Quality Implications of Efficiency-Based Clinician Profiling

Mark C. Rattray, MD; James Andrianos, MBA; Dorothy T. Stam, BA

ABSTRACT

BACKGROUND

"High performance networks" have been embraced recently by insurers and large purchasers in the effort to slow rising medical cost trends. These networks feature clinicians with favorable profiling scores, typically based on cost efficiency. An implicit belief is that cost efficiency signals quality in health care delivery, but there have been no large, rigorous studies of this assumption.

OBJECTIVE

To improve the understanding of the relationships between efficiency and quality ratings for individual clinicians.

DESIGN, SETTING, AND MEASURES

Descriptive study of performance measures derived from approximately 40 million claim lines representing a large insurer's entire commercial business during the 2 year period ending September 2002. Efficiency was defined as condition-specific cost variation over the course of patient care. Quality was defined as process adherence to selected evidence-based medical practice guidelines. Leading commercial tools were employed to assess efficiency and quality performance, respectively. For each type of performance, an index was constructed to summarize an individual clinician's composite achievement.

RESULTS

Correlation between clinicians' efficiency and quality index scores was significant ($p<0.05$) for 3 specialties, all in primary care. For these specialties, efficiency declined as quality compliance increased. Generally, the relationships were weak and do not support inferring one type of performance from the other. For all other specialties, no significant correlations emerged.

CONCLUSIONS

Clinician cost efficiency and process quality compliance are poor proxies for each other, as measured by leading commercial tools for each domain. Pending the results of additional work in this emerging topic, insurers and large purchasers should measure each practice characteristic separately and resist inferring one from the other.

QUALITY IMPLICATIONS OF EFFICIENCY-BASED CLINICIAN PROFILING

MARK C. RATTRAY, M.D
JAMES ANDRIANOS, M.B.A
DOROTHY T. STAM, B.A.‡

In health care, a number of recent studies have identified and attempted to measure a large and growing "value deficit"[1-5]. Rapidly increasing health costs amidst seemingly lackluster quality are motivating investigators to improve our understanding of underlying dynamics, in part so improvement efforts might be more focused and effective.

In the private sector, large self-insured employers are considering "high performance networks" in a bid to moderate medical cost increases. These networks feature clinicians with favorable profiling scores, typically based on cost efficiency as measured by commercial tools and methods. An implicit belief is that efficiency signals quality in health care, but this relationship has not been well studied.

This study seeks a clearer understanding of how clinician efficiency and quality ratings correspond, using the commercial measurement tools likely to be employed in private sector interventions.

‡ Mark C. Rattray, M.D., Clinical Faculty, University of Washington School of Medicine; Tel: 206-287-7295; Fax: 206-287-5493; mxrattr@regence.com.

James T. Andrianos, M.B.A., Principal, Calculated Risk, Inc.; Tel: 206-382-8575; Fax: 206-382-0581; ja_cri@msn.com.

Dorothy T. Stam, B.A., Manager, Clinical Performance Measurement, Regence BlueShield; Tel: 206-287-5973; Fax: 206-287-5844; dtstam@regence.com.

The authors recognize Zakaria Ramadan-Jradi, M.D., M.B.A., for his significant guidance and insight into clinical performance measurement.

METHOD

We examined measurable relationships between the quality and efficiency of individual clinicians. We defined quality as rates of process adherence to selected evidence-based medical practice guidelines. We examined cost in terms of efficiency, which was defined as condition-specific cost variation over the course of patient care. We relied on a single large repository of detailed claims to study in parallel the quality and efficiency performance of discrete clinicians. Using the claim repository, quality care was identified and credited to individual clinicians. The repository also permitted each patient's total medical cost to be parsed into its constituent clinical episodes of care and their corresponding costs. Clinicians were compared only to other clinicians of the same specialty.

