Published 2018. This article is a U.S. Government work and is in the public domain in the USA
DOI: 10.1111/1475-6773.12975
RESEARCH ARTICLE

The Relationship between Costs and Quality in Veterans Health Administration Community Living Centers: An Analysis Using Longitudinal Data

James F. Burgess Jr. , *Michael Shwartz*, *Kelly Stolzmann*, and *Jennifer L. Sullivan*

Objective. To examine the relationship between cost and quality in Veterans Health Administration (VA) nursing homes (called Community Living Centers, CLCs) using longitudinal data.

Data Sources/Study Setting. One hundred and thirty CLCs over 13 quarters (from FY2009 to FY2012) were studied. Costs, resident days, and resident severity (RUGs score) were obtained from the VA Managerial Cost Accounting System. Clinical quality measures were obtained from the Minimum Data Set, and resident-centered care (RCC) was measured using the Artifacts of Culture Change Tool.

Study Design. We used a generalized estimating equation model with facilities included as fixed effects to examine the relationship between total cost and quality after controlling for resident days and severity. The model included linear and squared terms for all independent variables and interactions with resident days.

Principal Findings. With the exception of RCC, all other variables had a statistically significant relationship with total costs. For most poorer performing smaller facilities (lower size quartile), improvements in quality were associated with higher costs. For most larger facilities, improvements in quality were associated with lower costs.

Conclusions. The relationship between cost and quality depends on facility size and current level of performance.

Key Words. Cost-quality tradeoff, composite quality measures, nursing home costs, nursing home outcomes

Increasing health care expenditures or costs across most of our health care delivery system and concerns about health care quality have given new importance to understanding the relationship between costs and quality. However,

the nature of the relationship differs by the context and structure of the different system components. In this study, we investigate the nursing home care environment. In this context, as in many others, different conceptual models lead to different hypotheses about the nature of the relationship. As pointed out by Weech-Maldonado, Shea, and Mor (2006), a model that emphasizes structural measures, for example, increased staffing levels and increased use of medical technology, hypothesizes that higher costs are necessary to achieve higher quality (Newhouse 1970); a model grounded in the quality management philosophy hypothesizes that well-designed processes with minimal waste lead to both lower costs and better quality (conceptualized as fewer “defects”) (Berwick and Finkelstein 2010). These different models have important policy implications and will apply differently in particular treatment environments. If higher spending is necessary to achieve higher quality, resources might justifiably be directed at those providers with the poorest quality. If both higher quality and lower costs follow from redesigned processes and other types of quality improvement, resources should be used to incentivize and support improvement activities.

Following a systematic review of the association between health care quality and cost across different types of providers, Hussey, Wertheimer, and Mehrotra (2013) concluded “Evidence of the direction of association between health care cost and quality is inconsistent. Most studies have found that the association between cost and quality is small to moderate, regardless of whether the direction is positive [higher costs are associated with higher quality] or negative [higher quality is associated with lower costs]” (p. 27).

In this study, we consider the relationship between cost and quality in the context of Veterans Health Administration (VA) nursing homes. As summarized in what follows, Hussey, Wertheimer, and Mehrotra’s (2013) findings about the relationship in nursing homes mirror their overall findings. Anderson, Hsieh, and Su (1998), based on cost and quality data from 494 nursing homes in Texas, found that nursing homes with the best outcomes (measured by a composite of 11 resident outcomes) had 7 percent higher costs per

Address correspondence to Jennifer L. Sullivan, Ph.D., Center for Healthcare Organization and Implementation Research (CHOIR), VA Boston Healthcare System, 150 S. Huntington Ave. (152M), Boston, MA 02130; e-mail: Jennifer.Sullivan@va.gov. James F. Burgess, Jr., Ph.D. (deceased), Michael Shwartz, Ph.D., and Kelly Stolzmann, M.S., are with the Center for Healthcare Organization and Implementation Research (CHOIR), VA Boston Healthcare System, Boston, MA. James F. Burgess, Jr., Ph.D., and Jennifer L. Sullivan, Ph.D., are with the Boston University School of Public Health, Boston, MA. Michael Shwartz, Ph.D., is with the Boston University Qualstrom School of Business, Boston, MA.

