

The Value of Ambulatory Care Measures: A Review of Clinical and Financial Impact from an Employer/Payer Perspective

Francois de Brantes, MBA; Paula S. Wickland, FSA, and John P. Williams, MD

To understand the value for payers and purchasers of primary care quality measures in an insured population, we conducted a 2-part analysis. In the first part, we reviewed the economic and clinical literature supporting 62 quality metrics spanning primary care that had been proposed for use in a physician recertification program and in a pay-for-performance program. We then ranked these metrics by both economic and clinical evidence of effectiveness. For many of the metrics, there was little clinical or economic support for inclusion in a pay-for-performance program. For the 20 with both clinical and economic evidence of effectiveness, we constructed actuarial models to understand the potential financial effect that attainment of these metrics would have in an insured population, from the perspective of a payer. Of those, 16 were found to be cost-saving in the short term with respect to direct medical costs incurred by payers. This analysis suggests that many recommended primary care quality measures may have little clinical evidence of effectiveness beyond expert opinion, and may provide scant clinical or economic benefit to payers if achieved. A minority, however, may deliver substantial savings in the short term. Given the current emphasis on pay-for-performance and pay-for-reporting programs, and recent studies showing a lack of relationship between measures and clinical/economic value, this analysis informs payers, purchasers, providers, and policymakers about the importance of choosing the right metrics and the methods for collecting them.

(Am J Manag Care. 2008;14(6):360-368)

For author information and disclosures, see end of text.

In 2007, for the first time in its history, the Medicare program tied a portion of a scheduled increase in physician fees to performance on a standard set of ambulatory care measures. This change in reimbursement strategy was prompted by (1) a recognition that measuring the value of Medicare physician spending has been, and continues to be, elusive; (2) a strong private sector movement to tie a portion of physician payment to demonstrated performance in delivering quality care; and (3) an acknowledgment that consumers deserve transparent information on the competence of physicians to meet certain quality thresholds.

As the Centers for Medicare & Medicaid Services (CMS) collects and disseminates these performance data, and as more than 100 similar efforts germinate in the private sector,¹ there is a paucity of robust studies on the relationship between the achievement of ambulatory care measures and healthcare cost and quality. Prior research has shown a link between performance measures and costs and quality of care.²⁻⁴ In other related articles, physicians who received recognition by the National Committee for Quality Assurance (NCQA) for demonstrating good outcomes in the management of patients with diabetes were shown to have lower costs.⁵⁻⁸ These studies are consistent with other studies that demonstrate similar results.⁹⁻¹¹ Their common denominator is the observation that a true measure of output is needed to compare the values created (or not created) by the care delivery process.

Output measures are best defined as those that most closely relate to the outcome of a patient's care, or that have the highest correlation with that outcome. For example, an important outcome for a patient with diabetes is to avoid complications such as amputation, myocardial infarction, and renal failure. The measures that are most closely related to the avoidance of these events are the proper management of the patient's glycosylated hemoglobin (A1C), low-density lipoprotein cholesterol (LDL-C), and blood pressure. Similarly, recent studies on the management of patients with cardiac disease demonstrate the importance of monitoring and measuring blood pressure.¹²

In a 2-part study, we reviewed 62 ambulatory care measures proposed for a specialty organization's recertification program and for a pay-for-performance initiative. These measures were selected by an expert panel, and 50 of them were endorsed by the National Quality Forum (NQF), the Ambulatory care Quality Alliance (AQA),

In this issue

Take-away Points / p367

www.ajmc.com

Full text and PDF

Web exclusive

e-Appendix Tables A and B

Value of Ambulatory Care Measures

and/or the NCQA. The measures span primary care, including coronary artery disease (CAD), heart failure (HF), diabetes mellitus, osteoarthritis, asthma, major depression, hypertension, and acute-care conditions. **eAppendix Table A** lists the metrics and their endorsement status (available at www.ajmc.com). The first part of the study consisted of ranking each measure according to an index that combined clinical and economic value, and the second part consisted of conducting detailed actuarial analyses of the subset of measures that had the highest index score.

Our findings imply that many payers, including CMS, should carefully consider what measures to focus on.

METHODS

To understand the benefit of each measure, we conducted a clinical and economic literature review, emphasizing meta-analyses demonstrating support for the measures. Given the preponderance of meta-analyses in our review, we captured a very large number of peer-reviewed articles. **eAppendix Table B** presents a review of the articles (available at www.ajmc.com). After we assembled the evidence for the measures, we created a point-based ranking system for both the clinical and the economic value of each measure. In basing our ranking systems on well-known methods published in the literature, our intent was to use an approach for capturing clinical and economic value that had been independently validated and was completely transparent. However, it is possible that our clinical and economic ranking systems, although comprehensive, did not capture all the elements of clinical and economic value that might be contained in a quality measure.

