Flex Monitoring Team Briefing Paper No. 23

The Potential Impact of Pay-for-Performance on the Financial Health of Critical Access Hospitals

December 2009

A Performance Monitoring Resource for
Critical Access Hospitals, States, and Communities

Flex Monitoring Team
University of Minnesota
University of North Carolina at Chapel Hill
University of Southern Maine

The Flex Monitoring Team is a consortium of the Rural Health Research Centers located at the Universities of Minnesota, North Carolina at Chapel Hill, and Southern Maine. Under contract with the Federal Office of Rural Health Policy (PHS Grant No. U27RH01080), the Flex Monitoring Team is cooperatively conducting a performance monitoring project for the Medicare Rural Hospital Flexibility Program (Flex Program). The monitoring project is assessing the impact of the Flex Program on rural hospitals and communities and the role of states in achieving overall program objectives, including improving access to and the quality of health care services; improving the financial performance of Critical Access Hospitals; and engaging rural communities in health care system development.

This report was prepared by Robert Town, PhD and Ira Moscovice, PhD, University of Minnesota.

Questions regarding the report should be addressed to: Ira Moscovice, 612-624-8618, mosco001@umn.edu.

Flex Monitoring Team
http://www.flexmonitoring.org

University of Minnesota
Division of Health Services Research & Policy
420 Delaware Street, SE, Mayo Mail Code 729
Minneapolis, MN 55455-0392
612.624.8618

University of North Carolina at Chapel Hill
Cecil B. Sheps Center for Health Services Research
725 Martin Luther King Jr. Boulevard, CB #7590
Chapel Hill, NC 27599-7590
919.966.5541

University of Southern Maine
Muskie School of Public Service
PO Box 9300
Portland, ME 04104-9300
207.780.4435

The Medicare Rural Hospital Flexibility Program

The Medicare Rural Hospital Flexibility Program (Flex Program), created by Congress in 1997, allows small hospitals to be licensed as Critical Access Hospitals (CAHs) and offers grants to States to help implement initiatives to strengthen the rural health care infrastructure. To participate in the Flex Grant Program, States are required to develop a rural health care plan that provides for the creation of one or more rural health networks; promotes regionalization of rural health services in the State; and improves the quality of and access to hospital and other health services for rural residents of the State. Consistent with their rural health care plans, states may designate eligible rural hospitals as CAHs.

CAHs must be located in a rural area or an area treated as rural; be more than 35 miles (or 15 miles in areas with mountainous terrain or only secondary roads available) from another hospital or be certified before January 1, 2006 by the State as being a necessary provider of health care services. CAHs are required to make available 24-hour emergency care services that a State determines are necessary. CAHs may have a maximum of 25 acute care and swing beds, and must maintain an annual average length of stay of 96 hours or less for their acute care patients. CAHs are reimbursed by Medicare on a cost basis (i.e., for the reasonable costs of providing inpatient, outpatient and swing bed services).

The legislative authority for the Flex Program and cost-based reimbursement for CAHs are described in the Social Security Act, Title XVIII, Sections 1814 and 1820, available at http://www.ssa.gov/OP_Home/ssact/title18/1800.htm

Table of Contents

List of Table and Figures

EXECUTIVE SUMMARY

Introduction

This paper models the impact of different pay-for-performance (P4P) incentives on the financial health of Critical Assess Hospitals (CAHs).

Methods

The impact of P4P on CAHs is modeled by simulating the change in Medicare revenue using different exchange functions. The exchange function translates hospital quality outcomes into payments. Data sources for the study include 1) Hospital Compare, CMS's public reporting system for hospitals, 2) Hospital Cost Report Information System (HCRIS), and 3) the Flex Monitoring Team's census of CAHs. The analysis is limited to CAHs that had converted by 2006. The quality performance measures are composite quality scores for the conditions of pneumonia and heart failure. The financial measure used is an estimate of Medicare inpatient revenue based on HCRIS data.

Results

The analysis finds that for pneumonia and heart failure, hospitals that provide higher quality of care also are more profitable as measured by net revenue. P4P incentives likely reduce the financial health of hospitals already in financial distress. However, the impact of commonly used P4P incentive structures on CAHs is modest (i.e., the number of CAHs in financial distress increases by approximately one percentage point).

Conclusions

An obvious concern with P4P is that it may negatively affect the financial stability of hospitals that are in a precarious position. CAHs are prime candidates for P4P programs to have such unintended consequences. However, our work suggests that P4P incentives are likely to

have, at best, only a modest impact on the financial stability of the CAHs that are already under significant financial pressure. Thus, if P4P programs are able to induce hospitals to increase quality, those benefits need not be weighed against the risk of putting already financially distressed CAHs in greater financial jeopardy. The results suggest that CAHs should be in included in future P4P initiatives so we can better understand how payment incentives affect the quality of care in small rural hospitals.