resident day than nursing homes with the lowest cost per day. Mukamel and Spector (2000) looked at the relationship between costs and quality in over 500 free-standing private and public nursing homes in New York State using cost and quality data from 1991 to 1992. Quality was measured using three risk-adjusted outcomes: decline in functional status, worsening pressure ulcers, and mortality. In general, they found an inverted U-shaped relationship between costs and quality: at lower levels of quality, quality improvements added to costs; at higher levels of quality, quality improvements reduced costs. Weech-Maldonado, Neff, and Mor (2003) analyzed 1996 cost and quality data from 781 nursing homes in five states. They found that higher outcome quality (lower values of an index comprised of pressure ulcer prevalence, worsening pressure ulcers, cognitive decline, and mood decline) was associated with lower costs. However, there was no direct relationship between a process quality index (created from prevalence of physical restraint and catheter use) and costs. Hicks et al. (2004), using 1999 data on all certified nursing homes in Missouri, examined four risk-adjusted quality indicators: decline in ADLs (activities of daily living), development of pressure ulcers, weight loss, and psychotropic drug use. For ADL decline and weight loss, the authors also found a U-shaped curve, but it was not inverted: at lower levels of quality, quality improvements reduced costs; at higher levels, quality improvements added to costs. Weech-Maldonado, Shea, and Mor (2006) analyzed data from the mid-1990s on 749 nursing homes in five states and considered two outcome measures: pressure ulcer worsening and mood decline. As seems appropriate given the conflicting findings of Mukamel and Spector (2000) and Hicks et al. (2004), Weech-Maldonado, Shea, and Mor (2006) found an inverted U-shaped relationship for pressure ulcers, but the opposite pattern for mood decline, although in both cases, the shape of the curve at lower quality levels was much flatter (particularly for mood decline) than suggested by the earlier literature. It is noteworthy that these studies use data that are now quite old and that results are based upon cross-sectional analyses.

In our study, we used 13 quarters of data from fiscal year (FY) 2009 through FY 2012 to examine the relationship between costs and quality in 130 Veterans Health Administration (VA) nursing homes, called community living centers (CLCs) by the VA. The VA is the largest integrated health care system in the United States and has a state-of-the-art electronic health record system (Spetz, Burgess, and Phibbs 2012) and a bottom-up activity-based costing system (Carey and Burgess 2000). Although in a government-run system, there are not the types of incentives to reduce cost that exist in non-VA nursing homes, budgetary pressures to reduce annual costs are still imposed by the

capitated resource allocation system (Yaisawarng and Burgess 2006). More important, there is a strong culture, driven by the mission of the VA to honor its debt to veterans, to provide high-quality cost-effective care (Kizer and Dudley 2009). Our study includes both traditional medically oriented measures of nursing home quality (the number of different types of unfavorable events) and more broad-based measures that focus on resident quality of life and resident-centered care (RCC). In addition, we used a longitudinal model that estimates overall results about the cost–quality relationship from the relationship between cost and quality within facilities over time. This approach reduces the impact of confounding factors that plague cross-sectional analyses.

METHODS

In what follows, we discuss how we operationalize four general concepts important when considering the quality/cost relationship in a nursing home environment and then describe the model used to link cost and quality. First, quality has a clinically oriented multidimensional aspect that is assessed through an accepted and validated periodic measurement tool. However, perhaps the most essential aspect of the nursing home environment over other health care settings is that the patient also is a resident. We use an established tool to measure aspects of resident-centered care. Third, as we are interested in costs, we use a validated case mix classification system based on resource needs. Finally, we measure costs employing the strengths of the VA detailed bottom-up costing methodology.

Clinically Oriented Quality Measures

As part of efforts to ensure that nursing home residents receive high-quality care, the federal government requires Medicare reimbursed nursing homes to undertake assessments of long-stay residents (those staying over 90 days) using the Resident Assessment Instrument (RAI) and to develop care plans based on these assessments. Full assessments are carried out yearly or when there is a significant change in resident status; partial assessments are carried out quarterly. Data from the RAI are referred to as the Minimum Data Set (MDS). A set of 24 quality indicators (QIs), which measures unfavorable events (e.g., falls) or patient states (e.g., depression), has been developed from the MDS data to measure the quality of nursing home care (Zimmerman 2003). Four of the QIs are stratified into high- and low-risk categories,

Table 1: Descriptive Statistics on Individual Variables and on the Variables Averaged over Quarter within Facility

Variable	N*	Mean (SD)	Min	5th %	25th %	Median	75th %	95th %	Max
Individual values of the variables									
Total cost [†]	1,672	773 (747)	0.1	64	251	557	1,021	2,242	4,672
Days [‡]	1,672	179 (190)	0.1	12	55	119	228	621	1,126
Severity	1,671	13.1 (3.8)	0.3	6.9	10.7	13.1	15.4	19.2	31.0
RCC	1,635	235 (37)	115	172	211	236	258	294	345
QS	1,690	0.14 (0.02)	0.07	0.10	0.12	0.13	0.15	0.18	0.27
Variables averaged across quarters within facility									
Total cost [†]	130	765 (679)	27	98	299	582	990	2,170	4,160
Days [‡]	130	177 (178)	36	19	66	127	209	584	1,068
Severity	130	13.1 (2.4)	4.9	9.2	11.7	12.9	15.0	16.8	18.7
RCC	130	234 (31)	137	187	215	232	253	287	312
QS	130	0.14 (0.02)	0.08	0.10	0.12	0.13	0.15	0.17	0.21

*For certain facilities in certain quarters, cost, days, and severity information were missing. Artifacts data were missing for two quarters.