For the clinical evidence rankings, we used a methodology adapted from that of the GRADE Working Group.¹³ The GRADE Working Group is an international collaboration that has critiqued the assortment of evaluation tools used to rate clinical guidelines and has generated a standardized evaluation process.¹⁴ Quality of evidence was scored on a 5-point scale based on the study design for the supporting evidence:

- Meta-analysis in support—5 points.
- Multiple randomized controlled trials in support—4 points.
- Single randomized controlled trial in support—3 points.
- Observational studies only in support—2 points.
- Expert opinion in support—1 point.

Scores were reduced if there were questions of study quality, consistency, bias, directness, or imprecise/sparse data as follows:

- Serious limitations in study quality (–1).
- Important inconsistencies (–1).
- High probability of reporting bias (–1).
- Major concern about directness (ie, how does the outcome studied in the evidence align with the measure's outcome?) (–1).
- Imprecise or sparse data (–1).

Conversely, scores were increased if there was evidence of strong association or dose response according to the following schema:

- Significant evidence of a strong association between measure and outcome (relative risk or odds ratio of >2 for morbidity or mortality outcome) (+1).
- Very significant evidence of a strong association between measure and outcome (relative risk or odds ratio of >5 for morbidity or mortality outcome) (+2).
- Evidence of a dose response gradient (+1).

As a result of this scoring, the maximum number of points awarded to any measure for clinical effectiveness in our analysis was 6.

For the economic ranking system, we adapted the method of Chiou et al.¹⁵ Points were first allocated on the basis of strength of evidence with:

- More than 1 study showing evidence of cost savings—4 points.
- More than 1 study showing evidence of cost-effectiveness or cost utility at <\$50,000 per life-year saved; or 1 study showing cost savings in some scenarios—3 points.
- One study showing evidence of cost-effectiveness or cost utility at <\$50,000 per life-year saved—2 points.
- No published cost studies—1 point.

Scores were increased or decreased based on the following questions applied to the highest-scoring individual evidence:

- Was uncertainty handled by (1) statistical analysis to address random events and (2) sensitivity analysis to cover a range of assumptions? Yes +.5. No –.5.
- Were the perspective of the analysis (eg, societal, third-party payer) and reasons for its selection evident? Yes +.5. No –.5.
- Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described? Yes +.5. No –.5.

As a result of this scoring scheme, the maximum number of points allocated for financial effectiveness was 5.5; therefore, the maximum number of points for the total combined score was 33, which represents the product of the clinical and

economic scores (Table 1). The primary reason for using a product-based combined score was to numerically highlight the metrics that have been the subject of rigorous studies of both clinical and economic effectiveness. Moreover, a combined ranking based on the product of the separate clinical and economic scores provides a more balanced index, and avoids assigning undue weight, for example, to measures with strong clinical effectiveness scores but weak economic value, or vice-versa. Figure 1 illustrates the Pareto-like distribution of metrics by total combined points, where 19 metrics received 20 or more points and the remainder of the metrics received an average of 3 points or less.

We then performed a cost-benefit calculation using the measures with the highest combined rankings, because the metrics with low scores had little or no evidence of economic and clinical effectiveness. The actuarial models assessed the value of reductions in adverse outcomes when high-scoring metrics were achieved. To generate each model, we calculated the per capita benefits of treatment by determining the number, type, and average cost of morbidity events prevented by attainment of each metric, as determined from the literature and validated through the Thomson Medstat MarketScan database (Thomson Medstat Inc, Ann Arbor, Michigan), a large integrated claims database of commercially insured employees of mainly large corporations. See Figure 2 and Figure 3 for specific examples. When cost figures were outdated, we inflated them to 2006 levels using the medical Consumer Price Index. We assumed study populations were 50% male and 50% female, and where ethnicity was relevant (for the cholesterol and hypertension models), we assumed the population was 90% white and 10% black.