INTRODUCTION

Ever since the Institute of Medicine's report "To Err is Human," there have been widespread calls for reforming Medicare's payment mechanism (Kohn, et al., 2000). Pay-for-performance (P4P) advocates argue that providing higher payments to providers for higher quality care will lead providers to respond to those incentives by devoting resources to quality improvement. Some recent evidence suggests that modest quality incentives can have a significant impact on quality of care. In a randomized trial of the impact of P4P on hospital quality , Lindenaur, et al., (2007) found that a 2% and 1% increase in Medicare payments to hospitals in the first and second deciles, respectively, increased clinical performance for heart failure, pneumonia and acute myocardial infarction (AMI).

While P4P has the potential to improve clinical quality and the patient's experience receiving care, it also may have a broader impact on the health care infrastructure. In order for P4P to have its desired consequences, it must put providers at meaningful financial risk. Thus, financially struggling providers might find themselves in even worse financial condition under a P4P initiative. Werner et al. (2008) have shown that under P4P safety net hospitals would experience a reduction in revenues and that, in turn, may ultimately compromise the quality of care for underserved populations.

In this analysis, we explore the financial impact of different possible P4P programs on Critical Access Hospitals (CAH). To be designated as a CAH, a hospital must be located in a rural area or an area treated as rural; be more than 35 miles (or 15 miles in areas with mountainous terrain or only secondary roads available) from another hospital or be certified before January 1, 2007 by the State as being a necessary provider of health care services. CAHs are required to make available 24-hour emergency care services that a State determines are

necessary. CAHs may have a maximum of 25 acute care and swing beds, and must maintain an annual average length of stay of 96 hours or less for their acute care patients. CAHs are reimbursed by Medicare on a cost basis (i.e., for the reasonable costs of providing inpatient, outpatient and swing bed services). A total of 1,305 hospitals are currently certified as CAHs (Flex Monitoring Team Website, 10/01/09).

In principle, CAHs can use the increase in revenue from cost-based reimbursement to improve their financial health and invest in quality improvement, HIT and other initiatives. However, P4P incentive plans have the potential to unwind, at least for some hospitals, the goal of improving their financial health. Thus, these two public policy initiatives may counteract one another. Clearly, if public programs are to be effectively designed and coordinated, they should not work to simultaneously undo the work of another public initiative. Werner et al. (2008) documents how P4P initiatives can have a significant and negative impact on the financial stability of safety net hospitals which in turn can undermine the goal of improving health care access.

We investigate the impact of P4P on CAHs by simulating the change in Medicare revenue using different exchange functions. The exchange function translates hospital quality outcomes into payments. The impact of a given P4P strategy on the overall financial health of all hospitals depends upon the initial distribution of financial well-being and, importantly, on the relationship between initial quality and hospital profitability. Our analysis examines this relationship and finds that, for pneumonia and heart failure (the conditions the analysis focuses upon); hospitals that provide higher quality care also are more profitable as measured by net revenue.

While we find that CAHs with higher quality of care metrics also are more profitable, our analysis suggests that P4P incentives programs will not meaningfully increase (approximately one percentage point) the number of hospitals in financial distress.

METHODOLOGY

Our simulation strategy is straightforward. Let R_i denotes the revenue of hospital *i* prior to any P4P adjustments and let $x(Q)$ be the exchange correspondence that maps the vector of hospital quality metrics into the change in revenue vector, ΔR. For a given exchange function and cost function, *C(Q),* the post-P4P adjusted net revenue is:

$$\pi(Q, R_i, x) = R_i - C(Q) - x_i(Q)$$

where $x_i(Q)$ refers to the i^{th} element of $x(\cdot)$ and is hospital *i*'s change in revenue from the P4P initiative given its performance and the performance of all the other hospitals.

Our analysis focuses on the relationship between different exchange functions and the distribution of $\pi_i(Q, R_i, x)$ focusing on the lower tail of the distribution. If lower quality hospitals are less profitable then the P4P initiative will exacerbate their financial difficulties. We compare the current distribution of net revenue (which we define more precisely below) to the distribution that arises under different exchange functions including:

Scenario 1: Pay the top 1st and 2nd deciles of the quality distribution 2% and 1% more, respectively, and reduce the payments to the 9th and 10th deciles by 1% and 2%, respectively.

Scenario 2: Pay the top 1st and 2nd deciles of the quality distribution 3% and 1% more, respectively, and reduce the payments to the 9th and 10th deciles by 2% and 4%,

respectively. The asymmetry in payments between the high and low quality hospitals is necessary to induce budget neutrality.

Scenario 3: The change in Medicare inpatient revenue is given by $\Delta R_i = .1 \cdot R_i \cdot (Q_i - Q_{median})$ where Q_{Median} is the median quality score of the other hospitals. This exchange function rewards all hospitals with above median quality and penalizes all hospitals with below median quality. The size of the payment increases/decreases with an increase/decrease in quality over the entire quality distribution.