[†]Measured in 000s.

[‡]Measured in 0s.

resulting in a total of 28 QIs (see Supplemental Material, Table 1). These indicators are provided to each non-VA nursing home and are used by regulators as part of the certification process (Castle and Ferguson 2010). The VA also monitors these 28 indicators and sends monthly reports to each facility. It is important to note that, as the MDS QIs measure unfavorable events, lower rates of the QIs indicate higher quality.

We created a composite measure (Institute of Medicine, 2006) from the 28 QIs. However, similar to Sullivan et al. (2013), we first adjusted for differences in the reliability of estimates of QI rates from facilities of varying sizes using a Bayesian hierarchical model (i.e., a multivariate normal-binomial model) to calculate “shrunkened” rates of each QI at each facility for each month. Schwartz et al. (2013) have shown that MDS QI data are consistent with this model and that the “shrunkened” rates calculated from the model are better predictors of the next year’s observed rates than the current year’s observed rates.

We combined the 28 “shrunkened” QI rates at each facility for each month into a composite measure using facility-specific opportunity-based weights, the approach used by the Centers for Medicare & Medicaid Services (CMS) in its initial pay-for-performance program (Kahn et al. 2006). This approach treats the impact of the adverse event measured by each QI as equally important (Schwartz, Restuccia, and Rosen 2015). The resulting composite can be thought of as the adjusted probability that an average resident experiences a

QI event. Because the MDS indicators traditionally are used as measures of nursing home quality, we refer to the MDS-based composite measure as the “quality score” (QS). We calculated a quarterly QS as the average of the Qs in the 3 months comprising each quarter.

Resident-Centered Care

The Omnibus Budget Reconciliation Act in 1987 created incentives for nursing homes to provide care more consistent with a medical model, one that emphasized standardization, safety, and clinical quality (White-Chu et al. 2009). However, beginning in the mid-90s, a number of nursing homes began to implement various culture change models that emphasized resident-centered care (RCC). These models involved shifting the care paradigm to promote resident quality of life and resident self-direction, in particular, more individualized care, transformed physical environments, and changed staff roles (Ronch 2003; Rahman and Schnelle 2008).

The VA implemented RCC in 2005 and, as part of this initiative, renamed its nursing homes to Community Living Centers (CLCs). The Artifacts of Culture Change Tool, which was developed by CMS in 2001 (Bowman and Schoeneman 2006), is used by the VA to measure changes over time in the extent of RCC implementation in CLCs. It is completed quarterly by each facility. Beginning with the last assessments in FY 2010, VA began to use an automated version of the Artifacts Tools to be consistent with CMS (see Sullivan et al. 2013). However, the VA only selected artifacts which were applicable to the VA. For the purposes of this study, we re-scored the Artifacts responses from the end of FY10 to FY12 to be consistent with the scoring approach used by VA in the prior years.

The tool has six domain scores: care practices (45 points), workplace processes (55 points), environment (180 points), leadership (25 points), family and community (30 points), and outcomes such as longevity and turnover (20 points). The domain scores are summed together to create a total Artifacts score (up to 355 points). Data from the 4th quarter of FY2009 and the 2nd and 4th quarter of FY2010 were unavailable, most likely due to internal changes in staffing at VA central office during those periods. We imputed the total Artifacts Total Score for the three missing quarters as the average of the quarter that preceded and followed the missing quarter. We refer to the Artifacts score as the “RCC score.” It is worth noting that the RCC score is a formal performance measure in the VA, and CLC directors’ incentive payments are influenced by the extent of increase in the score from year to year.

Resident Severity: Resource Utilization Groups

Resource Utilization Groups (RUGs) are a case mix classification system for nursing home residents. Using MDS data, residents are placed into one of a number of different categories based on their expected resource use (Fries et al. 1994). Over the years since RUGs were introduced, a number of different versions have been developed. At the time of our study, the VA was using the 53-RUG-III system (the maximum score assigned to any of the 39 categories is 53) that went into effect in January 2006 for Medicare payment to skilled nursing homes.

