

Bridges to Excellence—Recognizing High-Quality Care: Analysis of Physician Quality and Resource Use

Meredith B. Rosenthal, PhD; Francois S. de Brantes, MBA; Anna D. Sinaiko, MPP; Matthew Frankel, MBA; Russell D. Robbins, MD, MBA; and Sara Young, MBA

Objective: To examine whether physicians who sought and received Bridges to Excellence (BTE) recognition performed better than similar physicians on a standardized set of population-based performance measures.

Study Design: Cross-sectional comparison of performance data.

Methods: Using a claims dataset of all commercially insured members from 6 health plans in Massachusetts, we examined population-based measures of quality and resource use for physicians recognized by the BTE programs Physician Office Link and Diabetes Care Link, compared with nonrecognized physicians in the same specialties. Differences in performance were tested using generalized linear models.

Results: Physician Office Link-recognized physicians performed significantly better than their nonrecognized peers on measures of cervical cancer screening, mammography, and glycosylated hemoglobin testing. Diabetes Care Link-recognized physicians performed significantly better on all 4 diabetes process measures of quality, with the largest differences observed in microalbumin screening (17.7%). Patients of Physician Office Link-recognized physicians had a significantly greater percentage of their resource use accounted for by evaluation and management services (3.4%), and a smaller percentage accounted for by facility (-1.6%), inpatient ancillary (-0.1%), and nonmanagement outpatient services (-1.0%). After adjustment for patient age and sex, and case mix, Physician Office Link-recognized physicians had significantly fewer episodes per patient (0.13) and lower resource use per episode (\$130), but findings were mixed for Diabetes Care Link-recognized physicians.

Conclusions: Our findings suggest that the BTE approach to ascertaining physician quality identifies physicians who perform better on claims-based quality measures and primary care physicians who use a less resource-intensive practice style.

(*Am J Manag Care.* 2008;14(10):670-677)

For author information and disclosures, see end of text.

During the past several years, both public and private payers have adopted pay for performance in a wide variety of contexts with the hope of prompting more evidence-based and higher value patterns of care.¹⁻³ When it was launched in 2002 by a group of large employers collaborating with several health plans and provider organizations, Bridges to Excellence (BTE) was one of the first multistakeholder pay-for-performance programs.^{4,5} Its programs have relied on certification of performance based on self-reported (subject to audit) medical record and practice systems data. Physicians are invited by BTE to seek certification and begin receiving payments for all patients covered by the program's sponsors. Incentives for physicians who meet or exceed the performance criteria were initially determined from actuarial analyses that estimated savings to employers from better quality of care⁶ and are paid by participating employers and health plans directly to physicians, according to a specified bonus schedule that can vary between implementation regions.

Now operated through an independent not-for-profit entity with a board comprised of employer, health plan, and physician representatives, BTE's model is in operation nationwide through licenses to health plans, employer coalitions, and state agencies.⁷ Bridges to Excellence was first implemented in Massachusetts in 2003 with 2 major physician reward components: the Physician Office Link (POL) and the Diabetes Care Link (DCL). Attainment of the quality standards set by BTE is associated with both financial rewards and public recognition.

In 2003, approximately 20 physicians in Massachusetts were recognized by BTE under what was then its only program: DCL. By the end of 2006 there were more than 1000 BTE-recognized physicians in Massachusetts across all BTE programs and total payments to these physicians reached \$2.4 million. (In 2004 the Cardiac Care Link was introduced. Because that program is newer, we do not examine it in our analysis.) Nationally, 8500 physicians have been recognized and a total of \$10 million paid out. Although the BTE model continues to expand, comparative studies of BTE-recognized physicians and their practice patterns are lacking. In this article, we examine the quality, care delivery patterns, and resource use of BTE-recognized physicians compared with nonrecognized physicians in Massachusetts.

