JOURNAL OF HEALTH CARE FINANCE

Winter 2017

Linking Reimbursement to Performance in Acute Care Hospitals: Lessons from Maryland's Implementation Experience

Steven Garfinkel, PhD
Institute Fellow, American Institutes for Research

Ying Wang, PhD
Principal Researcher, American Institutes for Research (AIR)

Yi Lu, PhD
Economics Researcher, American Institutes for Research

Marilyn Moon
PhD, Senior Vice President, American Institutes for Research

Brenna Raines, MPP
Senior Manager, Plasma Protein Therapeutics Association

Christian Evensen, MS
Principal Researcher, American Institutes for Research

Acknowledgements: The authors wish to thank our funder, the Robert Wood Johnson Foundation (Grant No. 65463) and the staff members of the following organizations who supplied secondary data or participated in interviews: the Maryland Health Services Cost Review Commission, the Maryland Health Care Commission, the participating Maryland hospitals, the Maryland Hospital Association, and the Maryland Medical Assistance Program. The authors are entirely responsible for errors. This study was approved by American Institutes for Research's Institutional Review Board (IRB00000436).

Abstract

Background: In 2009, Maryland introduced two pay-for-performance (P4P) initiatives in its unique all-payer rate setting system; one using hospital acquired conditions (MHAC) and the other clinical process of care and HCAHPS measures (QBR).

Purpose: This study summarizes interviews with Maryland hospital managers and payers about their early program implementation experiences and implications for Medicare Value Based Purchasing.

Methodology/Approach: Rigorous qualitative analysis sampling, data collection, and analysis.

Results: The design of the QBR program—withholding a small percentage of each Diagnosis-Related Group payment and redistributing the pool based on performance—presented little threat to the hospitals, because the amount at risk is small and predictable. However, the initially proposed design on the MHAC program—refusal to pay for the marginal cost of 11 hospital acquired conditions—was extremely threatening, because the amount of revenue at risk would have been unpredictable and large. As implemented, though, the MHAC program used a withholding pool approach similar to the QBR and was much more acceptable to hospitals, even though the number of hospital acquired conditions subject to risk-based incentives was much larger—49, which increases the chances of incurring a penalty.

Conclusion: Hospitals were comfortable with the withholding pool approach because it is predictable, the amounts at risk were relatively low, and there is a potential benefit. However, refusal to reimburse for any costs associated with a single admission was threatening because of greater financial uncertainty and the focus on penalty without benefit, even when the trade-off means increasing the likelihood that a penalty might be incurred.

Practice Implications: The more acceptable withholding-pool incentive structure actually exposes hospitals to increased likelihood that a penalty will be incurred, but the alternative is much more threatening because it is less predictable, potential losses are high, and it offers no benefit for good performance.

Introduction

In 2008 and 2009, the state of Maryland modified its "all-payer" reimbursement system for acute inpatient hospital care by implementing the Quality Based Reimbursement (QBR) and Maryland Hospital Acquired Conditions (MHAC) programs. Both the QBR and MHAC programs intend to reward hospitals for better quality care, a rapidly growing trend in the United States.[1,2] The implementation of the QBR and MHAC programs in Maryland holds potentially valuable lessons for the similar programs currently in use nationwide as part of the Medicare value-based purchasing quality initiatives. This article compares and contrasts Maryland and CMS pay-for-performance (P4P) programs and draws lessons from the Maryland program implementation for CMS and other payers.

Maryland's experience with these two P4P programs[3] is important for acute inpatient care providers nationwide for two reasons. First, the Maryland QBR and MHAC programs are similar to the CMS Hospital VBP (76 Fed. Reg. 26490, 2011)[4] and Hospital Acquired Conditions (HAC) Reduction Programs.[5] Although not identical to the CMS programs, the Maryland system uses similar measures (i.e., preventable hospital acquired conditions) and revenue adjustment approaches (i.e., revenue withholding and redistribution based on performance) and can provide useful lessons for implementing CMS's national programs.

Second, understanding how hospital P4P programs work in Maryland provides unique and valuable lessons for CMS and for other states interested in establishing a "true" all-payer system. Maryland is the only state operating under a federal waiver that enables it to include Medicare beneficiaries in its all-payer rate-setting system. An independent agency, the Maryland Health Services Cost Review Commission (HSCRC) sets the reimbursement rates for *all* stays at all nonfederal acute care hospitals in Maryland. Therefore, the Maryland all-payer system eliminates hospitals' incentive to shift costs among various payers, including Medicare, and the mitigating impact of having a large proportion of patients (i.e., Medicare beneficiaries) remain outside the all-payer system. The Maryland experience provides insightful lessons for CMS as it tests and expands its payment reform models to include multiple payers (both public and private) as a result of the Affordable Care Act (ACA), and for other states (e.g., Vermont) interested in establishing all-payer systems.

---

[1] Kahn, C. N., Ault, T., Potetz, L., Walke, T., Chambers, J. H., & Burch, S. (2015). Assessing Medicare's hospital pay-for-performance programs and whether they are achieving their goals. Health Affairs, 34(8), 1281–1288.

[2] In 2014, Maryland began a new global budgeting system for acute care hospitals as a demonstration program for CMMI. However, the QBR and MHAC programs continue in conjunction with the global budgeting demonstration.

[3] Maryland also has implemented several readmission measures as part of its P4P program since 2011. At the time of our interview, these programs were not in place. Thus, our paper does not cover readmission measures. However, HCAHPS, clinical POC, and PPC measures make up the majority of measures in the CMS Hospital VBP program, so our findings from MHAC and QBR programs alone are still relevant to CMS's national programs.

[4] Medicare Program; Hospital Inpatient Value-Based Purchasing Program, 76 Fed. Reg. 26490 (July 5, 2011) (to be codified at 42 C.F.R. pts 422 & 480).

[5] Section 3008 of the Patient Protection and Affordable Care Act (ACA) established the Hospital-Acquired Condition (HAC) Reduction Program to reduce HACs in the hospitals. Payments to worst-performing hospitals in the program are reduced by 1% in terms of risk-adjusted HAC quality measures.

HSCRC staff has published information about the effectiveness of the QBR and MHAC programs based on hospital-level data,[6] but this article is the first to report the implementation experience of participating hospitals and payers using data from interviews with their staff members, conducted by independent investigators during the implementation period.