Scenario 4: Assume the CAHs participated in the CMS/Premier Hospital Quality Incentive Demonstration (HQID). In HQID the payment structure was identical to Scenario 1. The difference between Scenario 4 and 1 is that in Scenario 4 the comparison hospitals are all of those hospitals that participated in HQID whereas the set of comparison hospitals for Scenario 1 are CAH institutions.

Our calculations of the change in total revenue are based on our estimate of total Medicare inpatient revenue. Most P4P initiatives tie performance for a given condition or procedure to revenue for that procedure. Reliable quality of care metrics do not currently exist for most diagnoses. Without reliable quality metrics it is surely unwise, if not impossible, to use P4P incentives for those diagnoses. For the foreseeable future, most diagnoses will not likely be subject to P4P incentives. For this reason, our estimates will likely be an upper bound estimate of the impact of a P4P incentive system on the change in hospital revenue as our approach puts more of the hospital revenue at risk than most proposed P4P schemes.

DATA

Our data come from three primary sources: CMS's public reporting system for hospitals, Hospital Compare; Hospital Cost Report Information System (HCRIS); and the Flex Monitoring Team's census of CAHs. We limit our analysis to CAHs that had converted by 2006. We use data from the Flex Monitoring Team to identify CAHs and the dates of their conversion. From the Hospital Compare data we capture the recorded quality information for each CAH including the denominator information (the number of patients who should have received the recommended care) and the numerator (the number that actually received the recommended care). These data include information from 2003 to 2007.

Our quality performance measures are composite quality scores. Within the condition domains, Hospital Compare tracks the total number of patients that should have received a specific treatment (e.g. administering pneumonia vaccine) and the total number of patients that actually received that care. Composite quality scores are constructed by summing the total number of appropriate treatments for all the domains within a category and dividing that sum by the sum of the denominators across all the domains. HQID used the same approach to constructing performance measures.

Financial performance information was culled from the HCRIS data. These data were extracted and cleaned by the Flex Monitoring Team. We extracted revenue, cost and other hospital characteristic information from the Cost Reports and merged it with the Hospital Compare data. This step was somewhat challenging as hospitals have differing and changing reporting periods. We constructed annualized financial information for the calendar year. For hospitals that did not report on a calendar year basis we constructed weighted averages of the financial metrics where the weights are given by the overlap with the calendar year. For

example, if a hospital's reporting year was from July 1st to June 30th, then to construct the financials for 2006 to match to the Hospital Compare data we average the reports for 2006 and 2007. In our analysis, the primary measure of financial performance is net revenue.

Our initial analysis focuses on the initial three major conditions considered by Hospital Compare: acute myocardial infarction (AMI), Heart Failure (HF) and pneumonia. There are a number of challenges in assembling performance measures using the Hospital Compare data. Primary among these is that for many conditions there are a large number of missing values for the numerator and the denominator for the measures within each domain. For example, of the 1,212 CAHs in our data for 2006, 821 report complete quality information for pneumonia. For HF, 700 hospitals reported data sufficient to compute the composite score, while for acute myocardial infarction (AMI) only 286 CAHs reported quality information for all domains. We drop AMI from our analysis since the vast majority of hospitals did not report AMI values for the AMI quality domains.

There are two basic approaches for dealing with missing values. The first is to exclude hospitals with missing values. This approach greatly reduces the sample size for all conditions but pneumonia. The second approach assumes that missing values imply that no patient qualified for treatment within that domain and that a zero should be in place of a missing value. Our analysis uses the second approach. However, we performed the analysis making both assumptions and the results are insensitive to this choice. We also limit our sample to hospitals that treated at least 25 patients with either pneumonia or HF which is the criteria CMS uses for quality reporting.

For pneumonia, the quality measures include: 1) percentage of patients who were assessed for oxygenation; 2) percentage of patients who were given initial antibiotics within 4

hours after arrival; and 3) percentage of patients who were assessed and given pneumococcal vaccination. For heart failure, the measures include: 1) percentage of patients who were assessed for left ventricular function; and 2) percentage of patients who were given an ACE inhibitor or ARB for left ventricular systolic dysfunction.

Our work assesses the impact of potential Medicare P4P initiatives for inpatient care. To perform our simulations we need a measure of hospital inpatient revenue from Medicare. The HCRIS data contain an explicit measure of inpatient Medicare revenue, however, this value is missing for a large number of hospitals. Therefore, we construct an estimate of Medicare inpatient revenue by multiplying total inpatient revenue, which is available in the HCRIS data, by the percent of inpatient days that result from Medicare enrollees.

RESULTS

Table 1 presents some summary statistics for the initial sample providing a snapshot of the data prior to construction of the analysis sample. As mentioned above, the most striking feature of these data is the number of missing pneumonia and HF composite quality score observations. Since we are filling in missing values with zeros, missing composite scores are a consequence of the hospital not reporting any values for the individual measures within the composite score. This suggests that sample selection bias may be a concern, however without more information it is difficult to assess the importance of selection on our findings.