On quality measures alone, this comparison is of interest because of BTE's unusual reliance on voluntary certification through a site survey

In this issue
Take-away Points / p677
www.ajmc.com
Full text and PDF

and chart review that yields data regarding quality measures similar to those used in the Healthcare Effectiveness Data and Information Set (HEDIS). For the POL in particular, which is intended to reward overall population health management, substantial weight is placed on the structural measures of quality derived from the site survey. Thus, it is not clear that physicians with BTE certification would necessarily get better-than-average scores on the quality measures alone. Moreover, there is no comparative aspect to BTE recognition—the standards are set nationally and fixed (ie, they do not reflect relative rankings within a market), so it is unknown whether the physicians seeking certification are those whose performance is above average or simply those with larger numbers of eligible patients (and thus with more to gain). Finally, because BTE certifies physicians for 3 years based on a single retrospective review of 25 sampled charts for the quality measures, we also considered it important to ask whether a contemporaneous comparison between recognized and nonrecognized physicians would show better performance for BTE physicians. Even if recognized physicians performed better than average with the selected patients, these differences might not have persisted.

METHODS

We examined population-based measures of the quality of care and resource use for Massachusetts physicians participating in BTE's first 2 rewards programs, POL and DCL. Using an administrative dataset comprised of claims from all commercially insured members of 6 Massachusetts plans, we compared the performance of recognized physicians with the performance of nonrecognized physicians.

Bridges to Excellence Program Design

The POL program focuses on promoting the office practice's use of systems to enhance the quality of patient care. Today, this assessment is commonly being referred to as a core characteristic of a Medical Home.⁸ Practices that can demonstrate specific processes and systems of care can earn up to \$50 for each patient covered by a participating employer. (The rewards were designed to last 3 years, the period of time during which the recognition is valid.) To obtain the rewards available through the POL program, eligible physicians must pass the National Committee for Quality Assurance (NCQA) Physician Practice Connections assessment program or meet a comparable standard set by the Massachusetts Quality Improvement Organization: MassPRO. The standards for the POL, which are described in **Table 1**,⁹ include maintenance of patient registries for the purpose of identifying and following up with at-risk patients, provision of educational resources to patients, and use of electronic systems to maintain patient

records, provide decision support, enter orders for prescriptions and lab tests, and provide patient reminders.

The DCL program rewards physicians with \$80 per year for each diabetic patient covered by a participating employer. To obtain the rewards, eligible physicians must demonstrate that they provide high levels of diabetes care by passing NCQA's Diabetes Physician Recognition Program, which was developed in collaboration with the American Diabetes Association.¹⁰ To qualify for the 3-year recognition, physicians submit medical record data on glycosylated hemoglobin (A1C), blood pressure, and lipid testing, as well as data on eye, foot, and nephropathy exams, for a random sample of 25 diabetes patients. (The 25 charts are selected retrospectively for patients who have been in the practice's care for at least 1 year. A date is selected at random from a set window of time, and the next 25 charts for patients with a diagnosis for diabetes are pulled for analysis. Both the sampling method and the actual charts are reviewed and subject to NCQA audit.) The NCQA assesses the data submitted by the physician and attributes points based on the level of disease management demonstrated by the physician. All information submitted for POL and DCL certification is subject to audit.

Data Sources

We obtained data from BTE on the identities of physicians recognized for the POL and DCL programs between 2003 and 2006. Administrative data, including claims and enrollment files, were obtained from the Massachusetts Group Insurance Commission (GIC), which has created a database that includes all claims from all commercially insured members (not just those for GIC members) from the 6 health plans that serve the GIC (Harvard Pilgrim Healthcare, Tufts Health Plan, Unicare, Fallon, Neighborhood Health Plan, and Health New England). The GIC's data cover approximately 50% of privately insured residents and more than 7000 physicians in Massachusetts including all but 2 BTE-recognized physicians. Physician-level measures of quality and resource use per episode were computed by Mercer Human Resources Consulting using profiling software licensed from Resolution Health, Inc, and Ingenix, respectively.