Measuring Cost

The VA uses a bottom-up Management Cost Accounting System for resource allocation and management (Carey and Burgess 2000). This system allocates all fixed and variable costs within a department to intermediate products, which are then mapped into intermediate product departments (laboratory, nursing, pharmacy, radiology, surgery, and other) and aggregated to determine the costs of patient encounters. On average, about half of the costs are nursing, which are allocated using bed days of care as the intermediate product. These costs vary according to the pay of nurses. The Nurse Locality Pay System established under the VA Nurse Pay Act of 1990 sets RN wages based on wages in non-VA facilities in the same labor market (Staiger, Spetz, and Phibbs 2010). Therefore, we adjusted nursing costs using the RN average salary index to take account of these differences in input prices. We refer to the adjusted total costs as “total costs.”

Analysis

The data consist of the following variables for each of the 130 facilities for 13 quarters: total costs, patient days, quality score (QS), RCC score, and severity score. To examine the relationship between cost and quality, we ran a generalized estimating equation (GEE) model with an autoregressive (1) correlation structure, a log link, and a gamma distribution (Zeger, Liang, and Albert 1988). Total costs for each quarter at each facility were the dependent variables, and the quality scores for each quarter at each facility were the independent variables of primary interest. Covariates were quarterly patient days and severity score. Facility was treated as a fixed effect.

To evaluate the possibility of a nonlinear relationship between the independent variables and total cost and that the impact of the variables might

depend on facility size, we ran the model with both linear and squared terms for the independent variables and an interaction between patient days and both the linear and squared terms. The modified Park test suggested the Poisson distribution might be a reasonable alternative to the gamma distribution. However, because the GEE model with a Poisson distribution did not converge, we used the gamma distribution.

All analyses were performed using *SAS*, version 9.4 (SAS Institute, Cary, NC, USA). This study was approved by VA's Central Institutional Review Board.

RESULTS

Table 1 shows descriptive statistics for total costs, resident days, severity score, RCC score, and QS. The top part of the table shows variation of individual values of the variables. This variation reflects both variation across facilities and variation within facilities over time. The lowest QS in the database was 0.07 and the highest 0.27. The bottom part of the table shows variation of facility-level averages, where the averaging was performed over quarters within facilities. The highest performing facility in terms of quality had a QS of 0.08 (since the QS measures unfavorable events, lower scores are better) and the lowest performer a score of 0.21.

Table 2 shows, for each of the variables, trends across fiscal year, holding quarter constant, and trends across quarter, holding fiscal year constant. There were striking declines in costs, days, and severity over the course of a fiscal year. This trend was not consistent across all facilities, but it occurred in a large enough number to drive the overall trends. The reasons for this are not entirely clear, but they probably relate in part to the way the accounting system measures cost differences across services, which in practice does vary somewhat across sites. Also, in many facilities, budget allocations are made at the start of the fiscal year. As sicker long-stay residents either die or are moved into private sector nursing homes (most VA medical centers, in addition to having their own CLC, contract with the private sector for additional nursing home beds), beds are increasingly filled by short-stay residents (for whom MDS assessments are not performed and who are not included in our analysis) in order for the facility to maintain budget flexibility. The decline in severity over the course of a year is consistent with this hypothesis. Holding quarter constant, costs increased slightly over the fiscal years, with the exception of the third quarter where there was a slight drop, and days decreased

Table 2: Variables Over Quarter and Over Fiscal Year After Averaging Across Facilities

	Total Cost (000s)				Resident Days (0s)			
	FY09	FY10	FY11	FY12	FY09	FY10	FY11	FY12
Q1	–	998	1,021	1,028	–	239	233	226
Q2	–	864	881	897	–	207	201	193
Q3	733	720	716	715	181	174	164	154
Q4	494	474	505	–	123	118	117	–

	RCC Score				Quality Score			
	FY09	FY10	FY11	FY12	FY09	FY10	FY11	FY12
Q1	–	226	219	246	–	0.135	0.134	0.136
Q2	–	234	228	252	–	0.135	0.135	0.138
Q3	217	243	235	255	0.135	1.134	0.134	0.138
Q4	222	231	242	–	0.133	0.132	0.136	–

Severity Score				
Q1	–	14.5	13.7	14.0
Q2	–	14.3	13.4	13.4
Q3	13.4	14.2	12.5	12.7
Q4	11.4	11.8	11.2	–

consistently across the fiscal years. In general, severity decreased over the fiscal years, which may reflect the increasing use of contract beds for the sicker long-stay residents. As noted, a new automated system to collect Artifacts data was implemented in the last quarter of FY10 and the lower RCC scores in that period and the first three quarters of FY2011 probably reflect a time in which facilities were becoming familiar with the new scoring system. Aside from that period, there was a consistent increase in RCC scores both over quarters within years and over years within quarter. As noted, as RCC score is a formal incentive measure, increases in this score were not surprising. There were no trends either across quarter or fiscal year in the QS.