Table 1: Summary Statistics, Raw Data

Variable	N	Mean (Std dev)
Pneumonia Composite Quality Score	2,602	.84 (.089)
HF Composite Quality Score	1,388	.70 (.21)
Net Revenue	4,663	524,371 (1,482,935)
Operating Margin	4,663	.016 (.059)
Percentage Medicare Admissions	4,658	66.3 (17.3)

This concern is borne out in Table 2 which presents the summary statistics for our analysis samples – the panel of hospitals for which we have complete financial data and at least 25 patients that were included in the pneumonia or HF composite quality scores. Comparing the mean net revenue for the analysis sample to the starting sample reveals that net revenue is much higher for the analysis sample. Hospitals reporting pneumonia quality information have, on average, better financial performance than non-reporters.

Table 2: Summary Statistics, Analysis Data

Variable	N	Mean (Std dev)
Pneumonia Composite Quality Score	2,028	.85 (.084)
Heart Failure Composite Quality Score	1,368	.70 (.21)
Net Revenue ($)	2,028	833,719 (1,768,619)
Operating Margin	2,028	.022 (.048)
Percentage Medicare Admissions	2,028	66.8 (14.4)

This is a limitation of our analysis. We are not able to calculate the financial impact of P4P initiatives for those hospitals that did not submit any information to Hospital Compare or

did not treat enough patients to meet our cutoff criteria. For pneumonia, this is over 50% of hospitals. Non-reporting hospitals are not as profitable as reporting hospitals and thus are more likely to be pushed into financial distress by a P4P program. The degree to which this limitation impacts our conclusions depends on the relationship between the distribution of hospital quality and the distribution of financial performance. Below we discuss this relationship and also discuss the implications for our conclusions.

On average, hospitals in our sample are earning positive net revenue. Across all years, the average hospital net revenue is almost $834,000 with an operating margin of more than 2%. Not surprisingly, these hospitals treat a large share of Medicare enrollees—on average 67% of inpatient days are patients enrolled in Medicare.

Figures 1 and 2 present histograms of composite pneumonia and HF quality scores, respectively. For both quality measures, there is meaningful variation across CAHs. The pneumonia scores are comparable to the HQID results where the mean composite score was 86.4 (4^{th} quarter, 2005) – similar to the mean of 86.6% for 2005 and 2006 in our sample (Lindenauer, et al., 2007). However, CAHs do not perform as well for HF. The mean HF score in HQID was 88.0 in the third quarter of 2005 compared to the mean in our sample of .70.

Figure 1: Histogram of Pneumonia Quality Composite Measure

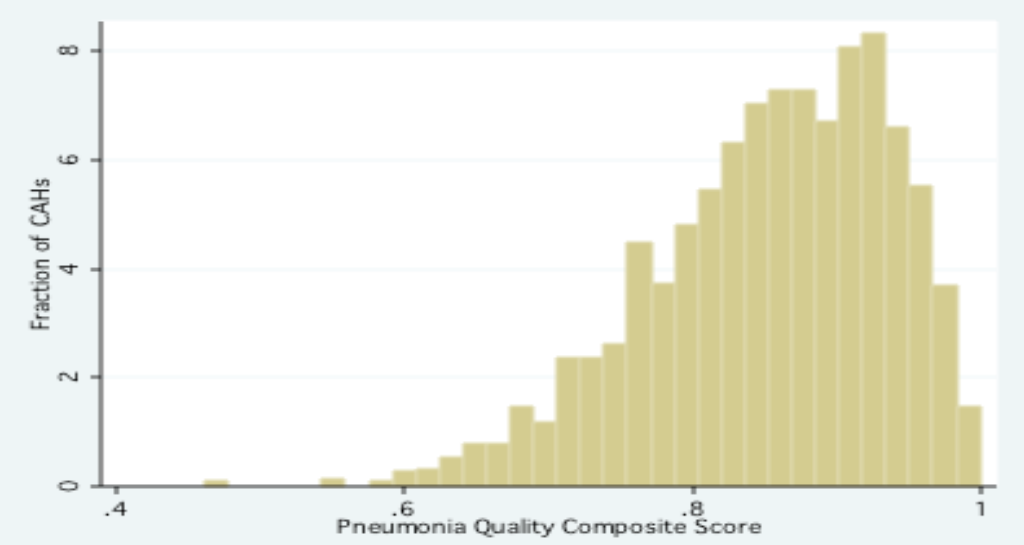
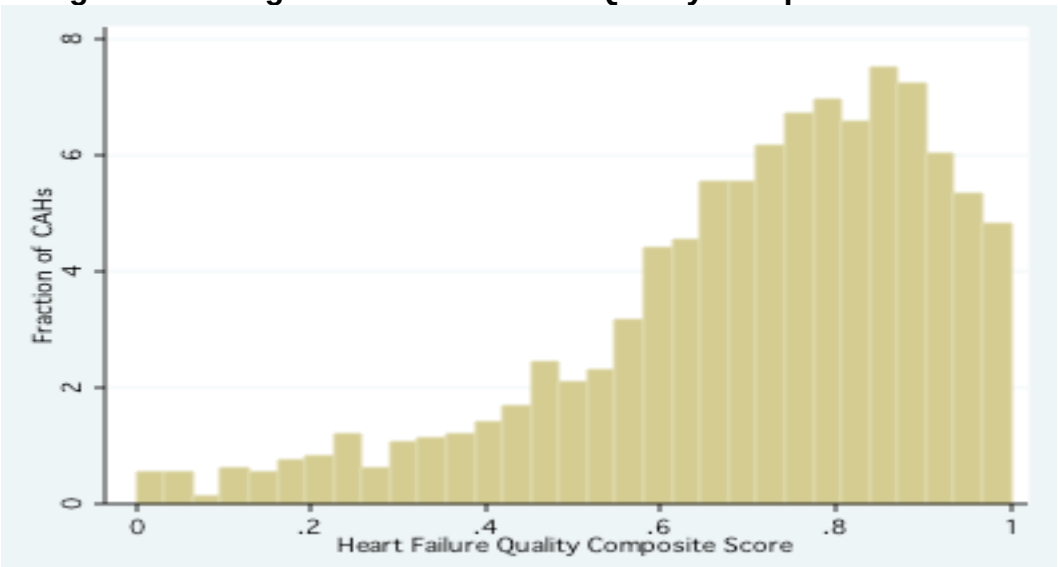