Quality and Cost Measures

To ascertain differences between recognized and nonrecognized physicians, we examined claims-based quality measures and average resource use per standardized episode for several categories of services and the total. (Episodes were created using Symmetry's Episode Treatment Groupers, a unit of Ingenix. These episodes are defined using proprietary algorithms that identify in claims data clinical events that trigger, break, and end an episode. Similarly, algorithms based on clinical

■ **Table 1.** Physician Practice Connections Survey Elements^a

PPC	Standard and Element Titles	Element Points	Standard Points
PPC 1	Access and Communication		
PPC 1A	Access and communication processes	4	
PPC 1B	Access and communication results	4	
	Total points for PPC 1		8
PPC 2	Patient Tracking and Registry Functions		
PPC 2A	Basic system for managing patient data	2	
PPC 2B	Electronic system for clinical data	3	
PPC 2C	Use of electronic clinical data	3	
PPC 2D	Organizing clinical data	6	
PPC 2E	Identifying important conditions	4	
PPC 2F	Use of system for population management	2	
	Total points for PPC 2		20
PPC 3	Care Management		
PPC 3A	Guidelines for important conditions	3	
PPC 3B	Preventive service clinician reminders	4	
PPC 3C	Practice organization	3	
PPC 3D ^b	Care management for important conditions	5	
PPC 3E	Continuity of care	5	
	Total points for PPC 3		20
PPC 4	Patient Self-Management Support		
PPC 4A	Documenting communication needs	2	
PPC 4B	Self-Management support	4	
	Total points for PPC 4		6
PPC 5^b	Electronic Prescribing		
PPC 5A	Electronic prescription writing	3	
PPC 5B	Electronic prescribing interoperability	3	
PPC 5C	Prescribing decision support—safety	3	
PPC 5D	Prescribing decision support—efficiency	2	
	Total points for PPC 5		11
PPC 6^b	Test Tracking		
PPC 6A	Test tracking and follow-up	6	
PPC 6B	Electronic system for managing tests	6	
	Total points for PPC 6		12
PPC 7^b	Referral Tracking		
PPC 7A	Referral tracking	4	
PPC 7B	Referral decision support	3	
	Total points for PPC 7		7
PPC 8	Performance Reporting and Improvement		
PPC 8A ^b	Measures of performance	3	
PPC 8B	Reporting to physicians	3	
PPC 8C	Setting goals and taking action	3	
PPC 8D	Reporting standardized measures	2	
PPC 8E	Electronic reporting—external entities	1	
	Total points for PPC 8		12
PPC 9	Interoperability		
PPC 9A	Use of prescribed standardized codes	1	
PPC 9B	Electronically receiving data	1	
PPC 9C	Electronically transmitting data	1	
PPC 9D	Using data for referral reports	1	
	Total points for PPC 9		4

PPC indicates Physician Practice Connections.

^aData were derived from reference 9.

^bPractices must pass this element of survey for recognition.

logic are used to group all claims related to that episode. For a more complete description, see reference 11.) For the POL cohort, we examined a set of measures typically used for examining primary care quality, including cervical cancer screening, mammography, A1C testing, cholesterol screening for individuals with coronary heart disease, and cholesterol screening for individuals with hypertension. For the DCL cohort, we examined 4 widely used claims-based measures of diabetes care quality. These were A1C testing, cholesterol testing, microalbumin testing, and diabetic retinal exams. Measure specifications were adapted by Resolution Health, Inc, from HEDIS and other national evidence-based guidelines to conform to the provider profiling context. The most recent 18 months of outpatient claims and pharmacy data were used for evaluation and calculation of each quality score (eg, A1C testing rate).

Eligible patients were attributed to multiple physicians based on the following steps. In the first step of the analysis, a member's claims were evaluated to determine which physicians had the greatest number of claims. In the second step, the rules (eg, diabetic patients should have A1C testing at least annually) that are applicable to an individual member are then assigned to all the physicians identified in step 1 who could be responsible for giving that type of care based on their specialty (eg, internal medicine or endocrinology could be responsible for diabetes care, but dermatology could not). The quality score for a physician was then calculated by determining what percentage of the patients for whom the physician was deemed "responsible" through this process met the criteria for the rule (eg, annual A1C testing).

The second set of measures we examined includes standardized measures of spending that we refer to as resource use (we used Medicare rates rather than the actual transaction prices, which are unique to each payer-provider pair). We decomposed resource use in 2 ways. First, we use the Episode Treatment Grouper software from Ingenix to tally resource use in 6 categories: evaluation and management, surgery (claims submitted by physicians for surgical interventions), facility (room and board only), inpatient ancillary (inpatient claims for services other than room and board), outpatient (claims for services provided in the outpatient setting other than evaluation and management), and prescription drugs.¹¹ Eligible patients were attributed to physicians who were responsible for the majority of the clinician fees in the episode. We compared average percentages of total patient resource use in each category for recognized and nonrecognized physicians. Then, we examined the number of episodes per patient and the total resource use per episode.

