Policies that Impact Care

Impact of Public Reporting on Quality of Postacute Care

Rachel M. Werner, R. Tamara Konetzka, Elizabeth A. Stuart, Edward C. Norton, Daniel Polsky, and Jeongyoung Park

Objective. Evidence supporting the use of public reporting of quality information to improve health care quality is mixed. While public reporting may improve reported quality, its effect on quality of care more broadly is uncertain. This study tests whether public reporting in the setting of nursing homes resulted in improvement of reported and broader but unreported quality of postacute care.

Data Sources/Study Setting. 1999–2005 nursing home Minimum Data Set and inpatient Medicare claims.

Study Design. We examined changes in postacute care quality in U.S. nursing homes in response to the initiation of public reporting on the Centers for Medicare and Medicaid Services website, Nursing Home Compare. We used small nursing homes that were not subject to public reporting as a contemporaneous control and also controlled for patient selection into nursing homes. Postacute care quality was measured using three publicly reported clinical quality measures and 30-day potentially preventable rehospitalization rates, an unreported measure of quality.

Principal Findings. Reported quality of postacute care improved after the initiation of public reporting for two of the three reported quality measures used in Nursing Home Compare. However, rates of potentially preventable rehospitalization did not significantly improve and, in some cases, worsened.

Conclusions. Public reporting of nursing home quality was associated with an improvement in most postacute care performance measures but not in the broader measure of rehospitalization.

Key Words. Quality of care, postacute care, nursing home quality, public reporting

Recent evidence about the poor quality of health care delivered in the United States (Kohn, Corrigan, and Donaldson 1999; Institute of Medicine 2001; McGlynn et al. 2003) has caused an outcry among health care consumers, providers, and policymakers. In an effort to improve quality of care, policymakers have turned to market-based reforms across the health care system. For example, the Centers for Medicare and Medicaid Services (CMS) have begun reporting health

care quality through their websites, such as Nursing Home Compare, which publicly rates the performance of nursing homes on certain quality measures.

Public reporting is designed to improve health care quality in two ways. First, public reporting may motivate improvements in the quality of individual providers, increasing provider-specific quality of care. Second, public reporting may increase the likelihood that patients select high-quality providers, thus increasing the number of patients receiving high-quality care. If either of these effects is realized, quality of care will improve on average.

Because public reporting has tremendous face validity, it has been widely adopted in many health care settings (Fung et al. 2008). Yet there is mixed evidence on whether these efforts truly improve quality of care and, if quality does improve, the mechanism by which this occurs—provider-driven quality improvement efforts or consumer use of the information to choose high-quality providers (Werner and Asch 2005; Fung et al. 2008). In addition, there remains concern that observed improvements in quality under public reporting may not represent "true" quality improvement but, rather, are the result of providers selecting healthier patients under public reporting so that they appear to have improved quality outcomes (Werner and Asch 2005).

Our objective in this paper is to test whether public reporting in the setting of postacute care in skilled nursing facilities (SNFs) stimulates provider-driven quality improvement while controlling for changes in market share and changes in patients' health risk.

BACKGROUND AND PRIOR LITERATURE

Poor quality of care has been pervasive in nursing homes for decades (Institute of Medicine 1986). Major regulatory policies have been implemented to im-

Address correspondence to Rachel M. Werner, M.D., Ph.D., 1230 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104; e-mail: rwerner@mail.med.upenn.edu. Rachel Werner is with the Center for Health Equity Research and Promotion, Philadelphia VAMC, the Division of General Internal Medicine, University of Pennsylvania School of Medicine, and the Leonard Davis Institute of Health Economics, University of Pennsylvania. R. Tamara Konetzka, Ph.D., is with the Department of Health Studies, University of Chicago, Chicago, IL. Elizabeth A. Stuart, Ph.D., is with the Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD. Edward C. Norton, Ph.D., is with the Department of Health Management and Policy and Department of Economics, University of Michigan, Ann Arbor, MI. Daniel Polsky, Ph.D., is with the Division of General Internal Medicine, University of Pennsylvania School of Medicine, and the Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA. Jeongyoung Park is with the Division of General Internal Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA.

prove nursing home quality, including the 1987 Nursing Home Reform Act or the Omnibus Budget Reconciliation Act (OBRA), which mandated that each Medicare- and Medicaid-certified nursing home be regularly inspected and submit regular comprehensive assessments of each resident. While OBRA led to some quality improvements (Kane et al. 1993; Shorr, Fought, and Ray 1994; Castle, Fogel, and Mor 1996; Fries et al. 1997; Mor et al. 1997; Snowden and Roy-Byrne 1998), significant problems with quality of care remained (Wunderlich and Kohler 2000).

More recently, with regulation failing to fully reform nursing home quality, the persistent problems of inadequate quality have been attributed to the lack of information about quality with which to stimulate consumer choice of care and provider competition for high-quality care (Mor 2005). Accordingly, quality improvement efforts have increasingly turned to market-based incentives such as public reporting of quality information. In 2001, the Department of Health and Human Services announced the formation of the Nursing Home Quality Initiative, with a major goal of improving the information available to consumers on the quality of care at nursing homes. As part of this effort, on November 12, 2002, CMS released Nursing Home Compare (http://www.medicare.gov/NHcompare), a guide detailing quality of care at over 17,000 Medicare- or Medicaid-certified nursing homes (Centers for Medicare and Medicaid 2002). When Nursing Home Compare was launched, it included 10 quality measures, three of which were measures of quality for patients in postacute care (Centers for Medicare and Medicaid 2002; Harris and Clauser 2002). For the postacute care measures, small SNFs, with fewer than 20 patients over 6 months who qualify for the denominator of a quality measure, are excluded from Nursing Home Compare.