We next calculated per capita costs, using average costs for generic versions of pharmacotherapy treatments (where available) obtained from an online Internet pharmacy and including other related medical costs from an amalgam of likely therapies. The cost of medication side effects in these particular applications was generally not considered, with 2 exceptions: aspirin use and switching to angiotensin receptor blockers (ARBs) because of intolerance to angiotensin-converting enzyme (ACE) inhibitors. Although the incremental cost of medication used for treatment was included in the model, the incremental cost of physician time to prescribe these treatments was not considered. None of the interventions listed here would lead to codable procedures, although it is conceivable that they could increase the acuity of individual visits. (For example, a level 2 or 3 visit might be justifiably “upcoded” to a level 3 or 4.)

After we derived the per capita benefits and costs of treatment, we summed them to yield the net financial effect—savings or cost—of the specific quality measure. The actuarial

models were conservative, considering only the direct medical cost of morbidity to employers/payers for patients less than 65 years of age, in a 1-year time frame. Based on the literature, we varied the specific morbidity effects for each measure. In the case of hypertension, for example, the literature documents reductions in end-stage renal disease (ESRD), CAD, and stroke. In the case of ACE inhibitor/ARB treatment for left ventricular systolic dysfunction (LVSD), there was a reduction in hospitalizations for congestive HF. There were additional morbidity effects that could be expected with each metric, such as the decrease in retinopathy with blood pressure reduction.²⁴

This focus on direct medical costs to the employer/payer meant that we did not include several major elements of the cost of care in the savings that we estimated. We ignored any medical costs directly paid by patients as well as the indirect and intangible costs to employers or patients. Additionally, we imputed savings only for the reductions in complications and other health factors that would occur prior to age 65 years and excluded the cost of mortality. (Implementation costs were excluded as there was no consensus on what they would be. The model was constructed, however, so that these costs could be easily accounted for when determined.) Finally, we assumed full patient compliance. Although we had no basis to assume full compliance and in fact had evidence to the contrary, there were no data to suggest how compliance might vary by condition and medication. Nevertheless, the model was developed so that the compliance factor could be changed by user preference, with reductions in compliance therefore affecting achievement of the outcome.

Clearly, there are a number of limitations to our methodology. First, preventing mortality is an important goal of payers and employers, and mortality effects were not considered in our models because we could not define an acceptable method of valuing life-years saved. Second, employers have reasons to care about direct costs and indirect costs, yet we ignored indirect costs because of the difficulties in adequately measuring them. Although placing these limitations on the model may seem overly conservative by reducing the net potential savings, the goal was to capture only the direct medical costs that would be paid by employers and payers, which most payers and purchasers use as a primary means to determine the value of a program.

RESULTS

From the list of 62 metrics, only 20 (each of which is endorsed by the NQF, AQA, or NCQA) received high combined scores for clinical and economic support in our ranking

Value of Ambulatory Care Measures

Table 1. Clinical Effectiveness, Cost-effectiveness, and Combined Scores by Metric

Measure	Clinical Effectiveness Score	Cost-effectiveness Score	Combined Score
Blood pressure <140/90 mm Hg (HTN)	6	5.5	33
Systolic blood pressure <140 mm Hg (HTN)	6	5.5	33
Diastolic blood pressure <90 mm Hg (HTN)	6	4.5	27
Blood pressure <140/90 mm Hg (DM)	6	4.5	27
A1C >9% (DM)	6	4.5	27
A1C <7% (DM)	6	4.5	27
LDL-C <100 mg/dL (DM)	6	4.5	27
LDL-C <130 mg/dL (DM)	6	4.5	27
LDL-C <100 mg/dL after discharge for AMI, CABG, or PCI (CAD)	6	4.5	27
LDL-C <130 mg/dL after discharge for AMI, CABG, or PCI (CAD)	6	4.5	27
LDL-C <100 mg/dL with any CAD	6	4.5	27
LDL-C <130 mg/dL with any CAD	6	4.5	27
Weight reduction (HTN)	5	4.5	23
BB use in HF	5	4.5	23
ACE inhibitor/ARB use in LVSD (HF)	5	4.5	23
BB post-MI with prescription 7 days after discharge (CAD)	5	4.5	23
BB post-MI with prescription 6 months after discharge (CAD)	5	4.5	23
Antiplatelet therapy in CAD—aspirin only (CAD)	5	4.5	23
ACE inhibitor/ARB in CAD with LVSD (CAD)	5	4.5	23
Back pain—bed rest >4 days (ACCs)	5	3.5	18
Upper respiratory infection—no use of antibiotics (ACCs)	5	3.5	18
UTI—antibiotic use for <7 days (ACCs)	5	3.5	18
Use of warfarin in HF and atrial fibrillation	4	3.5	14
Retinal eye screening (DM)	4	2.5	10
Microalbuminuria screening past year (DM)	4	1	4
Microalbuminuria screening (DM)	4	1	4
Use of appropriate medications in AST	4	1	4
Pharmacologic therapy (AST)	4	1	4
Effective acute-phase treatment in DEP	4	1	4
Decongestant used <4 days in nasal congestion (ACCs)	4	1	4
Effective continuation-phase treatment (DEP)	3	1	3
Assessment of volume overload (HF)	3	1	3
History taking for “red flags” in ALBP	2	1	2
History taking for cauda equina symptoms in ALBP	2	1	2
ALBP—physical exam	2	1	2
Optimal practitioner contacts (DEP)	2	1	2
Foot examination (DM)	2	1	2
Measurement of serum creatinine (HTN)	2	1	2
Avoiding contraindicated medications (ALBP)	1	1	1
Avoiding contraindicated physical treatments (ALBP)	1	1	1