Table 3, upper right quadrant, shows the correlations among the variables. All of the variables had a statistically significant correlation with costs ($p < .05$). Not surprisingly, there was a very high correlation between costs and resident days ($r = 0.96$). As expected, severity was positively correlated with costs ($r = 0.16$), as was the RCC score, although the correlation was low ($r = 0.06$). The QS was negatively correlated with cost ($r = -0.19$), indicating that improved clinical quality (i.e., a lower QS, as the score measures unfavorable events) was associated with higher costs. Larger facilities, that is, those with more resident days, tended to have a more severe case mix ($r = 0.13$) and

Table 3: Correlations of the Individual Variables and the Variables Averaged over Quarter within Facility[†]

<i>Variables</i>	<i>Total Cost</i>	<i>Days</i>	<i>Severity</i>	<i>RCC</i>	<i>QS</i>
Total cost	1.00	0.96***	0.16***	0.06*	−0.19***
Days	0.97***	1.00	0.13***	0.04	−0.24***
Severity	0.15	0.11	1.00	−0.00	−0.04
RCC	0.10	0.09	0.03	1.00	−0.08**
QS	−0.30**	−0.34***	−0.11	−0.13	1.00

[†]Upper right triangle is correlation among the individual variables; lower left is correlation among the variables averaged over quarter within facility.

* $p < 0.05$; ** $p < .01$; *** $p < .001$.

Table 4: Model Parameter Estimates*

<i>Variable</i>	<i>Coefficient (95% Confidence Interval)</i>	<i>p value</i>	<i>Exp (Coefficient)</i>
Days	0.0026 (0.0020, 0.0033)	<.0001	1.00
Days ²	−0.00000006 (−0.00000009, −0.00000003)	<.0001	1.00
Severity	0.2432 (0.2069, 0.2795)	<.0001	1.27
Severity ²	−0.0070 (−0.0083, −0.0058)	<.0001	0.99
Severity*Days	−0.00015 (−0.00019, −0.00011)	<.0001	1.00
Severity ² *Days	0.0000048 (0.0000035, 0.00000617)	<.0001	1.00
QS	19.02 (10.21, 27.84)	<.0001	1.83E+08
QS ²	−78.39 (−106.46, −50.32)	<.0001	9.02E-35
QS*Days	−0.0128 (−0.0198, −0.0058)	.0003	0.99
QS ² *Days	0.0602 (0.0325, 0.0880)	<.0001	1.06
Constant	9.0 (8.3, 9.7)	<.0001	8,038

*Sample size = 1,671.

higher clinical quality ($r = -0.24$). And higher levels of RCC care were associated with higher quality ($r = -0.08$). The lower left quadrant of the table shows correlations among facility-level averages. The negative correlation between QS and cost (−0.30) and between QS and days (−0.34) was stronger than when individual values of the variables were considered. At the facility level, the correlation between RCC and the QS was not statistically significant.

When we initially ran the GEE model, none of the terms that involved RCC were statistically significant. We then ran the model a second time including only the main effect for RCC, but it was still not statistically significant. Hence, we dropped RCC from the model. Table 4 shows the parameter estimates of the final GEE model. The statistical significance of the interaction

terms indicates the importance of the size of the facility when examining the relationship between cost and quality. To portray these relationships, we divided facilities into quartiles based on resident days (see Supplemental Material, Table 2 for descriptive statistics on the variables by resident day quartile). In Figure 1(a–d), we show the relationship between cost and quality for each quartile. For these runs, we used the median resident days and severity score for facilities in the quartile. The range of the quality score in the figures is from the minimum score to the maximum score of facilities in the quartile. To the extent median values are representative of the facilities in the quartile, one can conclude from these figures that for most of the smaller 25 percent of the facilities, improvements in quality were associated with higher costs; however, for most of the other 75 percent of the facilities, improvements in quality were associated with lower costs (except for the largest facilities with the very best quality scores). As the Y-axis is different in each of the figures, it is difficult to get a comparative sense of changes in cost associated with changes in quality across the quartiles. To provide some sense of the magnitude of cost changes associated with quality changes, we fit a simple linear regression model to the linear part of the cost/quality curves (eyeballed) in each quartile: 0.21 to 0.15 in quartile 1, 0.15 to 0.10 in quartile 2, 0.16 to 0.12 in quartile 3, and 0.15 to 0.11 in quartile 4. From the coefficient associated with the quality score, we estimated the change in total cost (as a percentage of