Studies of the effect of public reporting in nursing homes have thus far been mixed. Upon the release of Nursing Home Compare, Zinn et al. (2005) used data published on the Nursing Home Compare website to descriptively examine whether trends in the published quality measures improved in the postpublication period. They found that while 9 out of 10 published measures had statistically significant trends toward improvement, only pain control, use of physical restraints, and rates of delirium seemed to exhibit clear and clinically meaningful trends toward improvement. However, because the study used publicly reported data, it was unable to compare postpublication trends with preexisting trends and, thus, attribution of these trends to Nursing Home Compare was not possible. Similarly, work by Castle, Engberg, and Liu (2007) examined published nursing home quality measures in the post-Nursing Home Compare period with mixed results. A more recent study by Mukamel et al.

(2008) overcame the limitation of only examining the post-Nursing Home Compare period by reconstructing the quality measures during the period both before and after the launch of Nursing Home Compare. They examined changes in trends of a subset of quality measures among a small group of nursing homes and found that two quality measures (physical restraint use and pain control) improved after the launch of Nursing Home Compare. However, without a concurrent control group, it is difficult to attribute these changes to Nursing Home Compare. In addition, as with most studies of public reporting, it is unknown whether observed improvements are associated with improvements in quality of care more broadly and to what extent observed improvements in quality measures were due to changes in patient case mix.

Our work adds significantly to the existing evidence of response to Nursing Home Compare, and to the public reporting literature more generally. First, we explicitly examine within-SNF or SNF-specific changes in quality (versus changes in market share). Second, we use a difference-in-differences framework to compare trends in quality before and after Nursing Home Compare was launched and control for secular trends using a group of small SNFs not included in Nursing Home Compare. Finally, we also employ a broad measure of quality in addition to the individual reported measures, enabling us to examine whether improvement on Nursing Home Compare measures translated into broader improvements in quality of care for postacute care patients. Because the release of Nursing Home Compare may cause SNFs to select patients based on clinical characteristics, we also control for this changing pool of patients using propensity score matching. Thus, our estimates of the quality impact of Nursing Home Compare can be more definitively attributed to true improvements in quality rather than the result of a changing group of patients.

METHODS

We limit our analyses to quality of postacute (rather than long-term) care because the higher turnover rates and younger and less cognitively impaired residents in postacute care make it more likely to find an effect from public reporting. Postacute care residents can also be linked with important case mix adjusters from the qualifying Medicare-covered hospitalization, enhancing our ability to control carefully for changes in case mix severity.

Data

The primary data source for our analyses is the nursing home Minimum Data Set (MDS) for years 1999–2005, spanning 2002, when Nursing Home Compare

was released. The MDS contains detailed clinical data collected at regular intervals for every resident in a Medicare- or Medicaid-certified nursing home. Data on patients' health, physical functioning, mental status, and psycho-social well-being have been collected electronically since 1998. These data are used by nursing homes to assess the needs and develop a plan of care unique to each resident and by the CMS to calculate Medicare prospective reimbursement rates. Because of the reliability of these data (Gambassi et al. 1998; Mor et al. 2003) and the detailed clinical information contained therein, they are considered the best available data for measuring nursing home clinical quality and thus are the source for the quality measures reported on Nursing Home Compare. Calculating the quality measures directly from the MDS allows us to measure quality both before and after Nursing Home Compare was released.

We also used the 100 percent MedPAR data (with all Part A claims) to calculate rates of potentially preventable rehospitalizations over the same time period, an accepted indicator of SNF quality (MedPAC 2005) that broadly measures a major goal of postacute care—stabilization following acute hospitalization (Donelan-McCall et al. 2006). Rehospitalization has thus far not been included in Nursing Home Compare's quality measures because data on rehospitalization are unavailable using MDS, the current data source for the quality measures. However, we linked the MDS data to MedPAR data using unique patient identifiers to determine rehospitalizations, providing an opportunity to examine whether changes in the reported quality measures were associated with broader changes in the quality of care provided at SNFs.

Propensity Score Matching

Because public reporting of patient outcomes may cause providers to select (or cherry-pick) their patients based on their illness severity (Dranove et al. 2003; Werner and Asch 2005), patient characteristics may systematically vary with the launch of public reporting. To account for this possibility, we use propensity score matching to ensure the similarity of patients being compared (Rosenbaum and Rubin 1983; Heckman, Urzua, and Vytlacil 2006), matching patients from before Nursing Home Compare was released to patients with similar characteristics after Nursing Home Compare was released. By matching we assure similar distributions in observable characteristics between the patients being compared before and after Nursing Home Compare and limit our analyses to the units over which comparisons can be reliably made. This thereby reduces the possibility of model extrapolations and the biasing effects of model misspecification (Hill, Reiter, and Zanutto 2004; Ho et al. 2007).

Propensity scores were constructed using a large set of variables measured before or at nursing home admission. In particular, we conducted Mahalanobis matching on three variables (age, Cognitive Performance Scale [Morris et al. 1994], and RUG-III ADL Scale [Morris, Fries, and Morris 1999]) within propensity score calipers (Rubin and Thomas 2000). Each patient in the pre-period was matched to a patient in the postperiod within the same nursing home, and the matching was conducted "with replacement" in the sense that individuals in the postperiod could be selected as a match for up to 10 patients in the preperiod; the final analyses of the outcomes used weights to adjust for this. Propensity score matching thus constrains patients risk profiles to be the same before and after Nursing Home Compare was launched and also constrains each nursing home's market share to be the same. Thirty-three variables were included in the propensity score and an additional 20 variables were assessed and were included in the propensity score if they were out of balance after an initial match. (A full list of variables included in the propensity score and details of the propensity score matching are included in Appendix SA2.)