Figure 2: Histogram of Heart Failure Quality Composite Measure



A scatter plot of pneumonia and HF composite process scores against net revenue in 2006 is presented in Figures 3 and 4, respectively. We also graph the fit from a polynomial regression of net revenue on the composite quality scores and find there is a positive correlation between financial well-being and both composite quality scores. This suggests that P4P schemes may benefit hospitals in better financial condition at the expense of those hospitals that face resource

challenges. However, it is also clear from Figures 3 and 4 that the fit of this relationship is, at best, modest. This is confirmed in our regression analysis.

Figure 3: Relationship between Pneumonia Composite Quality Scores and Net Revenue/100,000 in 2006

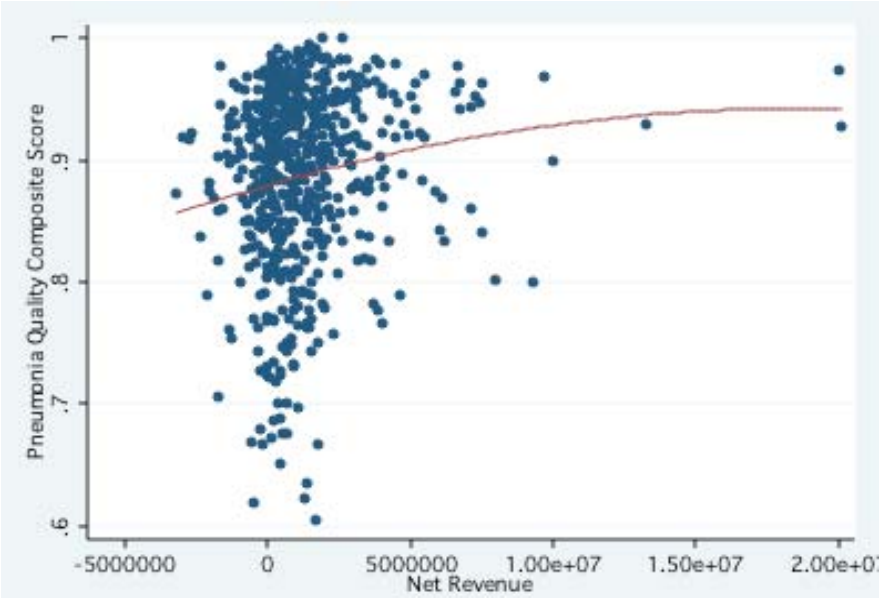


Figure 4: Relationship between Heart Failure Composite Quality Scores and Net Revenue/100,000 in 2006