Data Collection. We obtained 2 years of detailed claim records comprising the entire commercial "book of business" for Regence BlueShield, an independent licensee of the Blue Cross and Blue Shield Association in Washington State. The records represented the care delivered to approximately 1.1 million patients and amounted to nearly 40 million claim lines over a 2 year period ending September 30, 2002. Claims spanned both inpatient and outpatient care, whether delivered via professional, facility, or ancillary providers. All pharmacy claims were included.

Quality Evaluation Tool and Analysis. We utilized the services of Health Benchmarks, Inc. to conduct algorithmic examinations of the claim repository. The algorithms verified and measured process compliance with evidence-based medical guidelines. Algorithms for each indicator were adapted from existing published and peer-reviewed protocols comprising modern evidence-based medical guidelines.

Exhibit 1 lists the 33 process indicators for which compliance was evaluated. For each

indicator, a rate of compliance was determined for every qualifying clinician within the applicable specialty. This raw compliance rate was converted to a percentile score relative to the performance of other clinicians with the same specialty. Across indicators, all of a clinician's percentile scores were weighted into a single composite score. A clinician's quality index was his/her composite score divided by a specialty composite average. An index value of 1.0 approximated the statewide network average within a specialty, and therefore, not necessarily a performance target.

Efficiency Evaluation Tool and Analysis. Using the same claims, distinct clinical episodes of care were identified using dedicated grouping software. We licensed a clinical episode grouper from Symmetry Health Data Systems, Inc.. The grouper examines individual claim lines for clinically-related services, combining them into episodes of varying duration, depending on the appearance of pre-defined "quiet periods" in the claim sequence. The resulting episodes may be isolated, sequential, or concurrent. The grouper maintains diagnostic profiles of each patient, accounting for co-morbid conditions during Episode Treatment Group® (ETG) assignment. Each ETG approximates a clinically homogeneous classification, permitting severity adjustment when comparing episode sets of varying ETGs. There are 850 distinct ETGs.

We converted episode-of-care cost variation into relative efficiency scores for individual clinicians. For each episode, we identified the clinician most responsible for professional care, and we attributed the episode to this clinician. Clinicians with sufficient volumes of attributed episodes became eligible for efficiency comparison. For a given clinician's attributed episodes, expected resource utilization was compared to observed resource utilization, as measured by paid dollars. Paid dollars included insurer reimbursement, patient deductibles, and co-insurance. The efficiency index was the ratio of expected-to-observed episode costs. An index value of 1.0 approximated the statewide network average within a specialty, and therefore, not necessarily a performance target.

Detailed documentation of the methods used to collect, prepare, process, and analyze the claims records in this study is available upon request from the authors as a Technical Appendix[6].

RESULTS

We derived both quality and efficiency index scores for 2,271 clinicians representing 12 specialties. We examined the relationship between quality and efficiency index scores for each clinician by measuring the significance of within-specialty linear models. Exhibit 2 presents results from applying linear regression analysis to clinicians' quality and efficiency index scores by specialty.

Of 12 studied specialties, 3 models exhibited statistically significant ($p<0.05$) linear fits. Four specialties had statistically insignificant relationships.

We categorized an additional 5 specialties as lacking sufficient data, using a minimum of 25 clinician observations as a threshold.

The 7 specialties with conclusive relationships were Cardiovascular Disease, Family Practice, Family Practice (nurse practitioner), General Practice, Internal Medicine, Obstetrics/Gynecology, and Pulmonary Disease. Residuals for these specialties showed no evidence of non-linearity.

Family Practice, Family Practice (nurse practitioner), and General Practice showed significant relationships. Their results suggest quality improves as efficiency declines. Although these 3 relationships are significant, efficiency and quality appear to be poor predictors of one another (Exhibit 3).

COMMENT

This study concludes that clinician-level process adherence to evidence-based medical guidelines is largely uncorrelated with episode-of-care cost efficiency, as measured by widely used commercial tools and methods. Apparent exceptions exist in certain primary care specialties, where quality gains tend to correspond with declining efficiency. No evidence of relationships within the medical/surgical specialties was found.