Analytic Approach

We compared average performance for physicians recognized by the 2 BTE programs with the performance of similar

physicians who were not recognized. All physicians who were eligible for BTE rewards in Massachusetts were included in the comparison groups. The initial list was refined by excluding physicians who did not have at least 200 episodes and did not have enough data to receive a quality score. Additionally, comparison groups were selected from the same specialties as those of the recognized physicians: internal medicine and family practice for the POL and those 2 specialties plus endocrinology for the DCL. For the POL analysis, the comparison group was further restricted by excluding physicians who were outliers in their resource use (ie, whose average resource use per episode of care was more than 6 standard deviations above the mean resource use for all POL physicians). Analyses were conducted using the physician as the unit of analysis. The statistical significance of differences in average performance for each measure of resource use and quality was determined by using generalized linear models in which the explanatory variables were specialty, mean patient age, the percentage of patients who were female, and a case mix indicator, which is a measure of the resource intensiveness of the Episode Treatment Groupers of a physician's entire patient population compared with Episode Treatment Groupers for other physicians in the same specialty. We report predicted values from these regressions for both sets of physicians and used a bootstrapping (ie, resampling) method¹² to compute 95% confidence intervals (CIs) for the differences.

RESULTS

We identified 405 physicians who were recognized by BTE's POL program between 2003 and 2006 and 3916 nonrecognized physicians in the GIC's data (Table 2). Physicians recognized through the POL were more likely to be specialists in internal medicine than nonrecognized physicians and were less likely to be family practitioners. Patient sex was similar between POL-recognized and nonrecognized physicians, but patients were younger and case mix was substantially higher for nonrecognized physicians. Between 2003 and 2006, BTE's DCL program recognized 91 Massachusetts physicians; GIC's data showed 1204 nonrecognized physicians treating diabetic patients during the same time period. Physicians recognized through the DCL were much more likely to be specialists in endocrinology than nonrecognized physicians and much less likely to be family practitioners. Patient age and sex were similar between DCL-recognized and nonrecognized physicians, but case mix was substantially lower for nonrecognized physicians.

Comparison of Quality Measures

On all process measures of quality, recognized physicians performed better than nonrecognized physicians (Table 3).

■ **Table 2.** Sample Characteristics^a

Characteristic	POL Study Group			DCL Study Group		
	Recognized	Nonrecognized	P	Recognized	Nonrecognized	P
No. of providers	405	3916		91	1204	
Specialty, %			<.001			<.001
Internal medicine	87	77		62	77	
Family practice	13	23		7	20	
Endocrinology	NA	NA		32	2	
Average age, y	44.4	42.7	<.001	44.9	44.1	.10
Female, %	58.6	58.2	.65	55.4	54.0	.41
Average case mix for practice^b (SD)	0.90 (0.81)	1.05 (1.03)	.006	1.73 (1.36)	0.95 (0.52)	<.001
Total episodes^c	403,451	2,472,675		69,954	1,416,277	
Mean episodes per provider	996	631	<.001	769	1176	<.001

DCL indicates Diabetes Care Link; NA, not applicable; POL, Physician Office Link.

^aDerived from authors' analysis of Massachusetts Group Insurance Commission data.

^bCase mix is an index of the intensity of use for a physician's panel compared with others in the same specialty: a higher value is more resource intensive.

^cAn episode is defined as a group of services delivered to treat a condition over a period of time, delimited by events that begin, break, and end a coherent treatment regimen. Episodes here are defined based on Ingenix Symmetry Episode Treatment Group Software.

Differences were statistically significant for cervical cancer screening, mammography, and A1C testing for the comparisons of POL-recognized physicians with nonrecognized physicians. The DCL-recognized physicians performed statistically significantly better on all 4 diabetes process measures of quality, with the largest differences observed in microalbumin screening (17.7%; 95% CI = 14.0%, 21.4%). This second finding seems to affirm that the NCQA's Diabetes Physician Recognition Program concurs with claims-based measures of high-quality care for patients with diabetes.