(Continued)

■ **Table 1.** Clinical Effectiveness, Cost-effectiveness, and Combined Scores by Metric (*Continued*)

Measure	Clinical Effectiveness Score	Cost-effectiveness Score	Combined Score
Classification of asthma (AST)	1	1	1
Lung function testing (AST)	1	1	1
Influenza vaccination for asthma (AST)	1	1	1
Lipid profile (CAD)	1	1	1
LDL-C drug therapy (CAD)	1	1	1
Evaluation of activity level and anginal symptoms (CAD)	1	1	1
Annual A1C (DM)	1	1	1
Lipid profile (DM)	1	1	1
Assessment of left ventricular function (HF)	1	1	1
Weight measurement within last 6 months (HF)	1	1	1
Weight measurement (HF)	1	1	1
Activity level documentation (HF)	1	1	1
Plan of care (HTN)	1	1	1
Measurement of serum potassium (HTN)	1	1	1
Electrocardiogram (HTN)	1	1	1
NSAID and analgesic use in OA	1	1	1
Assessment of function and pain (OA)	1	1	1
History taking for knee pain (pain)	1	1	1
UTI—documentation of flank pain, fever, and dysuria (ACCs)	1	1	1
UTI—documentation of vaginal discharge (ACCs)	1	1	1
UTI—Bactrim as first-line therapy (ACCs)	1	1	1
Sodium restriction (HTN)	1	1	1

HTN indicates hypertension; DM, diabetes mellitus; A1C, glycosylated hemoglobin; LDL-C, low-density lipoprotein cholesterol; AMI, acute myocardial infarction; CABG, coronary artery bypass graft; PCI, percutaneous coronary intervention; CAD, coronary artery disease; BB, β -blocker; HF, heart failure; ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; LVSD, left ventricular systolic dysfunction; MI, myocardial infarction; ACCs, acute-care conditions; UTI, urinary tract infection; AST, asthma; DEP, depression; ALBP, acute lower back pain; NSAID, nonsteroidal anti-inflammatory drug; OA, osteoarthritis.

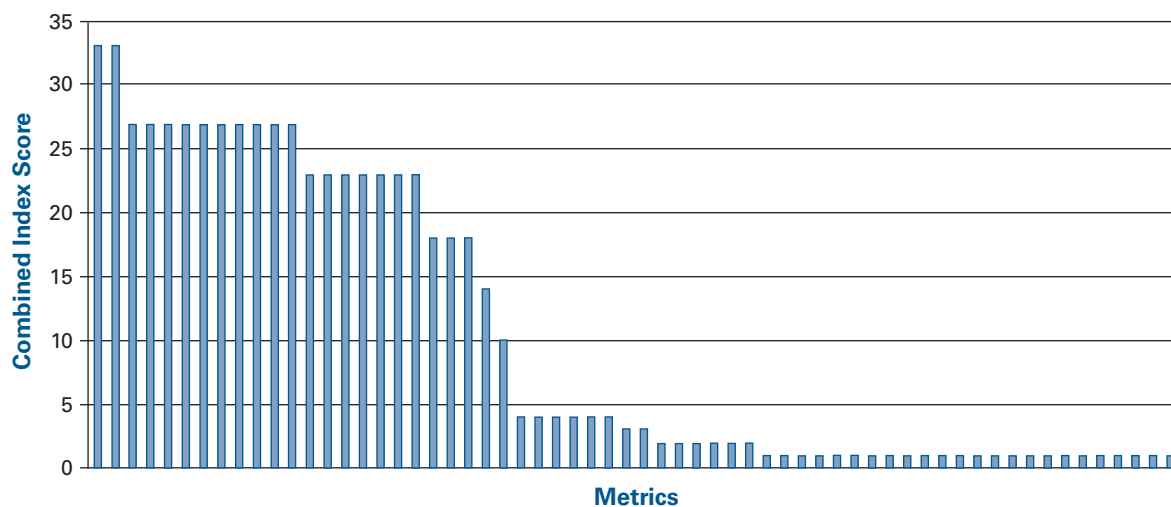
scheme. Of these 20, most were shown to be cost-saving in actuarial modeling based on the conservative assumptions described above (**Table 2**).