Limitations and Opportunities. Our methods did not attempt to make adjustments for efficient but inappropriate treatment decisions. Unlike compliance with clinical process guidelines, the existence of identified episodes of care need not validate the appropriateness of chosen treatment options. For example, an episode featuring a surgical intervention does not necessarily signal the appropriateness of a surgical versus a non-surgical course of care.

The earliest process indicators to be converted into claim algorithms were part of the Health Plan Employer Data and Information Set (HEDIS®). A portion of the quality indicators used in this study are HEDIS® or HEDIS®-like in nature. These indicators tend to apply to primary care specialties. In the future, the number of quality indicators available as claim-based measurement algorithms should be expanded, with special attention toward greater medical/surgical specialty coverage. Another opportunity will be to assemble larger claim repositories, perhaps even spanning health insurers.

Implications. The absence or weakness of quality-efficiency relationships has implications for the "high performance network" concept. For example, within the 7 specialties found to have conclusive relationships, a select sub-network of efficient clinicians would leave out many clinicians with above average quality performance. This study suggests about one-half of excluded clinicians would have above average quality ratings relative to their specialist peers. Moreover, for the specialties with significant (negative) relationships, this fraction would be higher, on average.

Without recognized standards for measuring professional quality and efficiency, we opted to employ the tools of leading commercial vendors. These tools are widely used to assess performance in real world settings, in many cases impacting clinical and business functions. Curiously, the instances of large claim repositories being evaluated in parallel for both quality and efficiency are few. This study is an early effort in this area. Even if a deeper or alternate correlation remains unrevealed, these results suggest that measuring quality and efficiency separately with dedicated tools is a prudent and balanced approach to gauging these aspects of clinician performance.

Funding. This project was funded by Regence BlueShield, an independent licensee of the Blue Cross and Blue Shield Association. Regence BlueShield funded the design and conduct of this study; the collection, management, analysis, and interpretation of the data; and the preparation, review, and approval of the manuscript.

REFERENCES

1. *McGlynn EA et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348(26):2635-2645.*
2. *Wenger NS et al. The quality of medical care provided to vulnerable community-dwelling older patients. Ann Intern Med. 2003;139(9):740-747.*
3. *Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. Ann Intern Med. 2003;138(4):273-287.*
4. *Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. Ann Intern Med. 2003;138(4):288-298.*
5. *Baicker K, Chandra A. Medicare Spending, The physician workforce and beneficiaries' quality of care. Health Aff. 2004:7 April. Available at: http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.184v1/DC1).*
6. *Readers: see Technical Appendix included at the end of this document.*

Exhibit 1: Selected Indicators of Conformance with Evidence-based Medicine	***Source: Health Benchmarks, Inc.***
Process Indicator	Number of Applicable Specialties
Mammography screening	12
Cervical cancer screening	11
Colorectal cancer screening	15
ACE inhibitor use in CHF	11
Use of long-term control drugs in asthma	12
Diabetic retinal exam	10
Glycosylated hemoglobin (Hb A1C) for diabetics	10
Compliance with anti-hypertensive drugs	11
Compliance with lipid lowering drugs	11
ENT: Evaluation of thyroid nodule	2
Follow-up after diagnosis of prostatic cancer	1
Orthopedics: X-ray prior to Knee MRI	4
Follow-up after hospitalization for mental illness	2
Asthma follow-up after emergency department visit	2
Lipid panel following stroke	12
Treatment of Major Depression: Acute Phase Management	11
Treatment of Major Depression: Optimal Practitioner Contacts	11
Lipid level monitoring for patients receiving Accutane	1
Colonoscopy or barium enema after a first time diagnosis of diverticulitis	3
Yearly visual field tests for glaucoma patients	1
Inappropriate Use of Antibiotics for Viral Upper Respiratory Infection	10
Osteoporosis Screening Following Fractures	8
Treatment of Coronary Artery Disease: Beta Blockers	8
CBC and LFT for Rheumatoid Arthritis Patients on Methotrexate	1
Creatinine Check For New Non-Selective NSAID Use	11
Breast Cancer and Radiation Therapy	2
Annual Endometrial Cancer Monitoring for Patients on Tamoxifen	5
Annual Creatinine, Potassium, Phosphorus, and Complete Blood Counts for ESRD & Pre ESRD Patients	1
TSH after dosage level change for patients on levothyroxine treatment	1
Follow-up examination after diagnosis and treatment of skin cancers	1
Liver Function Tests (LFT) and Blood Counts for patients initiated on Carbamazepine (Tegretol) or Valproic Acid (Depakene, Depakote)	1
Compliance with glaucoma medications	1
The use of semen analysis, prostatic smear, or prostate biopsy in the diagnostic work-up of chronic prostatitis	1