Figure 1: Relationship between Total Cost (in 000s) and Quality Score* by Patient Day (PD) Quartile (Q) of the Facility

*Note. A lower quality score indicates higher quality.

median cost for facilities in the quartile) for a 0.01 improvement in quality score. In quartile 1, each 0.01 improvement in quality score was associated with 6.8 percent higher costs; in quartiles 2 to 4, each 0.01 improvement in quality score was associated with a 1.1 percent, 3.5 percent and 10.0 percent reduction in total costs, respectively. In quartile 4, improvements in quality score from 0.09 to 0.07 were associated with a 4.6 percent increase in costs.

DISCUSSION

As noted in the Introduction, there are two models (each admittedly the extreme cases) about the relationship between cost and quality: One emphasizes the structural and major process changes that are necessary to achieve higher quality; and the other that well-designed processes can both lower costs and improve quality. These two models imply different assumptions about the direction of “causality.” The first assumes that more resources are necessary to achieve higher quality; the second tends to emphasize engaging front line staff in developing improved processes that will lead to both higher quality and, by reducing waste, lower costs. (Note the quote by Hussey, Wertheimer, and Mehrotra [2013] at the end of the second paragraph of the Introduction implies this shift in direction of implied causality.) As suggested by the inconsistency in earlier studies examining cost–quality relationships in nursing homes, these two models play out differently in different environments and perhaps for different types of quality indicators. Even in our environment, the relationship is complex and varies based on facility size and level of performance. However, overall, our analyses suggest that in the VA context, for most long-term care facilities, improvements in quality were associated with lower costs. The reductions in costs are not trivial. Using the median facility as representative of the facility size quartile, a 0.01 improvement in quality score was associated with cost reductions ranging from 1.1 percent of median total costs for most facilities in the second quartile to 10.0 percent of median costs for most of the larger facilities in the fourth quartile.

In both the correlation analysis at the facility level and in the multivariate model, the relationship between RCC and cost was not statistically significant, which suggests that concerns about higher costs may not be a major impediment to RCC implementation. Also, at the facility level, there was not a strong statistical relationship between RCC and clinical quality. Sullivan et al. (2018), using the same data set as us, examined the relationship between

RCC score and quality score using a GEE model with facility treated as a fixed effect. In this model, the statistical relationship between RCC and clinical quality also was not strong ($p = .22$).

One of the major challenges in examining quality in long-term care facilities is the difficulty of “adequately” adjusting for differences in patient risk. As a result, some of what is attributed, for example, to an improvement in quality may be the result of a less sick case mix. The RUGs score provides some adjustment, but to a large extent, RUGs measures level of treatment provided. The problem in adjusting for case mix in long-term care facilities is that the variables one would like to use for adjustment, for example various measures of frailty, are impacted by quality. As Mor (2005, p. 344): “. . . many nursing home measures are based on prevalence, because their residents are served for extended periods. Therefore, it is hard to identify a ‘baseline’ status for nursing home patients, which has not already been influenced by the quality of the nursing home. For example, being bedridden is predictive of acquiring a pressure ulcer. However, patients may become bedridden because of the inadequate mobility care earlier. Statistically controlling for this ‘effect’ could adjust away earlier poor care.” In MDS 2.0 (which we use in this paper), some risk adjustment is performed by exclusions from the denominator and, for 4 of the 24 indicators, by stratification into a low- and high-risk groups. Even in MDS 3.0, risk adjustment using resident-level covariates is only performed for 3 of 15 QIs used for long-stay residents. Despite these concerns, the QIs for which CLC nursing home managers are held responsible (which are the same ones used in the private sector at the time of the study) are the ones we used in this paper. It is the incentives created by these measures that motivate improvements.

As noted, the VA context, in which CLCs receive fixed budget allocations and are a part of an integrated delivery system, is different than that in which most private sector nursing homes operate. Hence, the extent to which our findings can be generalized is unclear. This is a limitation of our study. However, both within the VA and private sector, there are strong pressures to deliver high-quality care as efficiently as possible across all providers within the delivery system.