Dependent Variables: Quality Measures

Using the propensity score-matched cohorts, we applied the technical definitions of the quality measures provided by CMS (Morris et al. 2003; Ho et al. 2007) to calculate quality measures for postacute care patients over the time period of the study, 1999–2005. We calculated all postacute care quality measures that were publicly reported on Nursing Home Compare when it was launched in 2002: percent of short-stay patients who did not have moderate or severe pain; percent of short-stay patients without delirium; and percent of short-stay patients whose walking improved. Following the conventions of the CMS quality measures we calculated the postacute care quality measures only on those patients who stay in the facility long enough to have a 14-day assessment, and limit the SNFs to those that were included in Nursing Home Compare (i.e., those SNFs that had at least 20 eligible postacute care patients over 6 months). Our calculations of SNF-level quality measures on the full sample of postacute care patients consistently matched the quality measures reported by CMS on Nursing Home Compare. We rescaled all reported quality measures so that a higher score indicates higher quality of care. The means and standard deviations for these measures over the study period were as follows: no pain 76.3 (19.3); no delirium 96.4 (7.5); and improved walking 6.9 (10.5).

We also measured rates of potentially preventable rehospitalizations as a broader indicator of SNF quality. Whereas Nursing Home Compare quality

measures are defined based only on patients who stay in postacute care for at least 14 days, we measured rehospitalizations for all postacute care patients regardless of their length of stay. Potentially preventable rehospitalizations were defined based on the Agency for Healthcare Research and Quality Prevention Quality Indicators (Agency for Healthcare Research and Quality 2004) that were applicable to patients aged 65 and older (bacterial pneumonia, chronic obstructive pulmonary disease, dehydration, heart failure, hypertension, short-term diabetic complications, uncontrolled diabetes, and urinary infection), occurring within 30 days of admission to postacute care (White and Seagrave 2005). Higher rehospitalization rates indicate lower quality of care. Over the study period, an average of 7.0 percent of patients (standard deviation 5.9) had a potentially preventable rehospitalization.

Main Independent Variable: Nursing Home Compare Indicator Variables

We examined changes in the quality measures with the launch of Nursing Home Compare using two separate independent variables. First, we used a set of year indicator variables (2000–2005, omitting 1999) to examine changes in quality over the study period. These year indicator variables were used to examine changes in quality at the launch of Nursing Home Compare (in November 2002) by testing the difference between the coefficients on the 2002 and 2003 indicator variables. Second, we used a pre–post indicator variable equaling 1 in the period after Nursing Home Compare was launched (after April 22, 2002, in pilot states and November 12, 2002, in nonpilot states) and 0 otherwise. This pre–post indicator variable was used to test whether the quality level differed in the 3-year period after Nursing Home Compare was launched compared with the pre-Nursing Home Compare period.

Covariates

In all analyses we included the variables included in the propensity scores as covariates (see Appendix SA2) to adjust for any remaining small differences between the groups after the matching (Ho et al. 2007). In addition, to estimate changes in delirium we include prior residential history as a covariate, as specified by the CMS quality measure (Moore et al. 2005); and when estimating changes in rehospitalization rates, we include previously developed variables for risk adjustment of potentially preventable rehospitalization from postacute care (Donelan-McCall et al. 2006). Finally, because the Nursing Home Compare quality measures are calculated only on patients who remain in postacute care for at least 14 days, in all regressions we control for the "censoring" rate at

each facility using quarterly measures of the proportion of all postacute care admissions that remain in postacute care for 14 days at each SNF.

Empirical Specifications

To examine the effect of publicly reporting nursing home quality on nursing home quality of care, we test for within-SNF changes in facility-level quality. Because the propensity score matched cohorts explicitly constrains changes in market share (through 1:1 pre–post matching of patients within SNFs) we empirically isolate one way that Nursing Home Compare was designed to affect quality of care—through provider-driven quality improvements—while controlling for any consumer-driven changes in care.

First, we describe changes in postacute care quality of care using a pre–post specification. Within this pre–post specification, we test for within-SNF improvements in quality of care using individual-level linear probability models, where changes in quality for patient i in SNF j at time t were estimated as a function of Nursing Home Compare indicator variables, patient- and SNF-level covariates, and SNF fixed effects:

$$Quality_{i,j,t} = \beta_j + \beta_1 NHC_t + \beta X_{i,j,t} + \varepsilon_{i,j,t} \tag{1}$$

Second, we test whether the estimated changes in postacute care quality are attributable to Nursing Home Compare using a difference-in-differences specification, using small SNFs not subject to public reporting as a control group. Although one third of SNFs are excluded from Nursing Home Compare at any given time, many of these small SNFs are only intermittently excluded from Nursing Home Compare when their census drops below the 20-patient threshold. We include the 15 percent of SNFs that are *never* included in Nursing Home Compare as our control group. Empirically, we estimate changes in quality as a function of a Nursing Home Compare indicator variable, an SNF size indicator variable (*Large$_j$*, equaling 1 if a large SNF is included in Nursing Home Compare; 0 if an SNF is never included), its interaction with $NHC_{j,t}$, patient- and SNF-level covariates, and SNF fixed effects:

$$Quality_{i,j,t} = \beta_j + \beta_1 NHC_t + \beta_2 Large_j \times NHC_t + \beta X_{i,j,t} + \varepsilon_{i,j,t} \tag{2}$$