Approximately one third of the 62 metrics lacked conclusive clinical evidence to support their inclusion in pay-for-performance beyond expert opinion, even though all of these metrics are considered part of the standard of care in medical practice. In addition, the measures lacking supportive clinical evidence highly correlated with those lacking evidence of economic savings (data not shown). The low-scoring measures did share one important element—they tended to be process measures with distant relationships to outcomes. Although practices such as taking the patient’s medical history and performing a physical are time-honored parts of the

medical evaluation and may be prerequisites to interventions of proven clinical or economic benefit, they do not reduce morbidity or mortality directly—which is the measure of output that is important to payers and purchasers—and as a result have little or no actuarial value.

Although we limited the scope of potential savings to direct medical costs for patients under age 65 years, we found that most of the quality metrics with the strongest clinical and economic evidence of effectiveness were cost-saving in the 1-year frame of analysis. The range of savings was \$88 per patient per year for achieving systolic blood pressure of less than 140 mm Hg to \$781 per patient per year for use of ACE inhibitors/ARBs in LVSD or LVSD with CAD. In contrast, an LDL-C value below 130 or below 100 mg/dL was not cost-sav-

■ **Figure 1.** Distribution of Metrics by Combined Score



ing, yielding a net cost per patient per year of \$429 and \$412, respectively (Figure 3 and Table 2). In the case of LDL-C reduction, the absence of cost-savings was driven by the high cost of statin therapy and, to a lesser degree, by the relatively lower effect of reduced LDL-C (as opposed to blood pressure) on morbidity. The growing availability of generic statin therapy will likely change this valuation.

The short-term savings in the care of a working-age population related to compliance with these metrics also implies that their value would be far greater for CMS. In addition, because the cost of treatment needed to achieve high performance on some of these metrics is very small (eg, generic medication), and the benefits of complication avoidance grow with the time frame for analysis, a longer-term perspective would yield far greater net savings.

One important limitation of this study is that the actuarial models used to develop the net savings associated with each metric considered every metric independently of the other, and the results are not additive. For example, the sav-

■ **Figure 2.** Estimated Savings From Blood Pressure Management per Adult Age 18-64 Years With Hypertension (HTN)

Event:	ESRD	CHD	Stroke	Total
Absolute risk reduction per 100	0.812 ^a	0.095 ^b	0.043 ^c	
Event cost	\$22,128 ^d	\$9885 ^e	\$14,562 ^f	
Savings per person before cost of intervention	\$180	\$9	\$6	\$195
Intervention cost				\$107 ^g
Treatment compliance rate				100%
Savings per patient				\$88

ESRD indicates end-stage renal disease; CHD, coronary heart disease.

^aSources are references 16 and 17.

^bSources are references 16 and 18.

^cSource is reference 16.

^dSources are references 17, 19, and 20.

^eMedStat 2005 data inflated to 2006 values.

^fMedStat 2005 data inflated to 2006 values.

^gTreatment was a straight average of supply costs for enalapril 10 mg daily, hydrochlorothiazide 25 mg daily, and metoprolol 25 mg 3 times a day.

ings associated with the treatment of patients with CAD who have had a prior myocardial infarction cannot be derived simply by adding the savings associated with each CAD measure.

DISCUSSION

The discord that often exists between expert opinion and demonstrated clinical and economic evidence has important implications for pay-for-performance and pay-for-reporting

■ **Figure 3.** Estimated Savings From Management of Low-density Lipoprotein Cholesterol (LDL-C) for Patients With any Heart Disease

Event:	CAD	Stroke	Total
Absolute risk reduction per 100 per year: LDL-C <100 vs >100 mg/dL	2.494 ^a	0.0259 ^b	
Event cost	\$9885 ^c	\$14,562 ^d	
Savings before cost of intervention	\$246	\$4	\$250
Intervention cost			\$662 ^e
Treatment compliance rate			100%
Savings per patient			\$(412)

Event:	CAD	Stroke	Total
Absolute risk reduction per 100 per year: LDL-C <130 vs >130 mg/dL	2.321 ^a	0.0241 ^b	
Event cost	\$9885 ^c	\$14,562 ^d	
Savings before cost of intervention	\$229	\$4	\$233
Intervention cost			\$662 ^e
Treatment compliance rate			100%
Savings per patient			\$(429)

CAD indicates coronary artery disease.