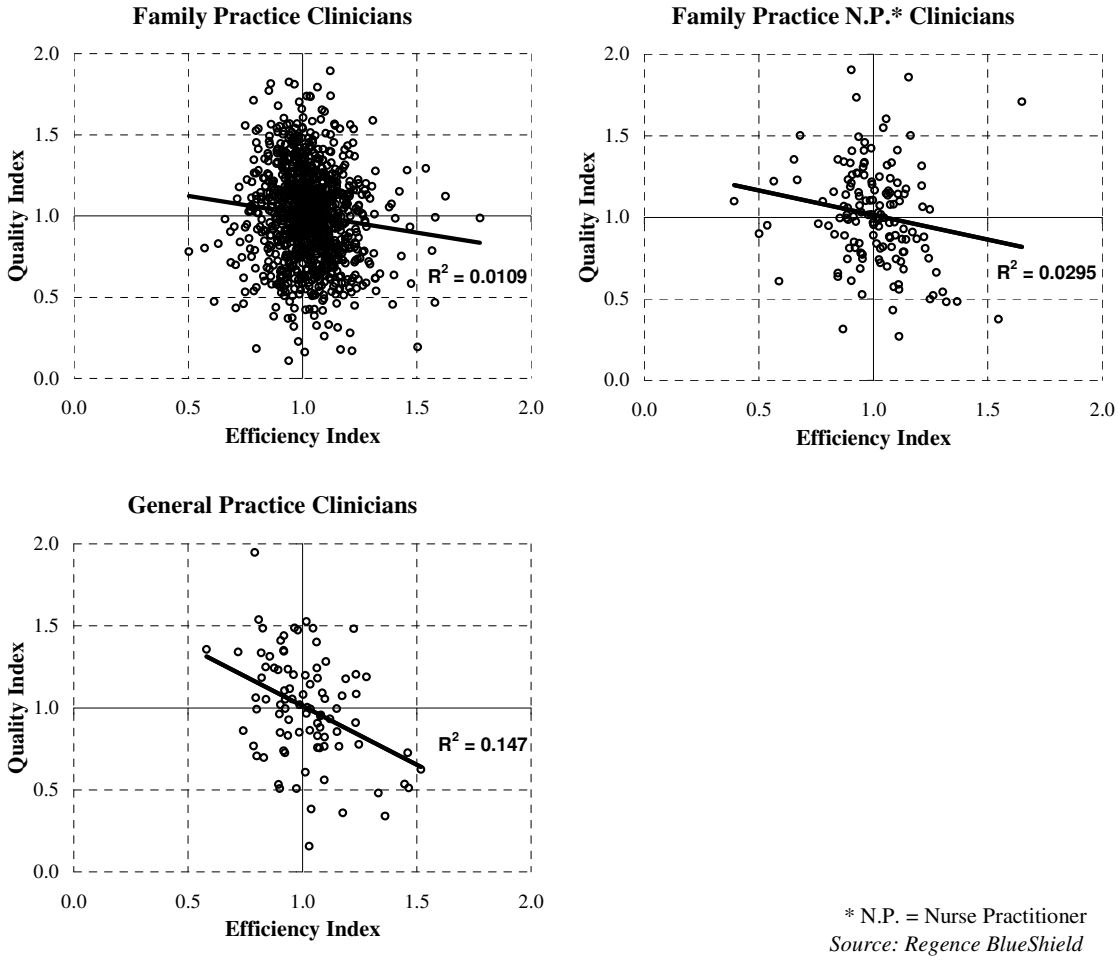
Exhibit 2: Linear Regression Results of Clinician Quality and Efficiency Index Scores