Although small SNFs are clearly different from larger SNFs on average, this is not a problem in that these average differences are controlled through the SNF fixed effects. This model does rely on the assumption that small and large SNFs would exhibit similar trends over time in the absence of Nursing Home Compare. To test the validity of this assumption, we test whether trends in quality improvement in the pre-Nursing Home Compare period were the same for

small and large SNFs using multiple *F*-tests and find that for all four measures there were no significant differences in quality trends between small and large SNFs before Nursing Home Compare. Nonetheless, because small SNFs are likely to be different than large SNFs in their response to quality improvement initiatives, we test the sensitivity of our results by limiting the treatment group to mid-sized SNFs, or those that are large enough to be included in Nursing Home Compare but more similar in size and other characteristics to the small SNFs (mid-sized SNFs were defined as those included in Nursing Home Compare with fewer than 110 total beds and <30 percent Medicare residents). In addition, because nursing homes with few postacute care residents may have larger numbers of long-stay residents, and thus be included in Nursing Home Compare for the long-stay measures, we tested the correlation between the short-stay and long-stay measures and found it to be low (range 0.04–0.06). Robust standard errors were used to account for nonindependence of observations from the same facility in all regressions (Huber 1967; White 1980).

RESULTS

A total of 8,137 SNFs from Nursing Home Compare were included in the study, covering 9,390,930 postacute care stays and 5,899,327 postacute care stays of at least 14 days. An additional 2,277 small SNFs, covering 442,952 postacute care stays and 214,094 postacute care stays of at least 14 days, were not included in Nursing Home Compare and serve as a control group. Characteristics of these SNFs, stratified by size, are summarized in Table 1. By definition, large and mid-sized SNFs (those included in Nursing Home Compare) had more beds. Mid-sized SNFs were more similar to small SNFs in size and percent of Medicare residents, but were more likely to be part of a chain and be for-profit facilities than small and large SNFs. Mid-sized facilities were also less likely to be hospital based and had fewer staff hours per resident day compared with large and small SNFs. These characteristics of mid-sized SNFs are often associated with lower SNF quality.

All three reported measures of quality improved over time, as did rates of potentially preventable rehospitalizations. (See Figure 1 for risk-adjusted trends in postacute quality of care.) Multivariate regression results testing within-SNF changes in quality of care associated with the launch of Nursing Home Compare are summarized in Table 2. The three reported quality measures were better in the 3 years after Nursing Home Compare was launched compared with before. Over those years, the percent of patients without moderate to severe pain improved by 2.0 percentage points (on a base of 76

Table 1:   Characteristics (Mean and Standard Deviation) of Skilled Nursing Facilities (SNFs) Included in Public Reporting through Nursing Home Compare (Large SNFs and Mid-Sized SNFs) and Those Excluded from Nursing Home Compare (Small SNFs)

	Included in Nursing Home Compare		*Not Included in Nursing Home Compare*
	Large SNFs (n = 5,970)	*Mid-Sized SNFs (n = 2,167)*	*Small SNFs (n = 2,277)*
Total residents	126.1 (74.8)	76.5 (18.1)	53.4 (32.2)
Total beds	152.2 (85.5)	86.7 (17.3)	71.3 (45.5)
Certified beds	143.5 (80.4)	85.4 (18.6)	67.0 (40.0)
Total occupancy	82.8% (18.8%)	88.3% (11.6%)	77.9% (21.4%)
% Medicaid	57.2% (26.6%)	61.8% (19.3%)	59.8% (28.8%)
% Medicare	21.6% (25.0%)	13.0% (6.3%)	11.9% (22.5%)
Chain	58.8% (49.2%)	68.7% (46.3%)	47.2% (49.9%)
Hospital based	12.2% (32.7%)	4.9% (21.7%)	12.2% (32.7%)
Total staff hours per resident day	4.4 (14.2)	3.8 (9.2)	5.0 (24.6)
Ownership			
For profit	67.8% (46.7%)	77.8% (41.6%)	59.0% (49.2%)
Not for profit	27.8% (44.8%)	20.2% (40.1%)	34.0% (47.4%)
Government	4.4% (20.5%)	2.0% (13.9%)	7.0% (26.5%)

Mid-sized SNFs were defined as nursing home with fewer than 110 total beds and <30% Medicare residents but large enough to be included in Nursing Home Compare. Values represent the averages from 1999 to 2005.

percent), the percentage of patients without delirium improved by 0.5 percentage points (on a base of 96 percent) and the percentage of patients with improved walking improved by 0.2 percentage points (on a base of 7 percent). While there was a significant reduction in rates of potentially preventable rehospitalizations in the 3-year period after Nursing Home Compare was launched, there was a slight worsening (an increase of 0.3 percentage points) in the year immediately following the launch of Nursing Home Compare (on a base of 7 percent). For most measures, the largest change in quality occurred between 2002 and 2003, when Nursing Home Compare was launched, with a subsequent continuing trend toward improvement.

Using a difference-in-differences framework with small SNFs as a control group produced similar results. Compared with before Nursing Home Compare was launched, pain and walking improved after the launch of Nursing Home Compare (Table 3). The magnitude of improvement in pain decreased when controlling for changes at small SNFs but the magnitude of improvement in walking increased. In contrast to the pre–post model, there