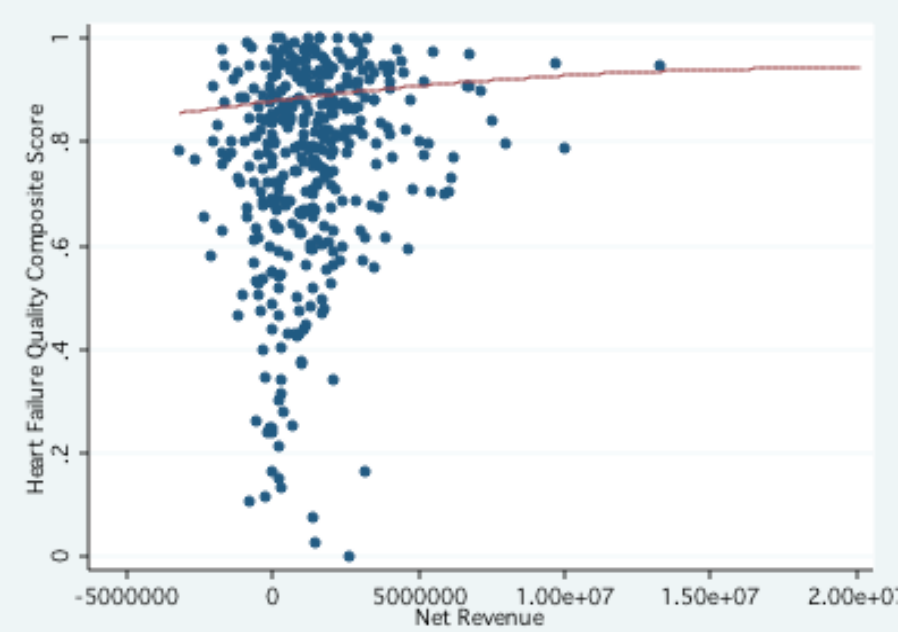


Table 3 presents the results of a random effects, linear regression analysis of the relationship between CAH net revenue and the two composite quality scores. In these analyses each hospital receives an independent and identically distributed, time invariant random error. The econometric evidence is consistent with the visual evidence in Figures 3 and 4. There is a positive and statistically significant relationship between net revenue and the composite quality scores. However, the quality of the fit is limited. Thus, on average, higher net revenue is associated with higher hospital quality performance, and P4P schemes can adversely impact the disparity in hospital financial performance. However, this relationship is weak as net revenue can only explain a fraction of the variation in hospital composite quality.

Table 3: Random Effects Estimates of the Relationship between Pneumonia and HF Composite Quality Scores and Net Revenue

	Pneumonia		Heart Failure	
Variable	**(1)**	**(2)**	**(3)**	**(4)**
Constant	.80 (.0036)	.79 (.010)	.64 (.010)	.63 (.010)
Net Revenue / 100,000	.00026** (.000084)	.00026** (.000097)	.0071** (.0023)	.00061** (.000097)
Number of Patients in Condition	---	.0019** (.000061)	---	.00016 (.00027)
Inpatient Days/100,000	---	-.32 (.98)	---	5.31* (2.52)
Percent Inpatient Days from Medicare	---	-.0044 (.015)	---	-.049 (.045)
N	2,028	2,028	1,444	1,444
R^2	.13	.13	.052	.064

Dependent Variable is the Pneumonia Composite Quality Score for (1) and (2) and the Heart Failure Composite Quality Score for (3) and (4). Annual indicator variables included but not reported.

**Significant at the 1% level of confidence.

*Significant at the 5% level of confidence.

It is possible that the relationship between net revenue and hospital quality is only relevant for those hospitals in the middle of the quality distribution. If so, many P4P initiatives would have little impact on financially struggling hospitals. In order to assess the impact of net revenue on the likelihood of very high and very low quality performance we estimate a series of probit regressions where the dependent variable is an indicator for whether the quality observation lies in a given decile of the quality distribution. We estimate four regressions for the 1st, top 2 deciles, bottom 2 deciles and the 10th decile of the pneumonia composite score distribution (Table 4). For the remainder of the analysis we focus solely on pneumonia as the results for HF are very similar and we have more pneumonia observations for the analysis.

Table 4: Probit Estimates of the Relationship between Pneumonia Quality Scores and Net Revenue

	Dependent Variable			
Variable	**1st Decile**	**1st or 2nd Deciles**	**9th or 10th Deciles**	**10th Decile**
Net Revenue / 100,000	.070*	.075**	-.070**	-.073**
	(.023)	(.021)	(.023)	(.028)
2004	.0055	-.012	.026	.036
	(.10)	(.083)	(.084)	(.10)
2005	-.030	-.046	.056	.066
	(.11)	(.098)	(.10)	(.13)
2006	-.064	-.095	.072	.082
	(.12)	(.10)	(.10)	(.13)
Constant	-1.33	-.87	-.84	-1.29
	(.10)	(.084)	(.074)	(.10)
N	2,015	2,015	2,028	2,028
Log Likelihood	-643.6	-995.1	-999.6	-643.2

Dependent Variable is the Pneumonia Composite Quality Score
*Significant at the 5% level of confidence.
**Significant at the 1% level of confidence.

The analysis is consistent with net revenue affecting quality throughout the entire quality distribution. The likelihood that a hospital has a quality score in the 10th decile is negatively related to net revenue while the probability that a hospital's quality score lies in the top decile is increasing in net revenue. Thus, P4P initiatives do have the potential to make poor hospitals poorer and rich hospitals richer.