Comparison of Practice Patterns and Resource Use Measures

Differences in practice patterns between physicians recognized by the POL program and the comparison group shared a common pattern but had mixed degrees of statistical significance (Table 4). In particular, patients of POL-recognized physicians had a significantly greater percentage of their resource use accounted for by management services (3.4%; 95% CI = 2.4%, 4.4%) and a smaller percentage accounted for by facility (−1.6%; 95% CI = −3.4%, −0.02%), inpatient ancillary services (−0.1%; 95% CI = −0.1%, −0.03%), and nonmanagement outpatient services (−0.1%; 95% CI = −1.7%, −0.3%). Patients of POL-recognized physicians also had a lower percentage of their total resource use accounted for by prescription drugs (−1.6%; 95% CI = −2.8%, −0.3%). The distribution of expenditures for DCL-recognized physicians compared with nonrecognized physicians treating diabetic patients were more similar, with smaller, but significant

differences only in the share of standardized resource use accounted for by management (2.0%; 95% CI = 0.5%, 3.6%) and surgery (−0.3%; 95% CI = −0.5%, −0.1%).

After adjustment for specialty, patient age and sex, and case mix, POL-recognized physicians had significantly fewer episodes per patient (0.13; 95% CI = 0.13, 0.15) and lower resource use per episode (\$130; 95% CI = \$119, \$140) (Table 5). For diabetes care, however, DCL-recognized physicians had more episodes per patient (0.17; 95% CI = 0.14, 0.20 for primary care physicians; and 0.08; 95% CI = 0.03, 0.13 for endocrinologists). Resource use per episode was lower for DCL-recognized primary care physicians (\$26; 95% CI = \$5, \$47) and higher, but not significantly so, for DCL-recognized endocrinologists (\$137; 95% CI = −\$45, \$320).

DISCUSSION

Bridges to Excellence is a national pay-for-performance program that relies on chart review and structural survey data to assess the quality of care in physician practices.^{5,13} Although analyses conducted during the design phase of BTE suggested that recognized physicians would not only deliver higher quality care but also would do so at lower total cost, ours is the first formal analysis comparing the practice patterns of recognized physicians with those of their peers using a standardized methodology that relied on a large claims database. Our findings suggest that recognition by BTE is associated with systematically better performance on claims-based measures of quality. More-

Table 3. Comparison of Claims-Based Quality Measures for Physicians Recognized by Bridges to Excellence and Nonrecognized Physicians^{a,b}

Measure (No. of Physicians in Sample)	Administrative Data, % (95% CI ^c)		Difference Between Recognized and Nonrecognized Physicians, % (95% CI ^c)
	Recognized Physicians	Nonrecognized Physicians	
POL study group			
Cervical cancer screening (3233)	88.9 (88.1, 89.7)	85.0 (84.6, 85.3)	3.9 (3.1, 4.8)
Mammography (2464)	88.1 (87.2, 89.1)	85.9 (85.5, 86.4)	2.2 (1.2, 3.2)
A1C testing (1469)	86.3 (84.8, 87.9)	82.3 (81.7, 83.0)	4.0 (2.4, 5.6)
Lipid panel: coronary heart disease (206)	89.6 (86.5, 92.6)	86.4 (84.6, 88.1)	3.2 (−0.2, 6.6)
Lipid panel: hypertension (402)	45.1 (41.6, 48.6)	43.8 (41.9, 45.8)	1.3 (−2.8, 5.4)
DCL study group			
A1C testing (1059)	89.6 (87.6, 91.6)	84.1 (83.5, 84.8)	5.5 (3.4, 7.6)
Diabetic retinal exam (1059)	99.2 (98.6, 99.7)	98.2 (98.1, 98.4)	0.9 (0.3, 1.5)
Lipid panel: diabetes (1059)	87.1 (85.2, 89.1)	82.1 (81.3, 83.0)	5.0 (2.9, 7.1)
Microalbumin screening (1059)	76.2 (72.6, 79.8)	58.5 (57.2, 59.7)	17.7 (14.0, 21.4)

A1C indicates glycosylated hemoglobin; CI, confidence interval; DCL, Diabetes Care Link; POL, Physician Office Link.
^aDerived from authors' analysis of Massachusetts Group Insurance Commission data.
^bTo be included in the analysis, physicians were required to have at least 10 patients eligible for a given measure.
^cConfidence intervals were based on generalized linear models controlling for physician specialty and patient characteristics (mean patient age, percent female, case mix).