Figure 1: Risk-Adjusted Trends in Postacute Care Quality
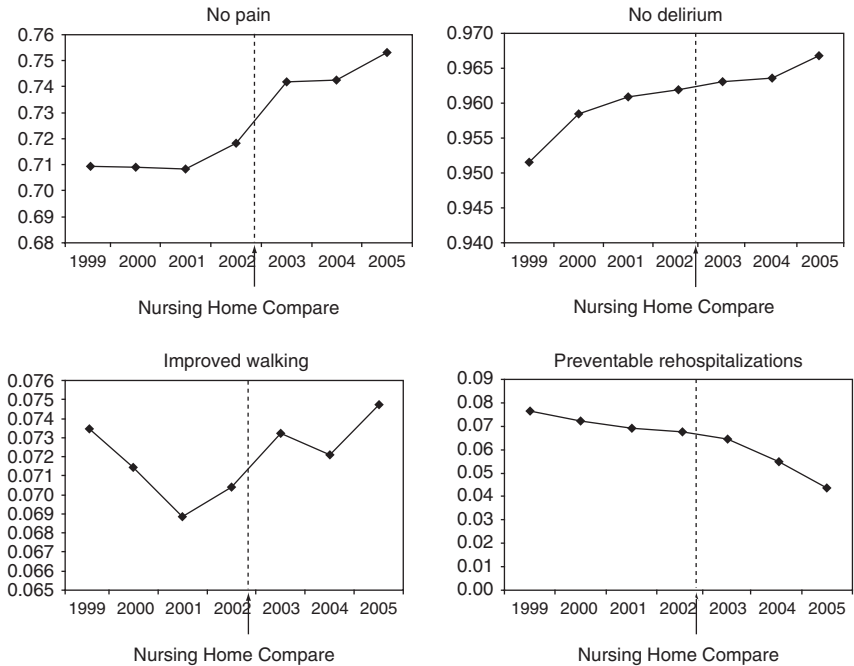
The vertical dashed line represents November 2002, when Nursing Home Compare was launched.

were no significant changes in within-SNF rates of delirium. Rates of rehospitalization worsened slightly in the year immediately after the launch of Nursing Home Compare, and they did not improve over the 3-year period after Nursing Home Compare in large SNFs relative to small SNFs.

Finally, we limited the comparison between large and small SNFs to midsized SNFs, rather than all large SNFs, by excluding SNFs with > 110 total beds and > 30 percent Medicare residents. This did not significantly change the results.

DISCUSSION

We found that most postacute care quality measures included in Nursing Home Compare significantly improved after the launch of Nursing Home Compare. The one measure that did not improve (delirium) had high performance levels before Nursing Home Compare and, thus, may not have had sufficient room to improve. While the quality improvements were statistically

Table 2:   Within-SNF Changes in Postacute Care Measures after Nursing Home Compare Was Released from Multivariate Regression with SNF Fixed Effects

	No Pain	No Delirium	Improved Walking	Rehospitalizations
1999 (omitted)				
2000	− 0.0004	0.0049***	− 0.0025***	− 0.0044***
	(0.0006)	(0.0004)	(0.0004)	(0.0004)
2001	0.0006	0.0066***	− 0.0036***	− 0.0113***
	(0.0007)	(0.0005)	(0.0005)	(0.0004)
2002	0.0082***	0.0082***	− 0.0033***	− 0.0104***
	(0.0008)	(0.0005)	(0.0005)	(0.0005)
2003	0.0213***	0.0102***	− 0.0014**	− 0.0077***
	(0.0010)	(0.0006)	(0.0006)	(0.0006)
2004	0.0227***	0.0104***	− 0.0007	− 0.0166***
	(0.0011)	(0.0006)	(0.0007)	(0.0006)
2005	0.0242***	0.0101***	0.0037***	− 0.0239***
	(0.0011)	(0.0006)	(0.0008)	(0.0006)
Constant	0.9111***	0.9886***	0.0739***	0.0191***
	(0.0035)	(0.0021)	(0.0023)	(0.0025)
Observations	5,529,400	5,443,268	5,400,913	8,915,689
Number of facilities	8,137	8,137	8,137	8,138
R^2	0.52	0.20	0.38	0.18
Change at implementation of NHC[†]	0.0131***	0.0020***	0.0019***	0.0027***
	(0.0009)	(0.0005)	(0.0005)	(0.0005)
Change between pre- and post-NHC[‡]	0.0210***	0.0051***	0.0024***	− 0.0081***
	(0.0007)	(0.0003)	(0.0004)	(0.0003)

Robust standard errors in parentheses.

***$p < .01$, **$p < .05$.

[†]Change at implementation of NHC is based on regression with year indicator variables using the difference between the 2002 and 2003 indicators.

[‡]Change between pre- and post-NHC is based on regression using pre- versus post-NHC indicator variable equaling 1 in the period after Nursing Home Compare was launched (after April 22, 2002, in pilot states and November 12, 2002, in nonpilot states) and 0 otherwise.

SNF, skilled nursing facility.

significant, in some cases the magnitude of change was small. The relative improvement in pain control attributable to Nursing Home Compare (based on the differences-in-differences specification) was <1 percent. However, with over 1.5 million patients admitted to postacute care annually, this quality improvement translates into approximately 12,000 fewer patients having moderate to severe pain. Improvement in walking had a relative improvement of 9 percent. While measured quality improved, changes in rates of potentially preventable rehospitalizations (a broader measure of quality) were inconsistent and, in some cases, worsened.

Table 3:　Within-SNF Changes in Postacute Care Measures after Nursing Home Compare Was Released Using Small SNFs as Controls from Multivariate Regressions with SNF Fixed Effects