^aSources are references 18 and 21.

^bSources are references 22 and 23.

^cMedStat 2005 data inflated to 2006 values.

^dMedStat 2005 data inflated to 2006 values.

^eCost is derived from the price of generic simvastatin used once daily.

projects. Without subjecting quality measures to the kind of analysis we conducted, policymakers risk creating a range of mandated measures that will not reduce healthcare costs, or even be neutral, but instead may be cost-additive. This is particularly relevant in programs that allow physicians to select a handful of metrics from a basket of metrics, or for programs that rely exclusively on measures that are not directly related to the reduction of morbidity or mortality of patients.

There are nascent measurement programs that emphasize the importance of outcomes. For example, the Minnesota Community Measurement collaborative has designed a performance assessment for practices relating to the care given to patients with diabetes that is limited to 5 measures, all of which are tightly linked to significant reductions in the risk of the patient experiencing an adverse event.²⁵ In these measurement systems, process measures are combined with intermediate and/or full outcome measures in a composite that

provides a very different analysis of performance than that achieved when each measure is considered as an independent variable. Essentially, the performance assessment process in the Minnesota Community Measurement collaborative explicitly combines metrics that have economic validity with metrics that don't but which still have clinical validity, thus still yielding a net positive economic impact for payers and purchasers.

POLICY IMPLICATIONS

If measuring outcomes or processes tightly linked with outcomes can result in cost-savings, why aren't payers focusing on these measures? An important reason is that there is an inherent difficulty in systematically collecting the data to assess performance

on these measures without incurring significant data collection costs.

The most ubiquitous, standardized, and inexpensive data collection process in force in the US healthcare system is the billing process. There are hundreds of billions of claims processed each year, and they can yield valuable information. For example, the *International Classification of Diseases, Ninth Revision, Clinical Modification* code V85 is the diagnosis code for the body mass index (BMI) of adults. V85.0 is given for a BMI less than 19, V85.1 for a BMI of 19 through 24, and so on.²⁶ Yet there is no widely used Current Procedural Terminology (CPT) designation that allows coding for a specific heart rate, LDL-C level, or A1C level. (As of 2006 and 2007, category II CPT codes have been issued that will allow coding of clinical values and will facilitate performance management through coding, but these have yet to be widely adopted.) Most importantly, despite the preponderance of evi-

Table 2. Actuarial Value of Net Savings per Metric

Measure	Clinical Domain	NQF/AQA Endorsed	Combined Score	Value, \$
Blood pressure <140/90 mm Hg	HTN	Yes	33	88
Systolic blood pressure <140 mm Hg	HTN	Yes	33	88
LDLC <100 mg/dL after discharge for AMI, CABG, or PCI	AMI	Yes	27	(412)
LDLC <130 mg/dL after discharge for AMI, CABG, PCI	AMI	Yes	27	(429)
LDLC <100 mg/dL with any CAD	CAD	Yes	27	(412)
LDLC <130 mg/dL with any CAD	CAD	Yes	27	(429)
Blood pressure <140/90 mm Hg (DM)	DM	Yes	27	166
A1C >9%	DM	Yes	27	177
A1C <7%	DM	Yes	27	96
LDLC <100 mg/dL (DM)	DM	Yes	27	250
LDLC <130 mg/dL (DM)	DM	Yes	27	150
Diastolic blood pressure <90 mm Hg	HTN	Yes	27	88
BB post-MI with prescription 7 days after discharge	AMI	Yes	23	288
BB post-MI with prescription 6 months after discharge	AMI	Yes	23	288
Antiplatelet therapy in CAD—aspirin only	CAD	Yes	23	489
ACE inhibitor/ARB in CAD with LVSD	CAD	Yes	23	781
BB use in HF	HF	Yes	23	362
ACE inhibitor/ARB use in LVSD	HF	Yes	23	781

NQF indicates National Quality Forum; AQA, Ambulatory care Quality Alliance; HTN, hypertension; LDL-C, low-density lipoprotein cholesterol; AMI, acute myocardial infarction; CABG, coronary artery bypass graft; PCI, percutaneous coronary intervention; CAD, coronary artery disease; DM, diabetes mellitus; A1C, glycosylated hemoglobin; BB, β-blocker; MI, myocardial infarction; ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; LVSD, left ventricular systolic dysfunction; HF, heart failure.

dence suggesting that morbidity and mortality are impacted by even small variations in blood pressure, there is, for example, no widely used coding that can differentiate between a systolic blood pressure of 140 mm Hg and one of 130 mm Hg.²⁷ New CPT category 2 codes might help mitigate some of this deficiency, although it is unclear by how much.