Simulation Results

The results of our simulations based on the pneumonia composite quality score are presented in Table 5.[1] The first column presents the data on the distribution of net revenue in our sample for 2006 prior to the simulation analysis. We focus on 2006 because it is several years after the initiation of Hospital Compare and hospital behavior with respect to quality is more

[1] We performed similar analysis using HF composite scores as the basis of the P4P and the results are similar to those using pneumonia.

likely to be closer to equilibrium as hospitals become more familiar with quality reporting. Column (2) presents the results for Scenario 1 that corresponds to the exchange function used in HQID. The resulting distribution of net revenue under P4P is not meaningfully different than without P4P. Under Scenario 1, the net revenue of 21.6% of hospitals is less than zero while 20.9% of hospitals earned negative net revenue without any P4P incentives. The percentage of hospitals with negative net revenue increased but the change is modest.

Table 5: Impact of Pay-for-Performance Incentives on the Distribution of Net Revenue for CAHs for 2006

	No P4P (1)	Scenario 1 (2)	Scenario 2 (3)	Scenario 3 (4)	Scenario 4 (5)
Mean	1,289,037	1,294,459	1,299,880	1,281,863	1,300,289
Median	887,068	872,982	871,391	870,219	865,172
Standard Deviation	2,169,380	2,207,632	2,250,165	2,195,587	2,212,760
Percent with Net Revenue < 0	20.9	21.6	21.9	21.6	22.1
Mean conditional on Net Revenue < 0	-682,691	-695,573	-708,454	-699,767	-692,904
Mean decrease in net revenue conditional on old Net Revenue < 0	---	-12,881	-25,762	-17,075	-10,213

Scenarios described in text. All amounts in 2006 dollars.

In Column (3) we present the results from Scenario 2 where the payoffs and penalties from the P4P incentives are twice as high as in Scenario 1. Again, the percentage of hospitals earning negative net revenue is approximately the same as the original, no P4P distribution. In this simulation the number of hospitals earning negative net revenue increases to 21.9%, a one percentage point increase over the original P4P distribution. Again, this is a modest increase in the disparity of net revenue across hospitals.

Column (4) presents the results from the more continuous exchange function of Scenario 3. The change in the number of hospitals earning negative net revenue is similar to Scenario 1 but less than under Scenario 2. Unlike the incentives provided in Scenarios 1 and 2, this

exchange function provides incentives, albeit modest, for all hospitals to improve their quality independent of their location in the quality distribution.

The last row of Table 5 presents the change in revenue for those hospitals whose net revenue was negative prior to any P4P adjustments. On average, the P4P incentives reduce the profitability of these already poor performing hospitals. However, the reduction in net revenue is, on average, small suggesting the P4P incentives will have only a modest negative impact on the sustainability of CAHs. For example, in Scenario 2 (the Scenario that generates the largest decline in net revenue), the decline in net revenue from the P4P plan is less than 3% of the median Medicare inpatient revenue.

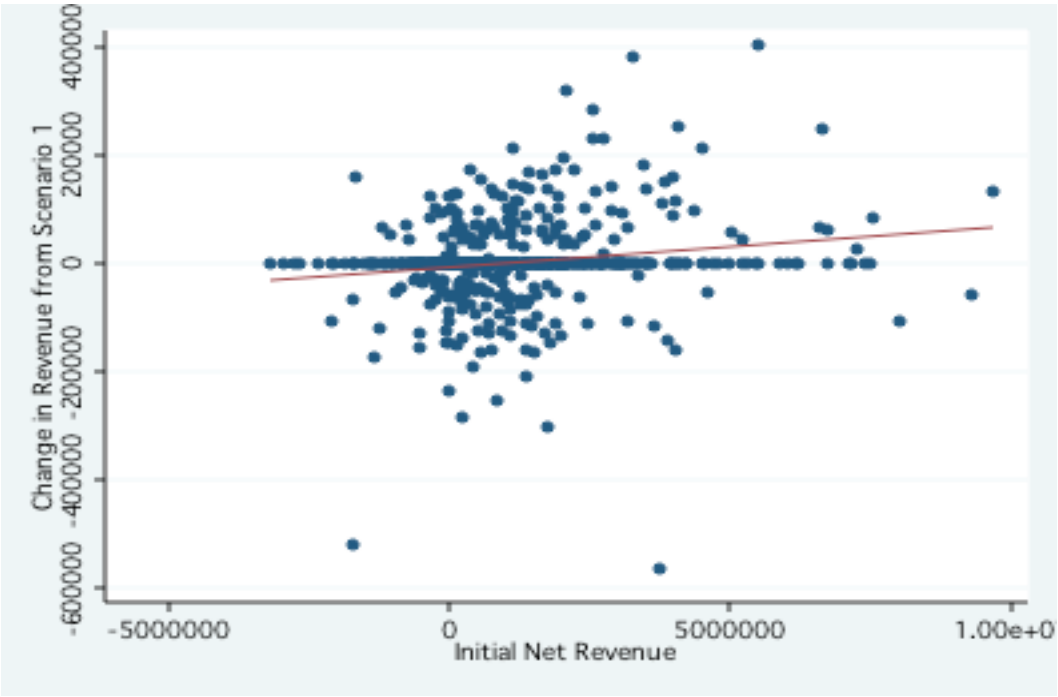
Our analysis has only compared CAHs to other CAHs. It is of interest to know how CAHs would fair if their P4P reimbursements were determined by their performance relative to all hospitals. Column (5) presents the results from this analysis. CAHs experience an increase in average net revenue as their performance for pneumonia is, on average, slightly better than those hospitals that participated in HQID. The percentage of hospitals earning negative net revenue in this simulation is similar to hospitals under Scenario 3 suggesting that, at least for pneumonia, including CAHs in P4P aimed at all hospitals would not place many CAHs in greater financial jeopardy. It is important to note that pneumonia may be the only quality condition for which this is true. As noted above, CAHs' HF composite measures were notably worse than for other hospitals in HQID.