A strength of the VA context is that all facilities use a common cost accounting system and there are system-wide protocols for collection and dissemination of severity, resident-centered care and MDS-based quality data. Also, our use of a composite measure derived from the individual QIs has clear advantages. A strong conceptual basis for hypothesizing quality indicator-specific relationships between cost and quality is rare in the

literature, even as quality indicators have proliferated. Although post hoc, one can justify relationships seen for different QIs, a priori it is hard to make the case that the relationship should differ for specific QIs. By focusing on the relationship between cost and any QI event, we reduce the small sample problem (some indicators show one relationship and others show another) arising when examining individual indicators and focus on the average collective relationship across all indicators. Also, the Bayesian model adjusts for differences in reliability of the individual QIs resulting from differences in the size of facilities, and thus provides an improved basis for comparing rates across facilities of different sizes. Finally, our use of a fixed effects model, which derives the overall relationship between cost and quality from the individual relationships between cost and quality over time within facility, significantly reduces the impact of unobserved confounders that differ across facilities. Thus, the validity of conclusions drawn from this type of model are stronger than those drawn from cross-sectional analyses, the approach used in other studies that have examined the relationship between costs and quality in long-term care facilities (and, for that matter, in most other types of facilities). However, changes within facilities over time still can confound results.

Finally, we are strong believers in employing mixed methods and qualitative investigations to sharpen our understanding of organizational characteristics and functioning that lead to variations in performance. We are finalizing such a study in 12 CLCs. Preliminary results suggest that high- and low-performing facilities differ on teamwork and communication, leadership support, and organizational culture (Sullivan et al. 2016). These types of studies have the potential to identify changes and interventions for improving performance in the lower-performing facilities.

CONCLUSIONS

There was not a strong statistical relationship between cost and RCC in VA CLCs. However, the statistical relationship between cost and clinical quality was much stronger and in general indicated that quality improvements are associated with lower costs. Efforts to improve clinical quality by the approximately 25 percent of smaller poor performing CLCs are likely to require more resources. However, for larger facilities, quality improvements were associated with lower costs.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: We would like to thank the project staff who assisted on the study, including Dr. Erol Peköz, Emily Rose Lopez, Katelyn Gormley, Melissa Afable, and Omonyélé Adjognon. We would also like to thank the reviewers whose comments and suggestions resulted in significant improvements in the paper. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the U.S. Department of Veterans Affairs.

Funding: This work was supported by VA HSR&D Grant IIR-11-356 (Sullivan, PI).

Previous dissemination: This work was presented at the 2016 Annual AcademyHealth meeting poster session (Boston, MA).

Disclosures: None.

Disclaimer: None.

REFERENCES

- Anderson, R. A., P. C. Hsieh, and H. F. Su. 1998. "Resource Allocation and Resident Outcomes in Nursing Homes: Comparisons Between the Best and Worst." *Research in Nursing and Health* 21 (4): 297–313.
- Berwick, D. M., and J. A. Finkelstein. 2010. "Preparing Medical Students for the Continual Improvement of Health and Health Care: Abraham Flexner and the New 'Public Interest'." *Academic Medicine* 85 (9 Suppl): S56–65.
- Bowman, C., and K. Schoeneman. 2006. *Development of the Artifacts of Culture Change Tool*. No. HHSM-500-2005-00076P. Baltimore, MD: Centers for Medicare & Medicaid Services.
- Carey, K., and J. F. Burgess Jr. 2000. "Hospital Costing: Experience from the VHA." *Financial Accountability and Management* 16 (4): 289–308.
- Castle, N. G., and J. C. Ferguson. 2010. "What is Nursing Home Quality and How Is It Measured?" *Gerontologist* 50 (4): 426–42.
- Fries, B. E., D. P. Schneider, W. J. Foley, M. Gavazzi, R. Burke, and E. Cornelius. 1994. "Refining a Case-Mix Measure for Nursing Homes: Resource Utilization Groups (RUG-III)." *Medical Care* 32 (7): 668–85.
- Hicks, L. L., M. J. Rantz, G. G. Petroski, and D. B. Mukamel. 2004. "Nursing Home Costs and Quality of Care Outcomes." *Nursing Economics* 22 (4): 178–92.
- Hussey, P. S., S. Wertheimer, and A. Mehrotra. 2013. "The Association between Health Care Quality and Cost: A Systematic Review." *Annals of Internal Medicine* 158 (1): 27–34.
- Institute of Medicine. 2006. *Performance Measurement: Accelerating Improvement*. Washington, DC: The National Academies Press.