Specialty	Clinician Count	Applicable Quality Indicators	p-value	Coefficient	Lower 95% C.I.	Upper 95% C.I.	R^2
Significant (p<0.05)							
Family Practice	1,070	15	p<.001	(0.226)	(0.356)	(0.097)	0.0109
Family Practice N.P.*	143	13	0.0403	(0.301)	(0.589)	(0.014)	0.0295
General Practice	90	15	p<.001	(0.717)	(1.082)	(0.351)	0.1470
Not significant (p>=0.05)							
Internal Medicine	503	16	0.0702	(0.127)	(0.264)	0.010	0.0065
Obstetrics/Gynecology	306	6	0.9928	(0.001)	(0.243)	0.240	0.0000
Cardiovascular Disease	34	5	0.8130	(0.042)	(0.397)	0.313	0.0018
Pulmonary Diseases	28	5	0.0763	(0.329)	(0.695)	0.037	0.1159
Insufficient data *(less than 25 observations)*							
Adult Medicine N.P.*	24	16	0.4751	(0.425)	(1.637)	0.787	0.0234
Midwife/OB/Gyn N.P.*	23	4	0.1235	1.105	(0.327)	2.538	0.1092
Endocrinology	21	5	0.5827	(0.183)	(0.869)	0.503	0.0162
Osteopathy	17	12	0.7710	(0.063)	(0.517)	0.391	0.0058
Gynecology	12	5	0.7457	(0.243)	(1.865)	1.380	0.0110
Total	**2,271**						

*N.P. = Nurse Practitioner

Source: Regence BlueShield

Exhibit 3: Scatter Plots of Specialties with Significant (p<0.05) Quality-Efficiency Relationships
Note: index scores greater than 1.0 indicate above average quality or efficiency



* N.P. = Nurse Practitioner
Source: Regence BlueShield

TECHNICAL APPENDIX

Figure 1 provides a Data Processing Overview, referenced throughout this section.

Data Collection. We extracted 2 years of claims for the commercially-insured, largely Preferred Provider and Point of Service populations of Regence BlueShield, an independent licensee of the Blue Cross and Blue Shield Association in Washington State. Approximately 1.1 million patients generated nearly 40 million claim lines over a two-year service period through September 2002. Claims spanned both inpatient and outpatient care, whether delivered via professional, facility, or ancillary providers. All pharmacy claims were included. Medicare, Medicaid, and government program claims were excluded, as were capitated programs. Each claim line included data elements found in typical adjudication systems, such as patient and clinician identifiers, service and eligibility dates, and procedure and diagnosis codes.
The same claims were used in evaluating process quality and cost efficiency.

PROCESS QUALITY

Quality Data Preparation. Health Benchmarks measured the frequency of distinct claim values in the procedure code, diagnosis code and National Drug Code data fields. Frequencies were compared with expectations from other commercial plans' claims. Where significant deviations appeared, including errant, unknown, or proprietary codes, an investigation was conducted. Sometimes, investigations resulted in revisions to the selected process indicators.
Eligibility history tables were evaluated to confirm that continuous patient eligibility reflected expected levels (Fig. 1, Cell Q1).