	No Pain	No Delirium	Improved Walking	Rehospitalizations
1999 (omitted)				
2000	0.0078***	0.0065***	− 0.0037**	− 0.0019
	(0.0029)	(0.0020)	(0.0018)	(0.0013)
2001	0.0019	0.0092***	− 0.0086***	− 0.0093***
	(0.0028)	(0.0022)	(0.0021)	(0.0015)
2002	0.0076**	0.0117***	− 0.0064**	− 0.0075***
	(0.0034)	(0.0023)	(0.0025)	(0.0018)
2003	0.0129 ***	0.0111***	− 0.0097***	− 0.0089***
	(0.0044)	(0.0029)	(0.0032)	(0.0022)
2004	0.0221***	0.0091***	− 0.0062**	− 0.0140***
	(0.0048)	(0.0028)	(0.0031)	(0.0022)
2005	0.0215 ***	0.0139***	− 0.0082***	− 0.0206***
	(0.0037)	(0.0025)	(0.0028)	(0.0022)
1999 × large SNF (omitted)				
2000 × large SNF	− 0.0082***	− 0.0015	0.0012	− 0.0025*
	(0.0029)	(0.0020)	(0.0019)	(0.0014)
2001 × large SNF	− 0.0013	− 0.0026	0.0050**	− 0.0019
	(0.0029)	(0.0022)	(0.0021)	(0.0015)
2002 × large SNF	0.0006	− 0.0035	0.0031	− 0.0028
	(0.0035)	(0.0024)	(0.0025)	(0.0019)
2003 × large SNF	0.0083*	− 0.0009	0.0083**	0.0013
	(0.0045)	(0.0030)	(0.0033)	(0.0023)
2004 × large SNF	0.0005	0.0012	0.0056*	− 0.0026
	(0.0049)	(0.0029)	(0.0032)	(0.0022)
2005 × large SNF	0.0026	− 0.0038	0.0119***	− 0.0032
	(0.0039)	(0.0026)	(0.0029)	(0.0023)
Constant	0.9110***	0.9884***	0.0746***	0.0166***
	(0.0034)	(0.0021)	(0.0022)	(0.0024)
Observations	5,743,494	5,653,807	5,609,763	9,358,641
Number of clusters	9,071	9,071	9,071	9,232
R^2	0.52	0.20	0.38	0.18
Change at implementation of NHC[†]	0.0078*	0.0026	0.0052*	0.0042*
	(0.0044)	(0.0029)	(0.0029)	(0.0024)
Change between pre- and post- NHC[‡]	0.0061**	− 0.0012	0.0066***	0.0022
	(0.0025)	(0.0015)	(0.0017)	(0.0014)

Robust standard errors in parentheses.

***$p<.01$, **$p<.05$, *$p<.1$.

[†]Change at implementation of NHC is based on regression with year indicator variables using the difference between the 2002 and 2003 indicators.

[‡]Change between pre- and post-NHC is based on regression using pre- versus post-NHC indicator variable equaling 1 in the period after Nursing Home Compare was launched (after April 22, 2002, in pilot states and November 12, 2002, in nonpilot states) and 0 otherwise.

SNF, skilled nursing facility.

The quality improvements we found were explained by within-SNF changes in quality of care rather than through other mechanisms such as increasing market share in high-quality facilities. It is unknown whether market share changed in response to public reporting in this setting; however, prior work has found that public reporting has a limited effect on market share (Mukamel and Mushlin 1998; Cutler, Huckman, and Landrum 2004; Romano and Zhou 2004; Howard and Kaplan 2006; Jha and Epstein 2006), possibly due to capacity constraints and informal quality information available before public reporting (Mukamel, Weimer, and Mushlin 2007).

There is ample literature suggesting that in the presence of external incentives, such as public reporting, measured quality improves on average. This has been documented in the case of hospitals (Williams et al. 2005), physicians (Hannan et al. 1994), health plans, and previously in nursing homes (Mukamel et al. 2008). Our findings, which improve upon the methods of prior studies, align with this prior empirical work. While there has been concern that some improvements due to public reporting might be due to changes in case mix (Werner and Asch 2005), we found that after extensive controls for this, improvements in measured quality remain.

We also found that potentially preventable rehospitalization rates did not change or worsened slightly after the launch of Nursing Home Compare. This may be counterintuitive. However, there are several potential explanations for this finding. First, prior work has found that narrow measures of quality are often not well correlated with broader measures of quality (Bradley et al. 2006; Werner and Bradlow 2006). Although strategies to improve measured quality may spill over to areas that affect quality more generally, and thus in the case of postacute care may reduce rates of hospitalization, this need not be the case. Indeed, efforts targeted at improving the three quality measures used in Nursing Home Compare may not affect rehospitalization from postacute care. Second, the divergence of Nursing Home Compare measures and potentially preventable rehospitalizations may be due to the different cohorts included in these measures. While rehospitalizations are measured for all postacute care admissions, the Nursing Home Compare measures reflect quality only among patients who stay in postacute care for at least 14 days. Characteristics of patients at 5 and 14 days may differ sufficiently to explain these divergent findings. One reason 5- and 14-day patients may differ is that SNFs may selectively discharge (and rehospitalize) postacute care patients before the 14-day assessment when Nursing Home Compare quality measurement occurs. In this way, it is possible that Nursing Home Compare contributes to higher rates of rehospitalization from postacute care. While we

control for changing case mix at SNFs using propensity score matching, and control for differential rates of discharge before day 14 across SNFs, we are unable to adequately control for rates of selective discharge in this analysis. Disentangling the effects of selective discharge from true changes in quality is an important area for future research.

Our results should be interpreted in light of the study's limitations. First, results based on propensity score-matched cohorts may not be representative of the changes in quality that occurred as we do not include all patients in our analyses. (Rather, we limit our analyses to the subset of patients with similar characteristics for whom we can best estimate the effects of Nursing Home Compare.) Nonetheless, these results are important from a policy standpoint, as they provide more precise estimates of the effect of public reporting on quality of care. Second, although we extensively control for patient selection based on observed differences between patients before versus after the launch of Nursing Home Compare, unobserved differences that are uncorrelated with observed differences remain a threat to the validity of our findings. Nonetheless, prior work suggests that using propensity score matching in the setting of an exogenous treatment decision, such as the launch of Nursing Home Compare, is a valid approach to account for differences across groups (Heckman, Urzua, and Vytlacil 2006). Finally, the quality changes we demonstrate may be due to changes in data accuracy rather than true quality changes, particularly given the subjective nature of the Nursing Home Compare quality measures. While other work has found that data changes explain some quality improvement (Green and Wintfeld 1995; Roski et al. 2003), this is less likely in the nursing home quality measures based on MDS. Electronic MDS data collection started in 1998, well before Nursing Home Compare was launched, and has been used to determine Medicare payment since that time, increasing nursing homes' incentive to accurately report these data for several years before the launch of Nursing Home Compare.