Background

Value-Based Purchasing. The QBR and MHAC programs are examples of P4P, one type of value-based purchasing (VBP).[7] VBP is an effort to "bend the cost curve" by reimbursing providers for care that produces value.[8] Value in this sense is defined as higher quality for a given cost or lower cost for a given level of quality. The Institute of Medicine defined quality as care that is safe, effective, patient-centered, timely, efficient, and equitable.[9] Value can be measured conceptually as "[favorable] health outcomes achieved per dollar spent."[10] The notion of value in health care is predicated on the belief that low-quality care is unacceptable and that high-quality care, as well as improvements in quality, ought to receive preferential financial rewards.

P4P pays directly for measurable quality or quality improvement using quality indicators. Quality indicators come from either administrative data or patient-reported data and can measure either process or outcome. Process measures capture provider behavior and are either reported in administrative records (e.g., insurance claims as in CMS's Hospital Inpatient Quality Reporting Program) or by asking patients about their experiences with providers (e.g., the Consumer Assessments of Healthcare Providers and Systems [CAHPS®] Hospital Survey, or HCAHPS). Outcome measures capture the impact of care on patients, such as readmissions and complications from administrative claims data and patient-reported outcome measures (PROMs). PROMs have not been widely used for quality improvement, but there is growing interest in using them, for instance, by systematically aggregating PROMs to the provider level.[11,12]

Maryland's QBR program rewards high performers and penalizes low performers based on both patient-reported HCAHPS measures and clinical process-of-care (POC) measures based on administrative data (e.g., pain management, aspirin on arrival for Acute Myocardial Condition [AMI] admission). The MHAC program does the same using outcome measures derived from

[6] Calikoglu, S., Murray, R., & Feeney, D. (2012). Hospital pay-for-performance programs in Maryland produced strong results, including reduced hospital-acquired conditions. Health Affairs, 31(12), 2649–2658.
[7] Damberg, C., Sorbero, M. E., Lovejoy, S. L., Martsolf, G. R., Raaen, L., & Mandel, D. (2014). Measuring success in health care value-based purchasing programs: Summary and recommendations. Retrieved from http://www.rand.org/pubs/research_reports/RR306.html
[8] Cutler, D. (2010). Analysis & commentary: How health care reform must bend the cost curve. Health Affairs, 29(6), 1131–1135.
[9] Institute of Medicine. (2001). Crossing the quality chasm: A new health system for the 21st century. Washington, DC: National Academy of Sciences.
[10] Porter, M. E. (2010). What is value in health care? New England Journal of Medicine, 363, 2477–2481.
[11] Black, N. (2013). Patient reported outcome measures could help transform healthcare. BMJ, 346, f167.
[12] Hostetter, M., & Klein, S. (2012). Using patient-reported outcomes to improve health care quality. Quality Matters, Dec 2011/Jan 2012. Retrieved from http://www.commonwealthfund.org/publications/newsletters/quality-matters/2011/december-january-2012/in-focus

administrative data (i.e., potentially preventable complications [PPCs], such as urinary tract infection not present on admission or collapsed lungs).[13]

How the QBR and MHAC Programs Work. Reimbursement for acute inpatient stays in Maryland is based on the All Patient Refined Diagnosis-Related Group (APR-DRG) classification, a product of 3M Health Information Systems, which expands the Medicare DRG classification with groups that are relevant to the non-Medicare population and creates subcategories based on severity.[14] The QBR and MHAC programs modify the original reimbursement by withholding a small percentage (0.5% at the start of each program) of each hospital's APR-DRG-based reimbursement, pooling the withheld amounts, and redistributing the pool based on performance over a 12-month reference period. Thus, the system is budget neutral: high performers win and low performers lose. Each hospital's share of the withheld funds is paid out as adjustments to its APR-DRG rates for the subsequent year, rather than as a lump sum payment at the end of the focal performance year. The percentage withheld for the QBR withholding pool has remained steady at 0.5%,[15,16] but the MHAC percentage has risen from 0.5% at the start to a maximum of 4.0% by 2016.[17] A detailed description of the system and changes over time for both QBR and MHAC is available on the HSCRC website.[18,19] Although the withheld percentages used by CMS and Maryland differ, their budget-neutral redistribution approaches are similar.

For fiscal year (FY) 2015, the QBR payment algorithm produced a score for each of the 43[20] acute care hospitals based on 10 CMS clinical POC measures for heart attack, heart failure, pneumonia, and surgical infection indicators, 8 patient experience measures from the HCAHPS,

[13] In 2014, CMS and the state of Maryland revised their Medicare waiver for acute inpatient care by adding global budgeting to the QBR and MHAC programs, replacing the conventional DRG-based payment per admission. Our findings precede the shift from per-admission reimbursement to global budgets.

[14] 3M Health Information Systems. (2003). All patient refined diagnosis related groups (APR-DRGs) methodology overview, version 27.0. Wallingford, CT: Author.

[15] In October 2015, HSCRC approved allocating 2% of hospital-approved inpatient revenue for QBR performance in FY 2018.

[16] Health Services Cost Review Commission. (2016). Maryland quality based reimbursement program measure standards, scaling determination, and other methodology changes for FY 2018. Baltimore, MD: Author. Retrieved from http://hscrc.state.md.us/documents/HSCRC_PolicyDocumentsReports/PolicyClarification/2016-02-19-QBR-MMO-to-Industry-FY-2018-Final.pdf].