Quality Data Processing. We selected 33 process indicators for evaluation. Selection criteria included overall coverage across the professional specialties, and clinical relevance given the commercial insurance-aged population. Some indicators apply to a single specialty; others apply to several. For instance, X-ray Prior to Knee MRI applies to Occupational Medicine, Physical Medicine, Rheumatology, and Orthopedic Surgery. The number of indicators applicable to a single specialty ranged between 1 and 15.
Where necessary, each indicator's algorithm was adapted to account for idiosyncratic claim aspects, such as proprietary procedure codes. The algorithms for all indicators were then run against the entire claim repository. Output for each indicator was examined for unexpected results, such as unusual implied disease prevalence rates. Any anomalies were investigated, algorithms refined, and rerun (Fig. 1, Cell Q2).
Assignment of cases to clinicians utilized a patient-centered, team-based approach. For each indicator, a case was opened for a clinician when:

1. The indicator in question applied to the clinician's specialty
2. The clinician had an office visit with the patient in question, and
3. The patient was continuously eligible for benefits

As an example, consider the mammography indicator. Suppose a continuously eligible patient has office visits to both a family practitioner and a gynecologist. A case is opened and assigned to both clinicians. If one or both clinicians meet the mammography recommendation, both clinicians receive compliance credit (Fig. 1, Cell Q3).

Quality Data Analysis. Results for each indicator included clinician identifier, specialty, assigned case counts, and number of cases for which the process standard was achieved. A raw compliance rate was

calculated for each clinician-indicator combination, where the clinician had at least 10 assigned cases. At the specialty-indicator level, a minimum of 10 such clinicians was required for subsequent analysis. Only clinicians with at least 3 distinct indicator compliance rates were eligible for the quality index calculation described below[1].

The clinician's raw compliance rate was restated as a percentile score, reflecting the performance of other same-specialty clinicians for the given indicator. A clinician's percentile scores from multiple indicators were weighted into a single composite score. The weighting system captured the practice relevance of a given indicator to a given specialty. For instance, mammography screening is a primary indicator in Obstetrics/Gynecology but a secondary indicator in Pulmonology. Primary indicators were weighted 2:1 relative to secondary indicators.

Using the individual composite scores for clinicians of the same specialty, an equally weighted average composite score was calculated for the specialty. A clinician's quality index was his/her composite score divided by the specialty composite average, as shown in Example 1 below. Thus, the composite score was indexed such that 1.0 approximated the average network compliance within the specialty.

Example 1: Quality Index Calculation

Requirements:
(1) Each clinician must have at least 3 process quality indicators
(2) Each clinician must have a minimum of 10 assigned cases per indicator, and
(3) There must be at least 10 clinicians meeting requirements 1 and 2 above (for each specialty-indicator combination)

Doctor 001, an Obstetrician/Gynecologist	A	B	C	D	E: C x D
Process Indicator	**Assigned Patients**	**Raw Compliance Rate**	**Ob/Gyn Percentile Score**	**Practice Relevance Weight**	**Weighted Percentile**
Colorectal cancer screening	49	45%	33	2	66
Cervical cancer screening	215	96%	95	2	190
Mammography screening	59	92%	69	2	138
Major Depression: Acute Phase Management	15	27%	6	1	6
Major Depression: Optimal Practitioner Contacts	13	8%	31	1	31
Total:	**351**			**8**	**431**
Composite Score (Sum E / Sum D):					**53.9**

All Obstetrician/Gynecologists	A	B	C: A / B
Clinician	**Clinician's Composite Score**	**Average Ob/Gyn Composite Score**	**Quality Index Score**
Doctor 001	53.9	49.4	1.1
Doctor 002	19.0	49.4	0.4
Doctor 003	39.0	49.4	0.8
.	.	.	.
.	.	.	.
Doctor 320	69.6	49.4	1.4
Total:	**15,794.5**		
Average Composite Score (Sum A / Clinician Count):	**49.36**		

In this way, we derived an aggregate, relative performance score for each clinician (Fig. 1, Cell Q4).
Only within-specialty clinician comparisons are relevant. Comparing the quality scores of clinicians with different specialties is problematic because the process indicators vary between specialties.
Comparisons were made to the observed statewide performance. This baseline is not intended to represent a standard of practice.

