent of facilities with respect to quality, but it appeared that quality improvements occurred until about the same thresholds in each case mix category. However, with more facilities in the worst 10 percent of facilities in the higher case mix categories, it makes sense for staffing requirements to be higher in higher risk facilities.

Total wages in dollars per resident day were associated with quality (i.e. not being in the worst decile) across the full range of wages for five of the quality measures, suggesting that quality keeps improving in these areas as staffing expenditures increase. While not as strong as the relationships with staffing levels by type of staff, which held for more quality measures, this relationship with total wages suggests that minimum staffing expenditures may be another way to assure that facilities are adequately staffed. If this strategy is used, it would be wise to also include a minimum requirement for licensed staff, which seems to be necessary to help prevent poor care in quality areas not associated with total wages (e.g. respiratory infections, weight loss). With no clear threshold for total wages, the tradeoff between cost of additional staff and impact on quality would need to be assessed to determine where to set a minimum.

Using California data, a strong relationship was found between nursing assistant retention and whether facilities were in the worst decile for five quality measures across almost the entire range of staff retention. Relationships between nursing assistant turnover and quality measures were weaker (only two quality measures showed a relationship) and there were thresholds above which turnover was no longer associated with quality. These findings persisted at different nursing assistant staffing levels. The retention results in particular demonstrate the importance of other staffing factors besides staffing levels in quality of nursing home care.

2.5 Conclusion

Using data from a representative sample of 10 states including over 5,379 facilities, our objective was to identify staffing thresholds below which quality of care was compromised and above which there was no further benefit of additional staffing with respect to quality. We were able to identify such thresholds in both a short-stay sample of Medicare SNF admissions and long-stay nursing home residents who were in the facility for at least 90 days. These staffing thresholds demonstrated incremental benefits of staffing, in terms of reducing the likelihood a facility would be in the worst 10 percent of all facilities, until thresholds were reached at which point there were no further benefits with respect to quality when additional staff were available. These staffing thresholds were at 2.4 hrs/resident day for nursing assistants and 1.14 hrs/resident day for licensed staff, including .55 hrs/resident day for RNs, to reduce hospital transfers in the short-stay sample. They were at 2.78 hrs/resident day for nursing assistants and 1.3 hrs/resident day for licensed staff, including .75 hrs/resident day for RNs, to improve outcomes and avoid selected care problems in the long-stay sample.

Staffing levels up until these thresholds all appear to be associated with improvements in quality suggesting that staffing minimums at or below these thresholds would assure a higher level of quality in nursing homes. The staffing minimums selected would need to balance additional costs with the quality benefits because quality improvements continue to occur up until staffing levels that the vast majority of facilities do not achieve. Total wages in dollars per resident day were also associated with quality in some areas, suggesting that minimum total staffing wages would also assure a higher level of nursing home quality in several, but not all, quality domains. Staffing retention was strongly associated with quality in California nursing homes demonstrating the importance of other staffing factors beyond staffing levels in assuring nursing home quality.

References

1. Feuerberg M, Reilly K, Kramer AM, Kane RA. "Review of selected research on nursing home staffing and resident outcomes. Chapter 6 in Appropriateness of Minimum Nurse Staffing Ratios in Nursing Homes." Health Care Financing Administration, 2000.

2. Bliesmer MM, Smayling M, Kane RL, Shannon I. The relationship between nursing staffing levels and nursing home outcomes. J Aging Health 1998; 10(3):351-371.

3. Cohen JW, Spector WD. The effect of Medicaid reimbursement on quality of care in nursing homes. J Health Econ 1996; 15:23-48.

4. Harrington C, Zimmerman D, Karon S, Robinson J, Bjorkling F. Nursing home staffing and its relationship to deficiencies. (This research was funded by the Health Care Financing Administration and the Agency for Health Policy & Research). 1999.

5. Spector WD, Takada HA. Characteristics of nursing homes that affect resident outcomes. J Aging Health 1991; 3(4):427-454.

6. Kramer AM, Eilertsen T, Lin M, Hutt E. "Effects of nurse staffing on hospital transfer quality measures for new admissions. Chapter 9 in Appropriateness of Minimum Nurse Staffing Ratios in Nursing Homes." Health Care Financing Administration, 2000.

7. Hutt E, Lin M, Kramer AM. "Effects of nurse staffing on selected quality measures for long term residents derived from MDS." Chapter 10 in Appropriateness of Minimum Nurse Staffing Ratios in Nursing Homes. Health Care Financing Administration, 2000.

8. Martau J, Lin M, Kramer AM. "Effects of Nurse Staffing on Nursing Home Quality Measures." Chapter 11 in Appropriateness in Minimum Nurse Staffing Ratios in Nursing Homes. Health Care Financing Administration, 2000.

9. Kramer AM, Eilertsen T, Lin M, Martau J, Hutt E. "Synthesis of findings on effects of staffing on quality of care. Chapter 12 in Appropriateness of Minimum Nurse Staffing Ratios in Nursing Homes. Health Care Financing Administration, 2000.

10. White et al. "Data sources of nursing home nurse staffing: reliability and validity analysis." Appropriateness of Minimum Nursing Staffing Ratios for Nursing Homes. Health Care Financing Administration, 2000.

11. Wade DT, Collin C. The Barthel ADL index: a standard measure of physical disability? Int Disabil Stud 1988; 10(2):64-67.

12. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965; 14:61-65.

13. Smith DM. Pressure ulcers in the nursing home. Ann Intern Med 1995; 123(6):433-442.

14. McKenna M, Moyers J, Feuerberg M. Review of non-regulatory quality improvement interventions. Report to Congress on HCFA survey. Health Care Financing Administration, 1998.

15. Kayser-Jones J, Schell E. The effect of staffing on the quality of care at mealtime. Nurs Outlook 1997; 45(2):64-72.

16. Kayser-Jones J. Inadequate staffing at mealtime: implications for nursing and health policy. J Gerontol Nurs 1997; 23(8):14-21.

17. Bowers B. "The relationship between staffing and quality in long term care facilities." HCFA study on Appropriateness of Minimum Nurse Staffing Ratios - Interim Report. Health Care Financing Administration, 1996.

18. Kayser-Jones J, Schell ES, Porter C, Barbaccia JC, Shaw H. Factors contributing to dehydration in nursing homes: inadequate staffing and lack of professional supervision. J Am Geriatr Soc 1999; 47(10):1187-1194.

19. Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V et al. MDS cognitive performance scale. J Gerontol A Bio Sci Med Sci 1994; 49(4):M174-M182.

20. Hartmaier SL, Sloane PD, Guess HA, Koch GG, Mitchel CM, Phillips CD. Validation of the minimum data set cognitive performance scale: agreement with the mini-mental state examination. J Gerontol A Bio Sci Med Sci 1995; 50A(2):M128-M133.

21. Saliba D, Kington R, Buchanan J, Bell R, Wang M, Lee M et al. Appropriateness of the decision to transfer nursing facility residents to the hospital. J Am Geriatr Soc 2000; 48(2):154-163.

22. Ouslander JG, Weinberg AD, Phillips V. Inappropriate hospitalization of nursing facility residents: a symptom of a sick system of care for frail older people. J Am Geriatr Soc 2000; 48(2):230-231.

23. Kayser-Jones JS, Wiener CL, Barbaccia JC. Factors contributing to the hospitalization of nursing home residents. Gerontologist 1989; 29(4):502-510.