[17] Health Services Cost Review Commission. (2015). Maryland hospital acquired condition (MHAC) program scaling for RY2017. Baltimore, MD: Author. Retrieved from http://www.hscrc.maryland.gov/documents/HSCRC_PolicyDocumentsReports/PolicyClarification/2015/MHAC-FY-2017-SCALING-Memo-04-02-2015.pdf

[18] Quality Based Reimbursement (QBR) program: http://hscrc.state.md.us/init_qi_qbr.cfm; Maryland Hospital Acquired Conditions program: http://hscrc.state.md.us/init_qi_MHAC.cfm

[19] Health Services Cost Review Commission. (2009). Final staff recommendations regarding HSCRC payment policy for highly preventable hospital acquired conditions. Baltimore, MD: Author. Retrieved from http://hscrc.state.md.us/documents/HSCRC_Initiatives/QualityImprovement/MHAC/MHAC_FinalStaffRecommendations-6-3-09.pdf

[20] McCready and Kernan have been excluded because of insufficient number of surveys and cases for both the clinical and HCAHPS measures in the base period or performance period. St. Joseph's has been excluded because it had only six months of data for the performance period.

and 1 mortality measure. Hospitals receive credit for meeting targets on specific clinical process indicators and for meeting all measures targets relevant to a patient's particular condition.[21]

During our study period (2009–2011), the MHAC payment algorithm scored each hospital on the difference between the observed and expected number of PPCs that patients experienced, defined by the method developer, 3M Health Information Systems. A PPC is a function of the relationship between the primary diagnosis, the secondary diagnoses, and whether the secondary diagnoses were coded as present on admission.[22] The expected experience of PPCs for each hospital is the statewide average PPC rate adjusted for the hospital's APR-DRG and severity of illness (SOI) patient mix. In 2009, the MHAC program adopted a list of 64 PPCs[23] and annually selected a subset of effective PPCs to be measured for performance.[24] The hospitals are scaled according to their marginal resource usage associated with greater- or fewer-than-expected PPCs. Marginal resource use is the difference between the average cost of a case in the same APR-DRG/SOI category with and without a PPC, estimated from a previous base year. The MHAC score is the combined marginal resources use (negative or positive) summed over each of the assigned PPCs. The withheld revenue pool is distributed to those best-performing hospitals (i.e., with low complication rates) based on the MHAC total score.

The MHAC program was originally designed to eliminate the marginal reimbursement associated with the additional costs that the PPC adds for each relevant discharge; this practice is essentially the HAC-based reimbursement policy, known as the "Hospital-Acquired Conditions Present on Admission Indicator (HAC POA)" program, which CMS began in 2008 and continues

[21] The hospitals receive credit on the clinical process indicators for meeting the goal on each measure relevant to the patient's condition (the "opportunity" score, which is the percentage of all process measure targets applicable to that patient that are achieved) and meeting the target for all measures relevant to the patient's condition (the "appropriateness" score, which is dichotomous—either all the targets were met for that patient or not). Opportunity accounts for 25% of the clinical process score, and appropriateness accounts for 75%. Each of the 21 clinical process measures are weighted equally in the clinical process score and each of the 8 patient experience domains are weighted equally in the patient experience score, but the clinical process score accounts for 70% of the total QBR score and the patient experience score for 30%.
The clinical process and patient experience measures are scaled according to both attainment and improvement, and the hospital receives the higher of the two scores. Attainment reflects the hospital's performance relative to all other hospitals during the current period. Improvement reflects the hospital's current performance relative to its own best performance in previous measurement periods, which is called the Baseline. For each measure, hospitals receive 1 attainment point if they exceed the 50th percentile and all 10 possible attainment points if they exceed the mean of the top decile, which is called the Benchmark. They receive from 0 to 9 improvement points per measure based on the ratio of their improvement (current minus Baseline) to the difference between the Benchmark and their Baseline score). Hospitals can receive up to 20 additional points that reflect the degree of consistency among the 8 patient experience measures.

[22] 3M Health Information Systems. (2013). 3M™ potentially preventable complications (PPCs) methodology overview, version 31.0. Wallingford, CT: Author. Retrieved from http://www.hhsc.state.tx.us/hhsc_projects/ECI/docs/PPC_methodology_overview.pdf

[23] The list is expanded to 65 PPCs by dividing PPC 22 (urinary tract infection) into PPC 65 (urinary tract infection without catheter) and PPC 66 (catheter-related urinary tract infection) in calendar year 2012.

[24] Excluded PPCs usually have low or insignificant marginal cost, clinical consistency problems, or coding accuracy concerns, as stated in HSCRC's memorandum "Re: Maryland Hospital Acquired Conditions (MHAC) Data and Reports for Full FY 2009", dated on December 4th, 2009 and available at http://hscrc.state.md.us/documents/HSCRC_Initiatives/QualityImprovement/MHAC/MemoFull_FY09_%20MHAC_DataReports9-12-04%20.pdf.

to use (73 Fed. Reg. 48434, 2008, p. 48471[25]). The MHAC included in its approach 11 PPCs that were considered similar to CMS's HACs and highly preventable. During the design process, stakeholders objected to the individual case-specific penalty and the use of so few PPCs, so the program was redesigned to withhold a fixed percentage of total reimbursement from all hospitals (i.e., not tied to the occurrence of a PPC) and reallocate it to hospitals that performed well on its overall PPC measure; this aggregate redistribution approach parallels the one used for the QBR program.

How the Maryland QBR and MHAC Programs Compare with CMS's Hospital P4P Programs. Since CMS established the Hospital Compare website in 2008 to provide consumers with information on how well hospitals provide recommended care to their patients, CMS has had four major P4P programs for acute inpatient care that compare with Maryland's programs: (1) the HAC POA Indicator program, which eliminates payment for the marginal cost of HACs in 14 HAC categories, beginning in 2008; (2) the HAC Reduction Program, which reduces the hospital's total Inpatient Prospective Payment System (IPPS) payments by 1% for hospitals in the lowest performing quartile of HAC scores, which began in FY 2015; (3) the Hospital VBP program, which withholds a small percentage of IPPS reimbursement and subsequently redistributes the withholding pool to hospitals based on clinical POC and HCAHPS measures and several mortality and efficiency measures, beginning in 2012; and (4) the Hospital Readmissions Reductions Program, which reduces IPPS payments to hospitals with excess readmissions, also starting in 2012.[26]

Maryland's QBR is analogous in concept, measure selection, and methodology to the CMS Hospital VBP program, although the percentages withheld and the redistribution algorithms differ. As noted previously, QBR withholds 0.5% and is based on 10 clinical POC measures, 8 HCAHPS measures, and 1 mortality measure. The CMS Hospital VBP program currently withholds 2.0%. In FY 2017, the CMS VBP program is based on 6 clinical care (process and outcome), 8 HCAHPS, 5 healthcare associated infection (HAI), 1 AHRQ Composite, and 1 efficiency measures.