COST EFFICIENCY

Cost Efficiency Data Preparation. The episode grouping software requires each claim line to be assigned to 1 of 3 Provider Types: Facility, Clinician, or Other. Facility-type providers include hospitals, skilled nursing facilities, etc. The Clinician category encompasses professionals who diagnose, treat, and recommend patient treatment. The Other category includes

remaining providers: laboratories, pharmacies, ambulance, etc.
Each claim line must also be assigned to 1 of 3 Types of Service: Room and Board, Medical-Surgical, or Ancillary. Room and Board generally designates inpatient confinements. Medical-Surgical normally denotes clinician-delivered services. Ancillary contains all remaining services.
The grouping software in part relies on Provider Type and Type of Service assignments to identify encounters justifiably beginning or extending a clinical episode. We prepared the claim extract with Provider Type and Type of Service assignments. The distribution of claim records across the assigned values was examined for conformance with observed ranges from national benchmarks representing the collective experience of other episode grouper licensees (Fig. 1, Cell E1).

Cost Efficiency Data Processing. Once the claim extract was prepared, episode grouping occurred. The grouper examined individual claim lines for related clinical encounters. Within collections of clinically related encounters, the software examined the claim sequence for periods without service activity. These "clean periods" vary by ETG and determine whether subsequent clinical activity should be considered part of an existing episode or the start of a new episode.
Only complete episodes were recognized during the grouping process. Episodes already underway before the earliest service date in the claims extract were not recognized; episodes still open by the latest service date were rejected. An exception was made to capture chronic conditions. Chronic episodes exceeding 1 year were split into sequential episodes lasting up to 1 year.
For each ETG, the grouping software contains outlier thresholds called "trim points". Trim points flag individual episodes as potentially aberrant if total charged dollars are too large or small. Trim point values are derived from a nationwide repository of episodic charge data, as supplied by individual grouper licensees.
Thus, the ETG software grouped clinically related encounters within the claim extract. Grouping results were checked against benchmarks and found to be within expected ranges. Benchmarks included frequency of ungrouped claim lines; distribution of complete and incomplete episodes; frequency of errant diagnosis, procedure, and drug codes; etc. (Fig. 1, Cell E2).
An important step in measuring efficiency is defining an episode attribution rule: the method for assigning an entire episode to the clinician most responsible for its constituent services, whether rendered, ordered, or referred to others. We used a modified version of a widely used attribution rule, which assigns the episode to the clinician with the greatest share of charges. We made 1 modification and 2 qualifications:

1. Use of paid dollars, rather than charges, thereby capturing total expected clinician reimbursement, including patient cost sharing
2. Restriction of the examined services to professional records only (office visits and medical/surgical procedures), and
3. Requirement that an attributed clinician account for at least 30% of the total paid dollars for these examined services[2]

For example, consider a complete episode comprising $1,000. Suppose professional office visits and surgical procedures amount to $450. If no clinician accounted for at least $135 of this service subset, or 30%, no episode attribution occurs. The clinician with the highest percentage greater than or equal to 30% is assigned the episode. When 2 or more clinicians are tied for the greatest percentage, no attribution occurs (Fig. 1, Cell E3).
We used version 4.4 of the Symmetry episode grouper. Data preparation and analysis activities were conducted using SAS for Windows version 8.01, with SAS Enterprise Guide version 2.0 as a graphical user interface. The computing platform was a dedicated Hewlett Packard Proliant DL

380 G3 with Intel Xeon 3.06 GHz processing and 4 gigabytes of RAM.

Cost Efficiency Data Analysis. From 39,380,577 claim lines, 2,912,006 distinct, complete episodes were identified. Of these episodes, 595,957 were eliminated because they qualified as outliers, or because the attribution rule resulted in no clinician assignment. An additional 131,701 episodes were eliminated because of a requirement that each specialty-ETG combination contain at least 100 episodes for subsequent processing. The remaining 2,184,348 episodes constituted the basis for deriving clinician efficiency index scores. These episodes accounted for $724,180,158, or 21.9% of the all paid dollars in the original claim extract, a retention rate within expected ranges for episode grouping runs.