Table 4. Share of Standardized Expenditures by Category of Service for Physicians Recognized by Bridges to Excellence Compared With Nonrecognized Physicians^a

Measure	Physician Office Link, % (95% CI ^b)			Diabetes Care Link, % (95% CI ^b)		
	Recognized Physicians	Nonrecognized Physicians	Difference Between Recognized and Nonrecognized Physicians	Recognized Physicians	Nonrecognized Physicians	Difference Between Recognized and Nonrecognized Physicians
Management	35.8 (34.8, 36.7)	32.3 (32.0, 32.7)	3.4 (2.4, 4.4)	31.8 (30.3, 33.3)	29.8 (29.3, 30.2)	2.0 (0.5, 3.6)
Surgery	1.8 (1.3, 2.3)	1.6 (1.4, 1.7)	0.2 (−0.3, 0.7)	0.6 (0.5, 0.8)	0.9 (0.8, 1.0)	−0.3 (−0.5, −0.1)
Facility (room and board)	14.3 (12.6, 16.0)	15.9 (15.4, 16.4)	−1.6 (−3.4, 0.02)	14.0 (11.3, 16.7)	15.8 (15.1, 16.5)	−1.8 (−4.7, 1.2)
Inpatient ancillary	0.1 (0.0, 0.1)	0.2 (0.1, 0.2)	−0.10 (−0.1, −0.03)	0.1 (0.03, 0.2)	0.2 (0.1, 0.2)	−0.02 (−0.1, 0.1)
Outpatient	18.5 (17.8, 19.1)	19.5 (19.2, 19.8)	−1.00 (−1.7, −0.3)	16.9 (15.8, 18.0)	17.8 (17.4, 18.1)	−0.8 (−2.0, 0.3)
Prescription drugs	29.0 (27.7, 30.2)	30.5 (30.1, 31.0)	−1.6 (−2.8, −0.3)	36.7 (34.8, 38.7)	35.6 (35.1, 36.1)	1.1 (−1.0, 3.2)

CI indicates confidence interval.
^aDerived from authors' analysis of Massachusetts Group Insurance Commission data.
^bConfidence intervals were based on generalized linear models controlling for physician specialty and patient characteristics (mean patient age, percent female, case mix).

over, recognized physicians appear to use a different style of practice than their peers, emphasizing patient management over ancillary services and procedures. Patterns of relative resource use for recognized versus nonrecognized physicians were more varied. Possibly because of better management, primary care physicians recognized by the POL had nearly

20% lower resource use per episode with approximately 5% fewer episodes per patient. The DCL-recognized primary care physicians had slightly lower resource use per episode for their patients, but DCL-recognized endocrinologists did not. Both primary care physicians and endocrinologists recognized by the DCL, however, had more episodes per

■ **Table 5.** Differences in Numbers of Episodes and Costs per Episode for Primary Care Physicians and Endocrinologists Recognized by Bridges to Excellence, Compared With Nonrecognized Physicians^a

Specialty	Physician Office Link (95% CI ^b)			Diabetes Care Link (95% CI ^b)		
	Recognized Physicians	Nonrecognized Physicians	Difference Between Recognized and Nonrecognized Physicians	Recognized Physicians	Nonrecognized Physicians	Difference Between Recognized and Nonrecognized Physicians
Primary care providers						
Average number of episodes/patient	2.09 (2.08, 2.09)	2.22 (2.22, 2.23)	-0.14 (-0.15, -0.13)	2.61 (2.58, 2.64)	2.44 (2.44, 2.45)	0.17 (0.14, 0.20)
No. of episodes	191,527	1,114,334		25,080	577,278	
Average standardized resource use/episode	\$570 (\$561, \$579)	\$700 (\$697, \$703)	-\$130 (-\$140, -\$119)	\$623 (\$602, \$643)	\$649 (\$645, \$653)	-\$26 (-\$47, -\$5)
Endocrinologists						
Average number of episodes/patient				1.66 (1.62, 1.70)	1.58 (1.55, 1.61)	0.08 (0.03, 0.13)
No. of episodes				2980	4037	
Average standardized resource use/episode				\$2671 (\$2556, \$2786)	\$2534 (\$2404, \$2664)	\$137 (-\$45, \$320)

CI indicates confidence interval.