Differences between the MHAC and CMS HAC POA programs were more substantial prior to 2010 but have become minimal since the HAC Reduction program was implemented in 2014. The CMS HAC POA program penalizes hospitals on a case-by-case basis by not paying a higher rate for excessive cost associated with HACs (HAC POA), while Maryland shifted a small portion of all reimbursement from low performers to high performers through a withholding pool based on overall annual performance. In 2014, the CMS HAC Reduction Program adopted an aggregate scoring approach similar to Maryland's. This approach also makes the implementation lessons from the MHAC program more relevant for the newer CMS HAC Reduction program.

[25] Medicare Program; Changes to the Hospital Inpatient Prospective Payment Systems and Fiscal Year 2009 Rates; Payments for Graduate Medical Education in Certain Emergency Situations; Changes to Disclosure of Physician Ownership in Hospitals and Physician Self-Referral Rules; Updates to the Long-Term Care Prospective Payment System; Updates to Certain IPPS-Excluded Hospitals; and Collection of Information Regarding Financial Relationships, 73 Fed. Reg. 48434 (August 19, 2008) (to be codified at 42 C.F.R. pts 411, 412, 413, 422, & 489).

[26] Centers for Medicare & Medicaid Services. (2016, February 8). Quality initiatives—general information. Retrieved from https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/qualityinitiativesgeninfo/index.html

Research Questions, Data, and Methods

Our objective was to evaluate the QBR and MHAC programs using a mixed-methods approach focusing on effectiveness and equity from both the program process and outcome perspectives. This paper focuses on the implementation process using qualitative data from in-depth key informant interviews with staff of four hospitals and the Maryland Medicaid program and a survey questionnaire completed by 12 hospitals. The research questions addressed covered by these data include: (1) How were the Maryland Quality Initiatives (QBR and MHAC) implemented? (2) How were the measures (process and outcome) selected and why? (3) What barriers or challenges were encountered during implementation? (4) How were barriers or challenges addressed? We examine program impact on quality of care, disparities, and clinical outcomes in other manuscripts currently under review or being prepared. The study received Institutional Review Board approval.

We attempted to recruit five hospitals to vary in size and location and three payers for case studies. One community hospital and two commercial insurance companies refused to participate. We recruited one participating hospital from each of four geographic regions of Maryland (Eastern Shore, Baltimore City, Washington, DC Metro, and Western Maryland). One hospital was urban, one was suburban, and two were rural. One was a university medical center and the others were community hospitals. The participating payer was the Maryland Medical Assistance (Medicaid) Program. Each organization agreed to an initial site visit and follow-up phone interviews.

We recruited leaders at each hospital who represented different perspectives on the effects of QBR and MHAC on quality, clinical, and financial operations. We conducted 51 semi-structured interviews with chief executive officers, chief nursing officers, chief medical officers, chief financial officers, and quality improvement directors; Maryland Hospital Association staff; and the Maryland Medicaid agency staff between August 2009 and April 2012. We divided the staff interviews into three rounds in September 2009: just after incentive-based reimbursement program began in July 2009; in the middle of the evaluation period; and in February 2012, close to the end of the evaluation period. We conducted thirty-eight of the 51 interviews with respondents from the hospitals. In addition, we surveyed senior administrators at all 47 nonfederal acute care hospitals operating in Maryland in April 2010 using a self-administered questionnaire. We received 12 responses, which we analyzed as part of the qualitative data set instead of treating them as quantitative data, because the response rate was low.

At least two investigators conducted all interviews, with one acting as a note taker. We interviewed participants individually and in small groups using a topic guide developed by the investigators; we taped and transcribed each hour-long interview. We used a computer software program (NVivo 9) to analyze data from the interview transcriptions, notes, and the self-administered survey, and coded and organized the responses into categories and themes. We compared anticipated effects discussed in the first round of interviews (at program initiation) with observed effects reported in the second and third interview rounds. We interviewed the same people in each round except when they had changed jobs or left employment at the participating institutions. Two members of the research team coded interviews, and they each independently coded a sample of transcripts to ensure inter-rater reliability.

Limitations

Our results may have limited generalizability, because they are based on four hospital case studies and 12 survey responses from among the 47 acute care hospitals operating in Maryland in 2009. These hospitals might not fully represent the experiences and attitudes of all 47 hospitals, especially since they tended to receive additional reimbursement under QBR and MHAC. However, we followed good qualitative research practice by identifying the key hospital characteristics that we expected to affect hospitals' experiences—geographic region, population density, and type of ownership (i.e., public, nonprofit)—and assuring that the major categories of each characteristic are represented in our sample. We also interviewed the staff of the Maryland Hospital Association to ensure that the views of Maryland hospital in general were represented.

As with any hospital quality improvement (QI) program, QBR and MHAC were implemented in the context of many other concurrent quality improvement efforts at each hospital, which makes it challenging to identify their unique effects. In the absence of a comprehensive, accurate inventory of all concurrent QI activities at each of the four case study hospitals, which we do not have, we cannot consider them explicitly in our analysis. Nevertheless, the questions in our interview guide specifically addressed QBR and MHAC, and the interviewers continually focused the discussion on QBR and MHAC and how they differ from other QI interventions. Consequently, we believe that the experiences and attitudes we have associated with QBR and MHAC are accurate reflections of those programs.

Results

Our discussion of the reactions of hospital staff and stakeholders to the QBR and MHAC programs is based on four aspects of these two programs: (1) general response, (2) program design, (3) operational design, and (4) perceived impact on hospitals and patients, including disadvantaged or vulnerable patients, patient experience, quality, operational costs, and clinician attitudes.

General Response. At the end of our study, overall reactions of interview participants to both the QBR and MHAC programs were positive. Participants described them as a logical and reasonable starting point for QI and emphasized that their organizations were committed to improving quality. However, many participants stated that the QBR and MHAC seem more like financial policies designed to reduce costs than quality improvement initiatives. The hospitals see HSCRC's mission as financial, not clinical or quality improvement, and so associated the QBR and MHAC programs with financial policy. They felt there should have been more clinical buy-in. They also reported that most hospitals do not align their quality improvement projects with their financial agenda. Several participants commented that it was illogical for all hospitals in Maryland to have the same quality priorities, because they serve different patient populations.