The alternative is to use clinical data contained in medical records. However, the process of collecting those data generally requires the abstraction of paper medical records, which consumes significant time and resources for physicians, or the abstraction of data from electronic health records, which have been adopted by only 15% to 20% of physicians. The emergence of regional organizations to manage health information exchange holds the promise of automating clinical data collection and aggregation,²⁸ but only in the future. Yet this analysis seems to suggest that private-sector and public-sector payers would be well served to invest in the collection of these data for some measures, while continuing to rely on claims for others.

REAL-WORLD IMPLICATIONS

As payers continue to develop incentive and reward programs for physicians, and as attention is increasingly focused on the aggregation of claims data, other administrative data, and medical record data, we all should consider the selection and weighting of performance metrics carefully in light of this analysis, and determine the actual return that might be realized by the monies spent on different types of data aggregation. Payers should find ways in their incentive and reward programs to emphasize the small set of metrics that have high

Take-away Points

Our research analyzed the clinical and financial value of 60 commonly used and generally approved physician quality measures from a payer-purchaser perspective.

- Only a handful of those measures had a significant clinical and financial impact.
- However, those measures are not routinely found in claims data, thus putting to question the amount of resources that should be devoted to large claims data aggregation efforts as opposed to other data collection efforts.

economic value. Public foundations and other contributors to data aggregation should consider the importance of and return on investing in data collections where the data will yield a far greater return than the cost of collection. Purchasers should continue to encourage their third-party payers to include a specific allowance in their incentive and reward programs for the adoption and use of electronic medical records, in recognition of the importance of having a systemic way of collecting, measuring, and scoring the performance of physicians on metrics that are both clinically and economically effective.

Author Affiliations: From Bridges to Excellence, Newtown, CT (FDB); Towers Perrin, Stamford, CT (PSW); and the WDE Group, LLC, Radnor, PA (JPW).

Funding Source: Funding for this work was provided by Bridges to Excellence, a not-for-profit 501-c-3 organization.

Author Disclosures: Mr de Brantes is the CEO of Bridges to Excellence, the company that provided funding for this work. Dr Williams has consulted for Bridges to Excellence. The other authors (FDB, PSW, JPW) report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (FDB, PSW, JPW); acquisition of data (JPW); analysis and interpretation of data (FDB, PSW, JPW); drafting of the manuscript (FDB, PSW, JPW); critical revision of the manuscript for important intellectual content (FDB, JPW); administrative, technical, or logistic support (FDB); and supervision (FDB).

Address correspondence to: Francois de Brantes, MBA, Bridges to Excellence, 13 Sugar St, Newtown, CT 06470. E-mail: francois.debrantes@bridgestoexcellence.org.

REFERENCES

1. **The Leapfrog Group.** Incentives & rewards. http://www.leapfroggroup.org/for_hospitals/fh-incentives_and_rewards. Accessed February 29, 2008.
2. **Towers Perrin.** *Cardiac Care Analysis—Savings Estimates*. December 29, 2003. http://www.bridgestoexcellence.org/Documents/bte_towersperrin.pdf. Accessed February 29, 2008.
3. **Towers Perrin.** *Diabetes Care Analysis—Savings Estimates*. December 6, 2005. http://www.bridgestoexcellence.org/Documents/DCL_analysis_1207051.pdf. Accessed February 29, 2008.
4. **Ingenix.** *Evaluation of the Diabetes Care Link*. February-March 2005. http://www.bridgestoexcellence.org/Documents/DPRP_Eval_2005.pdf. Accessed February 29, 2008.
5. **de Brantes F.** Lessons learned from a national pay-for-performance program. Healthcare Technology Project, Vol 4. January 31, 2007. http://www.hctproject.com/documents.asp?d_ID=4138. Accessed February 29, 2008.
6. **de Brantes F.** Pay-for-performance and beyond: a recipe for improving healthcare. In: *The Quality Conundrum: Practical Approaches for Enhancing Patient Care*. New York: PriceWaterhouseCoopers; 2007.
7. **de Brantes F, Fabius RJ, Galvin RS.** Transparency is a prerequisite of pay for performance. *Am Heart Hosp J*. 2004;2(3):170-174.
8. **de Brantes F.** Bridges to excellence: a program to start closing the quality chasm in healthcare. *J Healthc Qual*. 2003;25(2):2, 11.
9. **Bodenheimer T, Wagner EH, Grumbach K.** Improving primary care for patients with chronic illness. *JAMA*. 2002;288(14):1775-1779.