Despite these limitations, our study offers important new findings with regard to the role of public reporting in quality improvement. We find that most quality measures improve in response to public reporting even after controlling for secular trends. However, the clinical significance of these improvements may be limited given that improvements in narrow measures of quality of care may not translate into broader quality improvement. To achieve more robust quality improvement, stronger incentives to improve quality may be needed. One possible way to do this is to combine public reporting with pay for performance. While public reporting of quality information is important for reasons beyond quality improvement, such as

enhancing accountability of health care providers, alone it may play a positive but limited role in improving quality of care.

ACKNOWLEDGMENTS

REFERENCES

Agency for Healthcare Research and Quality. 2004. "Prevention Quality Indicators Overview. AHRQ Quality Indicators [accessed on March 5, 2009]." Available at http://www.qualityindicators.
ahrq.gov/pqi_overview.htm

Bradley, E. H., J. Herrin, B. Elbel, R. L. McNamara, D. J. Magid, B. K. Nallamothu, Y. Wang, S.-L. T. Normand, J. A. Spertus, and H. M. Krumholz. 2006. "Hospital Quality for Acute Myocardial Infarction: Correlation among Process Measures and Relationship with Short-Term Mortality." *Journal of the American Medical Association* 296 (1): 72–8.

Castle, N. G., J. Engberg, and D. Liu. 2007. "Have Nursing Home Compare Quality Measure Scores Changed over Time in Response to Competition?" *Quality and Safety in Health Care* 16 (3): 185–91.

Castle, N. G., B. S. Fogel, and V. Mor. 1996. "Study Shows Higher Quality of Care in Facilities Administered by ACHCA Members." *Journal of Long Term Care Administration* 24 (2): 11–6.

Centers for Medicare and Medicaid. 2003. "Nursing Home Quality Initiatives Overview" [accessed on March 5, 2009]. Available at http://www.cms.hhs.gov/NursingHomeQualityInits/35_NHQIArchives.asp

Cutler, D. M., R. S. Huckman, and M. B. Landrum. 2004. "The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery." *American Economic Review* 94 (2): 342–6.

Donelan-McCall, N., T. Eilertsen, R. Fish, and A. Kramer. 2006. *Small Patient Population and Low Frequency Event Effects on the Stability of SNF Quality Measures*. Washington, DC: MedPAC.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite. 2003. "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers." *Journal of Political Economy* 111 (3): 555–88.

Fries, B. E., C. Hawes, J. N. Morris, C. D. Phillips, V. Mor, and P. S. Park. 1997. "Effect of the National Resident Assessment Instrument on Selected Health Conditions and Problems." *Journal of American Geriatric Society* 45 (8): 994–1001.

Fung, C. H., Y.-W. Lim, S. Mattke, C. Damberg, and P. G. Shekelle. 2008. "Systematic Review: The Evidence That Publishing Patient Care Performance Data Improves Quality of Care." *Annals of Internal Medicine* 148 (2): 111–23.

Gambassi, G., F. Landi, L. Peng, C. Brostrup-Jensen, K. Calore, J. Hiris, L. Lipsitz, V. Mor, and R. Bernabei. 1998. "Validity of Diagnostic and Drug Data in Standardized Nursing Home Resident Assessments: Potential for Geriatric Pharmacoepidemiology. SAGE Study Group. Systematic Assessment of Geriatric Drug Use via Epidemiology." *Medical Care* 36 (2): 167–79.

Green, J., and N. Wintfeld. 1995. "Report Cards on Cardiac Surgeons. Assessing New York State's Approach." *New England Journal of Medicine* 332 (18): 1229–32.

Hannan, E. L., H., Kilburn Jr., M. Racz, E. Shields, and M. R. Chassin. 1994. "Improving the Outcomes of Coronary Bypass Surgery in New York State." *Journal of the American Medical Association* 271: 761–6.

Harris, Y., and S. B. Clauser. 2002. "Achieving Improvement through Nursing Home Quality Measurement." *Health Care Financial Review* 23 (4): 5–18.

Heckman, J. J., S. Urzua, and E. Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3): 389–432.

Hill, J. L., J. P. Reiter, and E. L. Zanutto. 2004. "A Comparison of Experimental and Observational Data Analyses." In *Applied Bayesian Modeling and Causal Inference From an Incomplete-Data Perspective*, Vol. 1, edited by A. Gelman and X.-L. Meng, pp. 44–56. New York: John Wiley.

Ho, D. E., K. Imai, G. King, and E. A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199–236.

Howard, D. H., and B. Kaplan. 2006. "Do Report Cards Influence Hospital Choice? The Case of Kidney Transplantation." *Inquiry: A Journal of Medical Care Organization, Provision and Financing* 43 (2): 150–9.

Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Non-Standard Conditions." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, edited by Lucien M. Le Cam and Jerzy Neyman, pp. 221–233. Berkeley, CA: University of California Press.

Institute of Medicine. 1986. *Improving the Quality of Care in Nursing Homes*. Washington, DC: National Academies Press.

———. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press.

Jha, A. K., and A. M. Epstein. 2006. "The Predictive Accuracy of the New York State Coronary Artery Bypass Surgery Report-Card System." *Health Affairs* 25 (3): 844–55.