Program Design. Although attitudes about the QBR and MHAC programs were positive in general, opinions about design varied by program and among hospitals. The chief difference in attitudes by program at the outset concerned their risk-bearing financial implications for hospitals. Hospital staff believed that the QBR was innocuous because (1) hospitals understood and were already reporting the QBR measures to CMS under the voluntary hospital reporting program, and (2) relatively little money was at risk. In contrast, they were extremely concerned about the originally proposed design of the MHAC program because it was expected to penalize hospitals for PPCs at the individual discharge level. The hospitals felt threatened by this

approach because (1) they did not understand the 3M PPCs scoring methodology, and (2) the costs of uncompensated care for PPCs at the individual discharge level were potentially enormous because the cost of treating complications can be much larger than the cost of treating the original reason for admission.

The four hospitals uniformly reported that the MHAC methodology was difficult to understand and that it was difficult to obtain explanations about how it would work. They questioned some of the specific indicators and were successful in getting some of them eliminated from the calculations. Several hospitals also questioned the large number of indicators in the MHAC, noting that some of them were for rare conditions. Further, they called the scoring methodology a "black box," because the data went to 3M to be analyzed and scores came back without detailed explanation.[27] Without understanding the scoring methodology, they believed they could not validate the scores it was producing, could not appeal unfavorable decisions effectively, and would have to spend significant staff resources on verifying the coding on which the PPCs were based for those cases where PPCs (and hence penalties) were flagged by the algorithm. The cost of treating complications, especially cascading complications, can be many times the costs covered by the APR-DRG payment, which placed an uncertain and potentially large amount of revenue at risk.

Ultimately, HSCRC rejected the individual discharge approach and adopted a methodology for the MHAC that paralleled the QBR methodology; measuring aggregate performance at the hospital level relative to the hospital's prior performance and statewide average performance and redistributing a withholding pool based on that performance. Under the aggregate system, the hospitals can predict their maximum dollar risk with considerable precision, which removed much uncertainty and made the MHAC much more acceptable. The MHAC began by withholding 0.5%, the same percentage as the QBR, which the hospitals felt (for QBR) placed a small amount at risk.

Although all four hospitals qualified for enhanced payments through QBR and MHAC (i.e., were "winners"), several respondents were concerned about the poorer performing hospitals, which could lose up to 0.5 % of annual revenue in our study period. Respondents explained that

> *Some [hospitals] may be on the bottom end; they would need additional resources to improve. Instead of getting money to improve, they get penalized.*
>
> *To implement and coordinate solutions will continue to be challenging for organizations that are hurt financially. We need more money, not less, for [this] mission.*

Operational Design. Respondents expressed concern about several operational features of QBR and MHAC, including the size of the risk, the way hospitals are ranked, the effect of small numbers, the focus of incentives on hospitals, and the implementation process.

- ***Risk for Revenue.*** Hospital representatives were concerned that the original MHAC design, which eliminated reimbursement for the marginal costs incurred because of a

[27] In 2010, HSCRC took over the PPC computations from 3M but continued to use the same algorithm. In the follow-up interviews, participants reported that the MHAC methodology was more transparent, but it was still difficult to understand.

PPC, exposed hospitals to too much financial risk. However, some respondents felt that the final model actually implemented—redistribution of a withholding pooling comprising 0.5% of total revenue from each hospital—did not offer a sufficient financial incentive to justify the additional cost of improved coding and clinical procedures needed to minimize the risk. They reasoned that (1) the potential loss or gain of revenue (0.5% at the start for both QBR and MHAC) was too small and (2) the potential winnings and losses from the withholding pool in the current year were not paid directly to the hospitals. Instead, they were reallocated in the form of adjustments to the next year's APR-DRG rates for each hospital, which diffused their impact. With respect to the first point, the one payer in the study, Maryland Medicaid, generally viewed both programs favorably but was concerned that the absolute amount of revenue at risk was too little to have an effect on hospitals and that the QBR measures were too oriented toward conditions of the elderly to benefit a significant portion of its beneficiary population.

The majority of respondents were unsure at the outset how the QBR and MHAC programs would affect revenue. One participant explained: "Uncertainty is there, and it's our enemy" and "I keep wishing we knew ahead of time…[in order to] organize the budget." Some respondents added that the uncertainty limits their ability to plan to use the at-risk money to expand or start new programs.[28]

At the final interview, hospitals varied in their reports of how they fared in terms of revenue. One hospital lost money in Year 1 but gained money in Year 2, ending with a small net benefit. Other hospitals reported a small increase in revenue "on the positive side but it's been an itty bitty thing." One respondent explained that despite being on the positive side, "the [implementation] expenses far outweigh the improvement in revenue from MHAC and QBR."

With respect to the second point, several participants commented that there are probably better ways to distribute the money: "If they brought in a bag of [money] in cash and plopped it on the table, you'd notice."

- ***Hospital Ranking.*** QBR ranks hospitals by comparing a hospital's performance against either the statewide average benchmark (attainment) or its own best previous performance (self-improvement). Some hospital representatives preferred to be compared with the statewide average rather than with the hospital's own self-improvement because they were concerned that they would top out on many of the measures using the self-improvement method.

  Some respondents indicated that the threat to the hospital's reputation from a low ranking carried more weight than the financial benefits or losses. One respondent said that

[28] Half of the hospitals we interviewed joined the Total Patient Revenue (TPR) Initiative in between our initial and follow-up interviews. Under the TPR program, hospitals negotiate a set revenue for the year with the HSCRC. Hospitals are guaranteed the global budget, so they have incentive to reduce readmissions, length of stay, and superfluous tests. The hospitals we interviewed who joined TPR knew how much they were getting up front, and therefore knew the dollar amount that was withheld, whereas other hospitals' at-risk money depended on the profit for that year. One hospital shared that it experienced low volume (inpatient admissions and LOS) in its first year, which would result in a $3.5 million excess revenue over costs based on the negotiated global budget.

"shame is [the] number-one [motivator]" because the rankings and scores are publicly reported.

- ***Low-Volume Hospitals.*** Staff at the smaller hospitals were concerned about the loss of precision when low-volume hospitals have few cases, as quoted below.

  > *We've got measures where if in a 30-day period one patient does not meet the measure, we've lost because we only have three patients total who were eligible. We are a victim of small numbers.*

  We learned that the process of introducing the QBR and MHAC systems was long and tedious, but one in which some small- and medium-sized hospitals did not engage fully, because they did not have staff with the knowledge or availability to devote to understanding the systems. Instead, they relied on the larger, more resource-rich hospitals that had knowledgeable staff available to represent their interests in negotiations with HSCRC, even though they acknowledged that their interests were not necessarily congruent with those of the larger hospitals.