^aDerived from authors' analysis of Massachusetts Group Insurance Commission data.

^bConfidence intervals were based on generalized linear models controlling for physician specialty and patient characteristics (mean patient age, percent female, case mix).

patient than nonrecognized physicians in their respective specialties.

Bridges to Excellence offers a potential complement to claims-data profiling of physicians for identifying and rewarding high-quality care. The advantages of this measurement approach derive primarily from its ability to capture clinical information and the associated greater credibility with physicians. Our analysis inadvertently highlighted another advantage compared with claims-based profiling methods as well. Even though the GIC's data cover 6 major payers, for some quality measures a segment of physicians could not be profiled because of an insufficient number of patients in the denominator (even using a relatively low cut-off of 10 patients per measure). By design, however, the BTE chart review approach can be applied to virtually all practices because it considers the universe of a practice's patients. Making "all-payer" data truly include all payers—including Medicare and Medicaid—and standardizing physician identifiers to maximize linkages among payers would reduce this advantage and dramatically increase the usefulness of claims data for profiling physicians, however.

The primary limitations of the BTE model for profiling physicians are twofold. First, because assessments are voluntary, BTE may only get the attention of already high-performing physicians and do little to encourage improvement

among the least prepared. Second, the collection and auditing of both the clinical and site survey data are expensive, particularly for small practices, relative to analysis of existing administrative data.

The association between POL recognition and lower resource use is consistent with findings from the actuarial model underpinning BTE. Each BTE program had been originally subjected to an actuarial analysis to determine ex ante whether cost savings would be realized by recognized physicians, and it is from these analyses that financial reward targets also were established (eg, \$80 yearly bonus per diabetic patient). Although we are unable to say at this point whether the relationship between recognition status and lower resource use is a causal one (such that more recognition will increase cost-efficiency), the knowledge of this association is important for payers and consumers who are seeking ways to increase the value obtained in the healthcare system. In particular, moving patients to these physicians might increase value, although this result may be limited to the extent that unmeasured heterogeneity of patients treated by recognized versus nonrecognized physicians is affecting the results. If nonrecognized physicians responded to such a loss of market share by adopting similar practice patterns and improving the quality and efficiency of their practice, the value of care delivered throughout the healthcare system would improve.

Our findings should be viewed in light of several limitations. First, our analysis is cross-sectional in nature, because time-series data were not available. Thus, we cannot make any claims about causal relationships between the BTE recognition program and quality or between quality and resource use. Second, because chart-review data were available only for recognized physicians, we did not have access to clinical data in our analysis for risk adjustment. Although our comparisons were adjusted for patient age and sex, and an indicator of case mix, residual patient heterogeneity correlated with recognition status may be confounding our findings. This is a particular issue for the resource use comparisons. Finally, we chose to examine Massachusetts physicians because of the availability of a large multipayer database, but our results may not generalize to markets where BTE might be launched later. Factors such as overall performance on quality measures, the organization of practice, and the intensity of use, all of which vary regionally, may be important mediators of the relationships we describe.

Although the evidence of impact is mixed,¹⁴⁻¹⁷ pay-for-performance programs like BTE have engaged payers and providers in an important dialogue about how to improve the quality and cost-efficiency of care. Our analysis demonstrated a positive relationship between BTE recognition and population-based measures of quality. It remains unknown, however, whether the program motivated physicians to reorganize their practices or adopt new treatment patterns. Although providing a means for high-performing physicians to identify themselves to payers and patients may be a worthwhile goal in itself, it also is critical to understand whether BTE and similar programs create a business case for poor performers to change. Diffusion of BTE's program in markets around the country should create opportunities to address this question.

Author Affiliations: From the Department of Health Policy and Management, Harvard School of Public Health (MBR, ADS), Boston, MA; Bridges to Excellence (FSD), Newtown, CT; Mercer Health Benefits (MF), Cleveland, OH; Mercer Health Benefits (RDR), Norwalk, CT; and Health Data Management Solutions (SY), Shaker Heights, OH.