To get a sense of the magnitude of the incentives necessary to cause a meaningful increase in the number of hospitals that are in financial difficulty, we explored the implications of a number of different P4P incentive structures. To increase the number of hospitals with negative net revenue by two percentage points, the P4P incentives would have to increase so that

those hospitals in the bottom decile would lose 10% of their Medicare inpatient revenue. This is an extremely aggressive incentive system and one that increases the hospital's revenue at risk five times the incentive system used in HQID.

An obvious question to ask is given the correlation between net revenue and pneumonia and HF quality metrics reported above, why does P4P have little impact on financially poorly performing hospitals? Figure 5 attempts to provide an answer to that question. In this figure we plot the change in net revenue under Scenario 1 against initial net revenue. While there is a correlation between the change in revenue from the P4P initiative and initial net revenue, the correlation is weak. This is because the variation in net revenue accounts for a minority of the variation in pneumonia composite quality scores.

Figure 5: Plot of the Change in Revenue from Scenario 1 and Initial Net Revenue for 2006



DISCUSSION

An obvious concern with P4P is that it may negatively affect the financial stability of hospitals that are in a precarious position (Werner and Asch, 2005). CAHs are prime candidates for P4P programs to have such unintended consequences. However, our work suggests that P4P incentives are likely to have, at best, only a modest impact on the financial stability of the CAHs that are already under significant financial pressure. Thus, if P4P programs are able to induce hospitals to increase quality, those benefits need not be weighed against the risk of putting already financially distressed CAHs in greater financial jeopardy.

There are a number of limitations to our analysis. First, data limitations prevent us from modeling the impact of P4P on composite quality scores for conditions other than pneumonia and HF. Our conclusions may be affected if we were able to examine a broader array of conditions.

Second, we are not able to model heterogeneity in hospital responses to P4P incentives. That is, we assume that hospital responses to P4P do not vary by financial status. This may not be a reasonable assumption, however, we do not possess hospital level data on responses to P4P incentives and thus cannot assess hospital quality responses to initial net revenue. Werner et al. (2008) make a similar assumption.

Third, we model the impact of P4P incentives for pneumonia and HF on total inpatient Medicare revenue. This likely overstates the impact of P4P on the financial health of hospitals as these diagnoses account for just a fraction of Medicare admissions. Most P4P systems link the revenue for a specific Diagnostic Related Group (DRG) to the quality scores for that DRG. Our simulation linked the outcomes from the pneumonia composite scores to the entire Medicare inpatient revenue.

In summary, while P4P incentives disproportionately reduce the revenue of hospitals that are more likely to be in financial distress, the average impact appears to be modest. If P4P programs can lead to improved hospital quality (as suggested by HQID [Lindenauer et al., 2007]), these benefits should not be accomplished at the expense of putting CAHs in financial jeopardy

As discussed in the IOM report, *Quality Through Collaboration: The Future of Rural Health* (2005), rural hospitals need to participate in demonstration projects aimed at improving the quality of care in these institutions. Our results suggest that CAHs should be included in future P4P initiatives so we can better understand how payment incentives affect the quality of care in small rural hospitals.

REFERENCES

Flex Monitoring Team. (2009, January 26). *Critical Access Hospital List.* [Online]. http://flexmonitoring.org/cahlistRA.cgi. [Accessed 2009, February 5].

Institute of Medicine. Committee on the Future of Rural Health Care Quality. Board on Health Care Services. (2005). *Quality through collaboration: The future of rural health care*. Washington, DC: National Academies Press.

Kohn, L.T., Corrigan, J.M., & Donaldson, M.S. (Eds.). (2000). *To err is human: Building a safer health system.* Washington, DC: National Academy Press.

Lindenauer, P.K., Remus, D., Roman, S., Rothberg, M.B., Benjamin, E.M., Ma, A. et al. (2007/2/1). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine, 356*(5), 486-496.

Werner, R.M., & Asch, D.A. (2005, March 9). The unintended consequences of publicly reporting quality information. *JAMA, 293*(10), 1239-1244.

Werner, R.M., Goldman, L.E., & Dudley, R.A. (2008, May 14). Comparison of change in quality of care between safety-net and non-safety-net hospitals. *JAMA, 299*(18), 2180-2187.