- Kahn III, C. N., T. Ault, H. Isenstein, L. Potetz, and S. V. Gelder. 2006. "Snapshot of Hospital Quality Reporting and Pay-for-Performance under Medicare." *Health Affair* 25 (1): 148–62.
- Kizer, K. W., and R. A. Dudley. 2009. "Extreme Makeover: Transformation of the Veterans Health Care System." *Annual Review of Public Health* 30: 313–39.
- Mor, V. 2005. "Improving the Quality of Long-Term Care with Better Information." *Milbank* 83 (3): 333–64.
- Mukamel, D. B., and W. D. Spector. 2000. "Nursing Home Costs and Risk-Adjusted Outcome Measures of Quality." *Medical Care* 38 (1): 78–89.
- Newhouse, J. P. 1970. "Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital." *American Economic Review* 60: 64–74.
- Rahman, A. N., and J. F. Schnelle. 2008. "The Nursing Home Culture-Change Movement: Recent Past, Present, and Future Directions for Research." *Gerontologist* 48 (2): 142–8.
- Ronch, J. L. 2003. "Leading Culture Change in Long-Term Care: A Map for the Road Ahead." In *Culture Change in Long-Term Care*, edited by J. L. Ronch, and A. S. Weiner, pp. 65–80. Binghampton, UK: Haworth Press.
- Shwartz, M., J. D. Restuccia, and A. K. Rosen. 2015. "Composite Measures of Health Care Provider Performance: A Description of Approaches." *Milbank Quarterly* 93 (4): 788–825.
- Shwartz, M., E. A. Pekoz, C. L. Christiansen, J. F. Burgess Jr, and D. Berlowitz. 2013. "Shrinkage Estimators for a Composite Measure of Quality Conceptualized as a Formative Construct." *Health Services Research* 48 (1): 271–89.
- Spetz, J., J. F. Burgess Jr, and C. S. Phibbs. 2012. "What Determines Successful Implementation of Inpatient Information Technology Systems?" *American Journal of Managed Care* 18 (3): 157–62.
- Staiger, D., J. Spetz, and C. S. Phibbs. 2010. "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment." *Journal of Labor Economics* 28 (2): 211–36.
- Sullivan, J. L., M. Shwartz, J. F. Burgess Jr, E. A. Peköz, C. L. Christiansen, M. Gerena-Melia, and D. Berlowitz. 2013. "Person Centered Care Practices and Quality in Department of Veterans Affairs Nursing Homes: Is There a Relationship?" *Medical Care* 51 (2): 165–71.
- Sullivan, J. L., R. L. Engle, M. Afable, K. Gormley, D. Tyler, M. Shwartz, K. Stolzmann, and V. A. Parker. 2016. *Is Variation in Resident Centered Care and Quality Performance Related to Health Systems Factors in Veterans Health Administration Community Living Centers?* Boston, MA. Poster session presented at the annual AcademyHealth Research Meeting.
- Sullivan, J. L., M. Shwartz, K. Stolzmann, M. Afable, and J. F. Burgess, Jr. 2018. "A Longitudinal Assessment of the Effect of Resident-Centered Care on Quality in Veterans Health Administration Community Living Centers." *Health Services Research* 53 (3): 1819–33. <https://doi.org/10.1111/1475-6773.12688>
- Weech-Maldonado, R., G. Neff, and V. Mor. 2003. "Does Quality of Care Lead to Better Financial Performance?: The Case of the Nursing Home Industry." *Health Care Management Review* 28 (3): 201–16.

- Weech-Maldonado, R., D. Shea, and V. Mor. 2006. "The Relationship between Quality of Care and Costs in Nursing Homes." *American Journal of Medical Quality* 21 (1): 40–8.
- White-Chu, E. F., W. J. Graves, S. M. Godfrey, A. Bonner, and P. Sloane. 2009. "Beyond the Medical Model: The Culture Change Revolution in Long-Term Care." *Journal of the American Medical Directors Association* 10 (6): 370–8.
- Yaisawarng, S., and J. F. Burgess Jr. 2006. "Performance-Based Budgeting in the Public Sector: An Illustration from the VA Health Care System." *Health Economics* 15 (3): 295–310.
- Zeger, S. L., K.-Y. Liang, and P. S. Albert. 1988. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." *Biometrics* 44 (4): 1049–60.
- Zimmerman, D. R. 2003. "Improving Nursing Home Quality of Care through Outcomes Data: The MDS Quality Indicators." *International Journal of Geriatric Psychiatry* 18 (3): 250–7.

SUPPORTING INFORMATION

Additional supplemental material may be found online in the Supporting Information section at the end of the article.

Appendix SA1: Author Matrix.

Table S1: Minimum Data Set (MDS) Quality Indicators (QIs).

Table S2: Descriptive Statistics for the Variables Averaged Over Quarter Within Facility Stratified by Quartile of Resident Days.