Funding Source: This work was supported by general institutional funds.

Author Disclosure: The authors (MBR, ADS, MF, RDR, SY) report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article. Mr de Brantes is the CEO of Bridges to Excellence, the organization whose pay-for-performance programs are discussed in this study.

Authorship Information: Concept and design (MBR, FSD, MF); acquisition of data (FSD, MF); analysis and interpretation of data (MBR, ADS, MF, RDR, SY); drafting of the manuscript (MBR, FSD, ADS, RDR); critical revision of the manuscript for important intellectual content (MBR, FSD,

Take-away Points

Bridges to Excellence (BTE) rewards physicians based on the results of site surveys and retrospective review of clinical data. Physicians recognized as high quality performed better than their peers on claims-based measures of quality and, in some cases, resource use measures.

- Physicians recognized by BTE appear to rely more on evaluation and management, and less on tests and procedures.
- Higher performance for recognized physicians may have predated the recognition program.
- Further research is needed to examine whether the recognition by BTE is associated with patient flow or performance improvement.

RDR); statistical analysis (MBR, ADS); administrative, technical, or logistic support (MBR); and supervision (MBR).

Address correspondence to: Meredith B. Rosenthal, PhD, Department of Health Policy and Management, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115. E-mail: meredith_rosenthal@harvard.edu.

REFERENCES

1. **Kuhmerker K, Hartman T.** *Pay-for-Performance in State Medicaid Programs: A Survey of State Medicaid Directors and Programs.* New York: The Commonwealth Fund; 2007.
2. **Trude S, Au M, Christianson JB.** Health plan pay-for-performance strategies. *Am J Manag Care.* 2006;12(9):537-542.
3. **Rosenthal MB, Landon BE, Normand SL, Frank RG, Epstein AM.** Paying for pay-for-performance in commercial HMOs. *N Engl J Med.* 2006;355(18):1895-1902.
4. **Epstein AM, Lee TH, Hamel MB.** Paying physicians for high-quality care. *N Engl J Med.* 2004;350(4):406-410.
5. **Rosenthal MB, Ferdinandopulle R, Song HR, Landon B.** Paying for quality: providers' incentives for quality improvement. *Health Aff.* 2004;23(2):127-141.
6. **Bridges to Excellence.** BTE research and analysis. <http://www.bridgestoexcellence.org/Content/ContentDisplay.aspx?ContentID=81>. Accessed March 31, 2008.
7. **Bridges to Excellence Web site.** <http://www.bridgestoexcellence.org>. Accessed November 30, 2007.
8. **Robert Graham Center.** *The Patient Centered Medical Home. History, Seven Core Features, Evidence and Transformational Change.* November 2007. <http://www.adfammed.org/documents/grahmcentermedicalhome.pdf>. Accessed March 31, 2008.
9. **National Committee for Quality Assurance.** Physician Practice Connections. <http://www.ncqa.org/tabid/141/Default.aspx>. Accessed May 8, 2008.
10. **National Committee for Quality Assurance.** Diabetes Physician Recognition Program. <http://www.ncqa.org/tabid/139/Default.aspx>. Accessed March 31, 2008.
11. **Ingenix Inc.** *Symmetry Episode Treatment Groups. A Condition Classification and Episode Building System.* White Paper. 2006. http://www.ingenix.com/content/attachments/ETG%206.0%20White%20Paper_01-17-07.pdf. Accessed May 7, 2008.
12. **Hall P, Horowitz JL.** Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica.* 1996; 64(4):891-916.
13. **The Leapfrog Group Web site.** <http://www.leapfroggroup.org/home>. Accessed November 30, 2007.
14. **Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S.** Does pay-for-performance improve the quality of health care? *Ann Intern Med.* 2006;145(4):265-272.
15. **Lindenauer PK, Remus D, Roman S, et al.** Public reporting and pay for performance in hospital quality improvement. *N Engl J Med.* 2007;356(5):488-496.
16. **Epstein AM.** Paying for performance at the tipping point. *N Engl J Med.* 2007;356(5):515-517.
17. **Young GJ, Meterko M, Beckman H, et al.** Effects of paying physicians based on their relative performance for quality. *J Gen Intern Med.* 2007;22(6):872-876. ■