- ***Lack of Physician Incentives within These Programs.*** The Maryland Hospital Association was skeptical about the value of the measures for improving quality because it believes that holding hospitals accountable without also holding the physicians accountable weakens the effectiveness of the incentives.

- ***Involvement of Hospitals in Program Design.*** Hospital involvement with the design of the QBR and MHAC programs varied among the hospitals. Some hospitals had no involvement, while others regularly attended HSCRC meetings, which were convened by the HSCRC at the Maryland Hospital Associations offices and were open to the public. Hospitals tended to view the HSCRC meetings, and the process in general, unfavorably. Respondents emphasized that they appreciated the opportunity to give feedback and that the HSCRC was gracious and good at following up. A MHAC final staff recommendations report in June 2009 indicated that the concerns from hospitals were carefully recorded and addressed, which eventually changed the case-specific and narrowly defined MHAC approach (11 HACs) to a broader, rate-based alternative (Health Services Cost Review Commission, 2009). However, respondents felt the meetings should have been held much earlier.

Perceived Impact on Hospitals and Patients. During the first round of interviews, we asked hospitals about effects they expected the programs to cause. During the follow-up interviews, we provided respondents with their initial responses and asked if the anticipated effects had occurred. Topics discussed pre- and post-implementation included disparities, patient experience, quality, operational cost, and clinician attitude toward the programs.

- ***Disparities among Disadvantaged Populations.*** Participants initially felt the implementation of the QBR and MHAC programs would not have an effect on disadvantaged populations. Hospital staffers responded that they treat patients equally regardless of race, ethnicity, socioeconomic status, or insurance status, describing that they "only look at the medical number" and that if the program caused them to discriminate, it will "cause [them] to discriminate to all patients equally." One respondent stated, "It can only help their outcomes and not necessarily hamper them or hurt them in

any way," and another explained, "It may identify disparity, but I don't think it is going to drive disparity." Overall, the respondents felt that if the programs improve care, then they would do so for all patients.

However, participants also believed that serving a more disadvantaged population would affect a hospital's performance in the QBR and MHAC systems. Uninsured or underinsured patients can encounter delays in seeking care, which can complicate their condition and lead to more comorbidities. Consequently, participants recognized the importance of coding POA accurately and were concerned about the feasibility and cost of doing so.

Looking back from the perspective of the follow-up interviews, participants tended to confirm their initial expectations, reporting that the programs are "applied pretty evenly" to patients and that they strictly focus on the clinical needs of patients. One respondent explained that because Maryland sets the same payment rates for all payers, all patients are treated the same. Therefore, it would be unlikely that persons typically subject to disparities would fall further behind because of the QBR or MHAC programs.

- ***Patient Experience.*** Initially, most participants felt that implementing the QBR and MHAC programs would not affect patient experience. One respondent said, "Avoidance of injury is always a satisfier, if they know it could happen. But [patients usually] focus on typical things like meal trays." One participant thought that providers might order rule-out tests more frequently and that the need to avoid penalties could drive patient care (unfavorably).

  The follow-up interviews about patient experience and outcomes were consistent with participants' initial assumptions. Participants agreed that the programs had no effect on patient experience measures. However, participants felt that, with time, patient experience might improve if patient outcomes improve because of the QBR and MHAC programs.

  > *The patient's clinical outcome itself has probably improved with less complications, etc. However, a patient's experience is a more holistic experience of everything that happens for them from the time they enter the hospital to the time they leave the hospital. It has to do with how kind [people] were to them, whether they were feeling cold or warm, were they getting blankets—that's patient satisfaction.... I don't think MHAC and QBR itself has had any impact on patient experience.*

- ***Quality.*** In general, participants initially believed that QBR measures would have a greater effect on quality improvement than some—not all—of the MHAC measures. A few participants acknowledged the potential value of the QBR program but were concerned about its focus on process rather than patient outcomes:

  > *It's hard to argue with the benefit [when] you provide all recommended care to patients. On the other hand, our inability to link performance on [QBR] core measures [to] overall improvement and disease process and outcomes makes you question the amount of effort and attention paid.*

Another issue mentioned in the initial interviews was whether the emphasis on the MHAC measures would crowd out existing individual hospital quality initiatives. Several respondents expressed concern that they were already undertaking efforts that would likely be pushed aside to respond to the demands that MHAC would place on the time of quality improvement staff.

In the follow-up interviews, participants had mixed feelings about whether the QBR and MHAC programs improved quality of care. Responses included the following: definitely improved, no effect, too soon to tell, and unsure. No respondents felt that the programs decreased quality of care. Some respondents felt the QBR program improved care more than the MHAC program because of the improved consistency and reliability of delivering care that it encourages. Other respondents felt that the MHAC program better improved quality, as it helped identify problem areas throughout the hospital better than the QBR measures could because the QBR measures focus on fewer conditions.

- ***Operational Cost.*** Participants expected the QBR and MHAC programs to increase costs in several departments. Hospitals did not receive funding to assist with implementation; therefore, any resources they use would have to be shifted from something else:

  > *[The QBR & MHAC programs are] not something that you just add on. It's going to be [an] additional expense.*

  This belief is consistent with the notion, discussed previously, that MHAC might reduce emphasis on existing local quality improvement initiatives.

  Incremental operational costs included collecting and tracking data and educating hospital staff on the new policies and procedures. Some participants felt that QBR would not increase costs, because they were already reporting the QBR measures to CMS and had already integrated them into their quality improvement culture, but all were concerned about the cost of MHAC. Participants cited the additional costs of changing the culture to improve charting and coding of POA conditions, which is needed to avoid having them treated as PPCs. They also were concerned about increased reliance on rule-out testing at admission; for example, increased urine testing to rule out urinary tract infections. One participant explained that even though all the systems are in place, it's going to take a lot more of everyone's time to ensure the whole process is up and running. One chief nursing officer expected that the nursing staff would have to spend time monitoring and collecting data, standardizing care, and decreasing variation. Other respondents planned to hire quality improvement staff to review records and educate clinicians and coders. The hospital respondents also felt that, to protect themselves from errors, they would have to validate the assignment of each PPC on the list of PPC cases they receive periodically from HSCRC, by assigning staff to reassess the medical and billing records to search for evidence that the condition triggering the PPC was present on admission, even if it was not originally coded as such. At one hospital, participants felt that this work could extend to having the investigating nurse interview the clinicians who entered the notes. Many participants also were concerned about the cost of preventing complications. For example, to prevent falls with fracture, one chief medical officer

explained that he would have to hire a sitter for every patient at a certain level of risk or figure out another strategy, also at a cost.