Kane, R. L., C. C. Williams, T. F. Williams, and R. A. Kane. 1993. "Restraining Restraints: Changes in a Standard of Care." *Annual Review of Public Health* 14: 545–84.

Kohn, L. T., J. M. Corrigan, and M. S. Donaldson. 1999. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academies Press.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. 2003. "The Quality of Health Care Delivered to Adults in the United States." *New England Journal of Medicine* 348: 2635–45.

MedPAC. 2005. *Report to Congress: Medicare Payment Policy (March 2005)*. Washington, DC: Medicare Payment Advisory Commission.

Moore, T., N. Wu, D. Kidder, B. Bell, D. Warner, L. Hadden, Y. Mackiernan, J. Morris, and R. Jones. 2005. *Design and Validation of Post-Acute Care Quality Measures*. Abt Associates Subcontract: RAND Prime Contract No. 500-00-0029 (TO 2), Abt Associates Inc.

Mor, V. 2005. "Improving the Quality of Long-Term Care with Better Information." *Milbank Quarterly* 83 (3): 333–64.

Mor, V., J. Angelelli, R. Jones, J. Roy, T. Moore, and J. Morris. 2003. "Inter-Rater Reliability of Nursing Home Quality Indicators in the U.S." *BMC Health Services Research* 3: 20.

Mor, V., O. Intrator, B. E. Fries, C. Phillips, J. Teno, J. Hiris, C. Hawes, and J. Morris. 1997. "Changes in Hospitalization Associated with Introducing the Resident Assessment Instrument." *Journal of American Geriatric Society* 45 (8): 1002–10.

Morris, J. N., B. E. Fries, D. R. Mehr, C. Hawes, C. Phillips, V. Mor, and L. A. Lipsitz. 1994. "MDS Cognitive Performance Scale." *Journal of Gerontology* 49 (4): M174–82.

Morris, J. N., B. E. Fries, and S. A. Morris. 1999. "Scaling ADLs within the MDS." *Journal of Gerontology Series A: Biological Sciences and Medical Sciences* 54 (11): M546–53.

Morris, J. N., T. Moore, R. Jones, V. Mor, J. Angelelli, K. Berg, C. Hale, S. Morriss, K. M. Murphy, and M. Rennison. 2003. *Validation of Long-Term and Post-Acute Care Quality Indicators*. Baltimore, MD: Centers for Medicare and Medicaid Services.

Mukamel, D. B., and A. I. Mushlin. 1998. "Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports." *Medical Care* 36: 945–54.

Mukamel, D. B., D. L. Weimer, and A. I. Mushlin. 2007. "Interpreting Market Share Changes as Evidence for Effectiveness of Quality Report Cards." *Medical Care* 45 (12): 1227–32.

Mukamel, D. B., D. L. Weimer, W. D. Spector, H. Ladd, and J. S. Zinn. 2008. "Publication of Quality Report Cards and Trends in Reported Quality Measures in Nursing Homes." *Health Services Research* 43 (4): 1244–62.

Romano, P. S., and H. Zhou. 2004. "Do Well-Publicized Risk-Adjusted Outcomes Reports Affect Hospital Volume?" *Medical Care* 42: 367–77.

Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41–55.

Roski, J., R. Jeddeloh, L. An, H. Lando, P. Hannan, C. Hall, and S. H. Zhu. 2003. "The Impact of Financial Incentives and a Patient Registry on Preventive Care Quality: Increasing Provider Adherence to Evidence-Based Smoking Cessation Practice Guidelines." *Preventive Medicine* 36 (3): 291–9.

Rubin, D. B., and N. Thomas. 2000. "Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates." *Journal of the American Statistical Association* 95 (450): 573–85.

Shorr, R. I., R. L. Fought, and W. A. Ray. 1994. "Changes in Antipsychotic Drug Use in Nursing Homes during Implementation of the OBRA-87 Regulations." *Journal of the American Medical Association* 271 (5): 358–62.

Snowden, M., and P. Roy-Byrne. 1998. "Mental Illness and Nursing Home Reform: OBRA-87 Ten Years Later. Omnibus Budget Reconciliation Act." *Psychiatric Services* 49 (2): 229–33.

Werner, R. M., and D. A. Asch. 2005. "The Unintended Consequences of Publicly Reporting Quality Information." *Journal of the American Medical Association* 293 (10): 1239–44.

Werner, R. M., and E. T. Bradlow. 2006. "Relationship between Medicare's Hospital Compare Performance Measures and Mortality Rates." *Journal of the American Medical Association* 296 (22): 2694–702.

White, C., and S. Seagrave. 2005. "What Happens When Hospital-Based Skilled Nursing Facilities Close? A Propensity Score Analysis." *Health Services Research* 40 (6, part 1): 1883–97.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48: 817–30.

Williams, S. C., S. P. Schmaltz, D. J. Morton, R. G. Koss, and J. M. Loeb. 2005. "Quality of Care in U.S. Hospitals as Reflected by Standardized Measures, 2002–2004." *New England Journal of Medicine* 353 (3): 255–64.

Wunderlich, G. S., and P. Kohler. 2000. *Improving the Quality of Long-Term Care.* Washington, DC: Division of Health Care Services, Institute of Medicine.

Zinn, J., W. Spector, L. Hsieh, and D. B. Mukamel. 2005. "Do Trends in the Reporting of Quality Measures on the Nursing Home Compare Web Site Differ by Nursing Home Characteristics?" *Gerontologist* 45 (6): 720–30.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.
Appendix SA2: The Impact of Public Reporting on Quality of Postacute Care: Propensity-Score Methods.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.