For these 2,184,348 episodes, the network-wide average paid dollars per episode was calculated for each combination of specialty and ETG, representing the expected cost.

Only individual clinicians with at least 30 attributed episodes across ETGs featuring at least 10 episodes were evaluated for efficiency. These clinicians' collective episode volume was 1,511,170, accounting for $352,865,766.

To measure a specific clinician's efficiency, that clinician's observed cost for each attributed episode was compared to the corresponding network expected cost. The ratio of total expected to total observed dollars became the clinician's relative efficiency index as demonstrated in Example 2 below. Note that an index score of 1.0 approximated the network-wide expectation for the studied mix of ETGs.

Example 2: Efficiency Index Calculation

Requirements:
(1) Each Specialty-ETG combination must contain at least 100 comparative episodes network-wide
(2) Only clinician-ETG combinations with at least 10 attributed episodes are retained
(3) Of retained episodes, a clinician must have at least 30 attributed episodes to be evaluated

Doctor ABC, a Family Practice Physician	A	B	C	D	E: D / C	F: E x A
Episode Treatment Group®	**Clinician Episodes**	**Clinician Cost**	**Network Episodes**	**Network Cost**	**Average Network Cost**	**Expected Cost: Average Network Cost * Clinician Episodes**
Otitis media, w/o surgery	14	$967	25,679	$2,429,351	$95	$1,324
Tonsillitis, adenoiditis or pharyngitis, w/o surgery	19	$1,707	43,417	$3,957,743	$91	$1,732
Acute bronchitis, w/o comorbidity, age 5+	17	$2,250	29,506	$3,708,763	$126	$2,137
Total:	**50**	**$4,924**				**$5,193**
Efficiency Index Score (Sum F / Sum B):						**1.05**

In this way, we derived an aggregate, relative performance score for each clinician (Fig. 1, Cell E4).

We opted not to compare the treatment patterns of physicians with differing specialties, even when treating the same condition. For example, we recognized an episode of sinusitis attributed to an otolaryngologist separately from one attributed to a family medicine physician. This acknowledgement enhances within-specialty severity adjustment.

As with quality comparisons, efficiency comparisons were made to statewide performance. This baseline is not intended to represent a standard of performance.

There were 2,739 clinicians (12 specialties) with quality index scores and 5,654 clinicians (69 specialties) with efficiency index scores; 2,271 clinicians (12 specialties) had both quality and efficiency index scores. A total of 26,426 distinct clinicians (150 specialties) comprised the professional network during the study period. Thus, we evaluated 8.6% of the total clinician network for both process quality and cost efficiency.

NOTES

1. Separate analysis had determined that clinicians without at least 3 distinct indicator compliance rates had more volatile index scores relative to their peers. Excluding these clinicians produced normal or normal-like distributions of index scores within each specialty.
2. Analysis showed a low frequency of episodes with multiple managing clinicians, regardless of the attribution threshold under consideration. At attribution thresholds beyond the 35%-40% range, we begin to see an increasing trend where the more costly episodes are not attributed to a responsible clinician.
3. The quality analysis, because of its focus on evidence-based medicine, only featured traditional medical-surgical professional specialties. The efficiency analysis considered all specialties found in health insurance claim data, both traditional and non-traditional (such as acupuncturist, naturopath, chiropractor, etc.).

Figure 1: Data Processing Overview for Deriving Professional Quality and Cost Efficiency Index Scores (see accompanying file "Quality Implications of Efficiency-Based Clinician Profiling - Technical Appendix Figure 1.pdf")

Figure 1 (Technical Appendix): Data Processing Overview for Deriving Professional Quality and Cost Efficiency Index Scores