The follow-up interviews revealed mixed outcomes in terms of cost and resources. The staff at one hospital unanimously agreed that the effects were "minimal, if anything," whereas all respondents from a different hospital agreed that both programs had increased their costs. Staff at a different hospital mentioned that they had not incurred additional costs, because the finance team decided not to divert any resources away from in-house quality projects to improve performance on the QBR and MHAC measures. Some hospitals did not incur new marginal costs in total (e.g., new hires, purchased software), but resources had to be reallocated from other uses to implement the QBR and MHAC programs.

The participants who perceived an increased marginal cost of care attributed it to new hires; data collection; PPC prevention materials (e.g., gloves, hand soap/sanitizer); and staff time spent understanding, preparing, and implementing new processes. A few participants described the implementation process as "labor intense." One respondent stated that there were significant start-up costs but felt that, after a year, those had neutralized.

- ***Clinician Attitudes.*** Initially, most respondents expected that it would be a challenge to obtain clinicians buy-in to the new programs. "It is hard to just go into a physician's office and get them to do things…. I need to be more persuasive rather than dictatorial." Most doctors don't have time to review every case with coders, and changing their culture was seen as a challenge. One respondent explained, "This may be a shock, but the way to get doctors' attention is to give them more or less money. The data may be interesting to them; their enthusiasm for nonfinancial performance is about equal to their enthusiasm for herpes. It's just not there." In addition to suggesting financial incentives for physicians, one respondent commented on the lack of physician accountability: "Hospital performance is a byproduct of physician performance, so…put something in it for doctors."

  In the follow-up interviews, hospital staff participants unanimously stated that the programs had no effect on clinical staff satisfaction. At one hospital, respondents stated that they mainly focused on training for coders and not physicians, as they had anticipated. Another participant described the documentation requirements of MHAC as "significantly excessive" and stated that clinicians spend too much time documenting and not enough time providing care at the bedside. One participant added, "This underscores a point that this MHAC data that we get periodically [have] not led to any profound clinical insights or interventions."

Discussion

The Maryland experience provides unusual insight into widespread adoption of performance-based reimbursement, because it is the only state that sets the same rates for Medicare beneficiaries and all other payers. Because Medicare accounts for more than one quarter of inpatient reimbursement, it presents a potentially significant counterweight to the incentives imposed by regulation of other payers in other all-payer states. The qualitative data we used for

this article (and the quantitative data we use for our impact analyses currently under review elsewhere) pre-date the Medicare HAC Reduction Program, which began in 2015, and the Medicare Hospital VBP Program, which began in 2012, and which use incentive structure similar to the Maryland MHAC and QBR programs, respectively. In this sense, Maryland offers a window into a world in which all reimbursement for acute inpatient care will be based on the same set of performance-based incentives.

The use of an aggregate, fixed withholding pool for MHAC and QBR caps risk by setting a specific percentage of DRG-based reimbursement to be redistributed. The hospitals considered that percentage to be relatively small compared with the penalties that could be incurred in the discharge-specific approach originally considered for the MHAC program, in which all marginal costs attributed to a PPC would not be reimbursed. But more importantly, the financial risk can be predicted with greater precision using the aggregate approach than using the discharge-specific approach. Once the financial risk of the MHAC program became comparatively low and predictable, hospital administrators' concerns diminished. With predictable risk and diminished concern, the need to invest heavily in the immediate improvement of POA recording and coding became less urgent.

The urgency of contesting PPC determinations by the MHAC program—and the costs of doing so—also diminished because the penalties were not immediate. Aggregate scoring was conducted annually, and the redistributed payments were made in the following year through adjustments to the APR-DRG rates. Thus, the pain—the percentage of the APR-DRG payments withheld—was immediate, while the potential rewards—the redistributed withholdings—were delayed. Delayed rewards tend to be worth less than immediate rewards from both the financial and the psychological perspectives, so it is possible that the hospitals did not work as hard to avoid the penalties as they would have if the penalties and rewards were more proximal. Thus, the changes in the financial incentive structure reduced concern as well as the likely impact of the program.

The Maryland hospitals found the initial plan to withhold payment for the marginal cost of PPCs in individual discharges problematic, because they felt that the specific PPCs were likely to be inappropriately adjusted, the financial risk was not capped, and the cost of implementation (i.e., improved recording and coding of POA conditions and reviewing records to validate the accuracy of PPC determinations and penalties) was high. The revised and implemented MHAC program echoes the design of the CMS HAC Reduction Program, where a hospital-specific weighted scoring method is used. The weighted score comprises eight Agency for Healthcare Research and Quality Patient Safety Indicator 90 composites and three Centers for Disease Control and Prevention–defined hospital-acquired infection measures.[29] The aggregate methods adopted in both the Maryland and CMS programs steered rule makers away from possible disagreements over the appropriateness of the case adjustments, and thus the debates over false positive and false negative cases.

[29] The FY 2016 HAC Reduction Program uses the Agency for Healthcare Research and Quality Patient Safety Indicator 90 Composite and the CDC National Healthcare Safety Network Central Line-Associated Bloodstream Infection, Catheter-Associated Urinary Tract Infection, and Surgical Site Infection measures.