10. **Bodenheimer T, Wagner EH, Grumbach K.** Improving primary care for patients with chronic illness: the chronic care model, part 2. *JAMA*. 2002;288(15):1909-1914.
11. **Gilmer TP, O'Connor PJ, Rush WA, et al.** Predictors of health care costs in adults with diabetes. *Diabetes Care*. 2005;28(1):59-64.
12. **Cutler D, Long G, Berndt ER, et al.** The value of antihypertensive drugs: a perspective on medical innovation. *Health Aff (Millwood)*. 2007;26(1):97-110.
13. **Atkins D, Briss PA, Eccles M, et al; GRADE Working Group.** Systems for grading the quality of evidence and the strength of recommendations, II: pilot study of a new system. *BMC Health Serv Res*. 2005;5(1):25.
14. **Atkins D, Best D, Briss PA, et al; GRADE Working Group.** Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490.
15. **Chiou CF, Hay JW, Wallace JF, et al.** Development and validation of a grading system for the quality of cost-effectiveness studies [erratum in *Med Care*. 2003;41(3):446]. *Med Care*. 2003;41(1):32-44.
16. **Burt V, Whelton P, Roccella EJ, et al.** Prevalence of hypertension in the US adult population. Results From the Third National Health and Nutrition Examination Survey, 1988-1991 [comment in *Hypertension*. 1995;25(3):303-304]. *Hypertension*. 1995;25(3):305-313.
17. **Haroun M, Jaar BG, Hoffman SC, Comstock GW, Klag MJ, Coresh J.** Risk factors for chronic kidney disease: a prospective study of 25,534 men and women in Washington County, Maryland. *J Am Soc Nephrol*. 2003;14(11):2934-2941.
18. **Wilson P, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB.** Prediction of coronary heart disease using risk factor categories [comments in *Circulation*. 1998;97(18):1761-1762 and *Circulation*. 1999;99(16):2219]. *Circulation*. 1998;97(18):1837-1847.
19. **Joyce A, Iacoviello JM, Nag S, et al.** End-stage renal disease-associated managed care costs among patients with and without diabetes. *Diabetes Care*. 2004;27(12):2829-2835.
20. **US Census Bureau.** Statistical Abstract of the United States, 2003. http://www.census.gov/prod/www/statistical-abstract-2001_2005.html. Accessed February 29, 2008.
21. **Grundy S, Cleeman J, Merz CN, et al; Coordinating Committee of the National Cholesterol Education Program.** Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III Guidelines. *J Am Coll Cardiol*. 2004;44(3):720-732.
22. **Corvol J, Bouzamondo A, Sirol M, Hulot JS, Sanchez P, Lechat P.** Differential effects of lipid-lowering therapies on stroke prevention. *Arch Intern Med*. 2003;163(6):669-676.
23. **Hu G, Sarti C, Jousilahti P, et al.** The impact of history of hypertension and type 2 diabetes at baseline on the incidence of stroke and stroke mortality [comment in *Stroke*. 2006;37(5):1150; author reply 1150]. *Stroke*. 2005;36(12):2538-2543.
24. **Yu T, Mitchell P, Berry G, Li W, Wang JJ.** Retinopathy in older persons without diabetes and its relationship to hypertension. *Arch Ophthalmol*. 1998;116(1):83-89.
25. **Minnesota Community Measurement.** Managing diabetes in Minnesota. <http://www.thed5.org>. Accessed February 29, 2008.
26. **ICD9. chrisendres.com.** Persons without reported diagnosis encountered during examination and investigation of individuals and populations (V70V85). <http://icd9cm.chrisendres.com/index.php?action=child&recordid=10981>. Accessed February 29, 2008.
27. **Chobanian AV, Bakris GL, Black HR, et al; National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure; National High Blood Pressure Education Program Coordinating Committee.** The Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560-2572. Epub 2003 May 14.
28. **de Brantes F, Emery DW, Overhage JM, Glaser J, Marchibroda J.** The potential of HIEs as infomediaries. *J Healthc Inf Manag*. 2007; 21(1):69-75. ■