The difference in hospitals' expectations for the QBR and the MHAC systems at the outset was one of the most striking findings. Hospital staff were generally skeptical at the outset that the QBR system would present much of an incentive, because the amount of money at risk was small, physicians were not at risk under the QBR system, and the financial benefits of good performance would not be paid directly, but instead would be distributed as adjustments to the APR-DRG rates in the following year. So, although they were pleased with the change in the MHAC design, they also acknowledged that the adoption of the QBR-type approach for the MHAC weakened its power as an incentive, even though the percentage withheld for the MHAC would eventually reach 4% compared with the 0.5% at risk under QBR. Hospitals had little concern about the QBR measures because they had been reporting POC measures to CMS for several years. In contrast, Maryland hospitals had little if any history recording or reporting the data needed to calculate the PPCs; believed they would incur significant future costs of recording, reporting, and validating PPC determinations; and did not understand how the PPCs were calculated. The substantial and meaningful decrease in hospitals' uncertainty resulting from the shift of MHAC from the individual discharge to the aggregate redistribution approach explains why hospital staff reported fewer concerns at the end of the study than at the beginning and why a number of participants reported that the significant implementation expenses they had anticipated did not materialize.

These findings also identified an interesting difference in the interpretation of risk by hospital administrators that should be considered when designing incentive programs. Throughout this article, we attribute the diminishing concern about the MHAC program to diminishing financial risk associated with the switch from refusal to pay for discharge-specific marginal PPC costs to redistribution of an aggregate withholding pool drawn from all discharges. However, this shift also increased the number of PPCs for which hospitals could be penalized from 11 under the original proposal to 47 under the MHAC as implemented and some administrators noted that the probability of being "shamed" by poor performance was a greater motivator than the amount of money at risk. It appears that the risk of experiencing a PPC, with its attendant shame, was less salient to these administrators than the predictability and the limit of financial losses afforded by the aggregate redistribution approach. We will investigate this issue further in a subsequent article.

Participants acknowledged the potential for positive impact on patients if the QBR and MHAC programs improve quality, but no one reported that they expected or experienced changes in disparities or patient experience of care because of the QBR and MHAC programs. They also reported at the end of the study that the QBR and MHAC systems had no impact on staff satisfaction or burden. Initially, they had expected the MHAC program to substantially increase staff burden by requiring training in recording and coding POA conditions and extensive follow-up by staff to validate each PPC finding. Although they incurred training costs, the follow-up costs did not materialize to the extent anticipated, because the risk of loss was much lower than expected and did not merit ongoing investment in risk mitigation. Whether the QBR and MHAC programs had an impact on quality of care and disparities will be the subject of forthcoming

articles based on patient-level analyses, though HSCRC has published findings that suggest a favorable impact.[30]

Practical Implications

Although CMS' Hospital VBP and HAC Reduction Programs apply only to Medicare beneficiaries and the Maryland P4P systems apply to all payers, they are similar on dimensions that the hospitals considered important. The major difference between the CMS and Maryland systems was the treatment of HACs before the CMS HAC Reduction Program began in 2015. At that time, CMS operated the HAC POA Indicator Program, in which it withheld all payment for the marginal costs associated with PPCs and Maryland used the withholding pool approach for PPCs. CMS added the MHAC-like HAC Reduction program in FY 2015. Including HACs represents an opportunity and demand to move beyond process measures alone to include clinical outcomes as one basis for quality-based reimbursement. So, Maryland's earlier experience in implementing a program similar to CMS's HAC Reduction program may provide useful insights into the implementation of HACs in the Medicare program beyond 2015.

Based on the Maryland experience, implementation of hospital P4P programs should take a number of important factors into consideration. If financial incentives and the number of performance measures (e.g., number of PPCs) overburden hospitals with operational costs that are perceived to exceed the benefits of quality improvement, the incentive for change will diminish and other hospital-initiated quality activities will likely be crowded out, especially at small-volume hospitals. Stronger financial incentives and increased risk exposure for hospitals is associated with more influence on hospital behavior and greater likelihood of change, but also potentially greater resistance (due, in part, to high costs of mitigating high risk). One possible implication is that programs should be introduced with low risk and ramped up over time as they become familiar, to manage the resistance while still creating sufficiently strong incentives to create behavior change. Whether the kinds of financial risk imposed by QBR and MHAC programs are sufficient to motivate hospitals to improve is another issue that we will attempt to address in subsequent publications that report on the impact of the QBR and MHAC programs on improvement in measures and ethnic and economic disparities.

Because physician services are an integral component of hospital care, study participants believed that embedding physician financial risk within hospital P4P programs is essential to success. Further, transparency of scoring methods is key to hospitals' acceptance of and commitment to P4P. Including providers' perspectives in program design decreases the resistance to program implementation and provides the program developers with better insights into the likely impact of the program during its development. Although many participants in this study felt that they had had little if any influence on the design of the QBR and MHAC programs, pressure from hospitals in the aggregate seems to have had an important effect on the decision to implement the MHAC as an more predictable aggregate redistribution program instead of a highly uncertain and potentially costly discharge-specific program that would pay none of the costs incurred for PPCs.

[30] Calikoglu, S., Murray, R., & Feeney, D. (2012). Hospital pay-for-performance programs in Maryland produced strong results, including reduced hospital-acquired conditions. Health Affairs, 31(12), 2649–2658.

Authors, Their Affiliations and Their Contact Information:

Steven Garfinkel, PhD, Institute Fellow, American Institutes for Research [Correspondent Author]
100 Europa Drive, Suite 315, Chapel Hill, NC 27517; Phone: 919-918-2306 (office);
Email: sgarfinkel@air.org, Fax: 919-960-6983

Ying Wang, PhD, Principal Researcher, American Institutes for Research (AIR)
201 Jones Road, Suite 1, Waltham, MA 02451; Phone: 781-373-7034 (office);
Email: yiwang@air.org

Yi Lu, PhD, Economics Researcher, American Institutes for Research
1000 Thomas Jefferson Street, NW, Washington, DC 20007; Phone: 202-403-6807 (office)
Email: ylu@air.org; Main Fax (202) 403-5454

Marilyn Moon, PhD, Senior Vice President, American Institutes for Research
6003 Executive Blvd., 3rd Floor, Rockville, MD 20852; Phone: 301-592-2101 (office);
Email: mmoon@air.org

Brenna Raines, MPP, Senior Manager, Plasma Protein Therapeutics Association
147 Old Solomons Island Road, Ste. 100, Annapolis, MD 21401, USA; Phone: 443.458.4667 (office); Email: BRaines@pptaglobal.org

Christian Evensen, MS, Principal Researcher, American Institutes for Research
100 Europa Drive, Suite 315, Chapel Hill, NC 27517; Phone: 919-918-2310 (phone);
Email: cevensen@air.org; Fax: 919-960-6983