Physician Tiering by Health Plans in Massachusetts

By Ajay D. Wadgaonkar, BS, Eric C. Schneider, MD, MSc, and Timothy Bhattacharyya, MD

Investigation performed at the National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland

Background: Physician tiering is an emerging health-care strategy that purports to grade physicians on the basis of cost-efficiency and quality-performance measures. We investigated the consistency of tiering of orthopaedic surgeons by examining tier agreement between health plans and physician factors associated with top-tier ranking.

Methods: Health plan tier, demographic, and training data were collected on 615 licensed orthopaedic surgeons who accepted one or more of three health plans and practiced in Massachusetts. We then computed the concordance of physician tier rankings between the health plans. We further examined the factors associated with top-tier ranking, such as malpractice claims and socioeconomic conditions of the practice area.

Results: The concordance of physician tiering between health plans was poor to fair (range, 8% to 28%, $\kappa = 0.06$ to 0.25). The percentage of physicians ranked as top-tier varied widely among the health plans, from 21% to 62%. Thirty-eight percent of physicians were not rated top-tier by any of the health plans, whereas only 5.2% of physicians were rated top-tier by all three health plans. Multivariate analysis showed that board certification, accepting Medicaid, and practicing in a suburban location were the independent factors associated with being ranked in the top tier. More years in practice or fewer malpractice claims were not related to tier.

Conclusions: Current methods of physician tiering have low consistency and manifest evidence of geographic and demographic biases.

In 2005, the Group Insurance Commission of Massachusetts required participating health plans to control costs by stratifying physicians into tiers¹. Physician tiering, or physician clinical performance assessment, is a new concept in health care that ranks physicians into two or more categories on the basis of cost efficiency and adherence to accepted clinical practice standards². “Top-tier” physicians, with better cost-efficiency or quality ratings, are assigned a lower copayment than physicians in the second tier. Ultimately, patients are encouraged to visit top-tier physicians because of lower out-of-pocket fees, while health insurance companies reduce spending by endorsing more cost-efficient practices. Tiering is primarily used in specialties with heavy spending, wide variation in resource use, and established practice guidelines².

At the present time, there is little information about programs of physician stratification through clinical performance assessment. We therefore performed a cross-sectional study, investigating the tiering of orthopaedic surgeons in Massachusetts. We hypothesized that if the tiering methodologies were robust, health plans would have a high degree of

concordance and would tend to rate the same physicians as top-tier. We also studied the physician demographic and practice characteristics associated with top-tier status.

Materials and Methods

Data Acquisition

We collected demographic, insurance, training, and licensing data on 615 orthopaedic surgeons practicing in the state of Massachusetts. We used the Commonwealth of Massachusetts Board of Registration in Medicine (www.massmedboard.org, updated July 2, 2007) to identify surgeons for the present study. These surgeons were then cross-referenced with physicians listed by the three Massachusetts health plans that publish physician tiering results. We included only surgeons who were licensed and practicing in Massachusetts and who accepted one or more of these three health plans. We excluded surgeons who were practicing outside of Massachusetts. Two physicians who were listed as orthopaedic surgeons in the insurance company registries were found to be physical medicine and rehabilitation specialists, and they were removed from the analysis.

Disclosure: In support of their research for or preparation of this work, one or more of the authors received, in any one year, outside funding or grants of less than \$10,000 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), National Institutes of Health. In addition, one or more of the authors or a member of his or her immediate family received, in any one year, payments or other benefits of less than \$10,000 or a commitment or agreement to provide such benefits from commercial entities (Stryker, AO North America, and Best Doctors, Inc.).

From the Commonwealth of Massachusetts Board of Registration in Medicine, we recorded demographic data on the surgeons, including board certification, Medicaid participation, and the number of malpractice claims filed against them in the last ten years. We calculated the number of years in practice by subtracting the year in which training was completed from 2007. We used the physician's practice address listed with the Board of Registration to determine practice zip code. From the 2000 United States Census (www.census.gov), we then gathered data on the practice zip code's population density, the percentage of individuals below the poverty level, and median household income. This approach to demographic data has been used in a number of other studies³⁻⁸. Population density data were further used to classify a surgeon's practice setting as urban (>3000 persons per square mile), suburban (1000 to 3000 persons per square mile), or rural (<1000 persons per square mile). Individual surgeon volume for the physicians performing more than ten joint replacements in 2005 was obtained from the Commonwealth of Massachusetts.

Health Plan Orthopaedic Surgeon Tiering Methods

Table I briefly reviews insurance company methods for stratifying orthopaedic surgeons into tiers for the 2007 study year. We investigated each health plan web site and queried the three insurance companies for details regarding their physician-stratification methodology. However, the availability of in-depth information describing tiering methods was very limited.

All three health plans stratified surgeons on the basis of specific cost-efficiency and performance-quality benchmarks. Plan C analyzed surgeons at the group level only. Each health plan used Symmetry's Episode Treatment Group (ETG) methodology to assess cost efficiency⁹. This model analyzes inpatient, outpatient, and pharmaceutical

and auxiliary claims data from a comprehensive database originally compiled by Mercer human resource consulting. Claims data were combined and classified into distinct, complete episode treatment groups, with each episode treatment group linked to the treating physician. For example, an episode of back pain would be identified on the basis of billing data and the rate of utilization of imaging studies measured. Next, the method computed a physician's average resource use for each episode treatment group and compared these values with average resource usage by physicians in the same specialty treating a clinically similar mixture of episode treatment groups.

The three health plans primarily based quality ratings on physician adherence to selected medical practice guidelines and other evidence-based standards. For example, Plan A used an osteoporosis-management measure identifying "the percentage of women sixty-seven years and older who suffered a first bone fracture and who had either a bone mineral density test or a prescription for a drug to treat or prevent osteoporosis during the six months after the date of fracture."¹⁰

Two health plans (Plans B and C) also relied on quantitative patient satisfaction ratings, use of e-prescribing, use of electronic medical records, and physician use of self-made patient registries to gauge quality. The quality measures that were used were diverse and differed by health plan. Quality measures were taken from such sources as the Healthcare Effectiveness Data and Information Set (HEDIS), Resolution Health, the Agency for Healthcare Research and Quality (AHRQ), and Massachusetts Health Quality Partners, among others. Quality scores were calculated and, in conjunction with cost-efficiency data, were used to determine final tier status. Plans A and B classified individual surgeons. Plan C placed all surgeons within a group in the same tier. Plan A required surgeons to meet both cost and quality metrics to reach the top tier.

TABLE I Summary of Health-Plan Tiers and Classification Methodology

Health Plan/Tier	Copay	Classification Method
Plan A		Specialists are individually classified into tiers on the basis of claims data from 2003 to 2005.
Tier 1*	\$10	Specialists are initially assigned to the "Tier 3" category. Specialists who meet both cost-efficiency and quality benchmarks are designated Tier 1. Specialists who meet either the cost-efficiency or quality benchmark, with insufficient data collected for the other variable, are designated Tier 2.
Tier 2	\$10	
Tier 3	\$20	Specialists who have practiced for two years or less in Massachusetts, with insufficient data collected for both quality and cost-efficiency, are also designated Tier 1.
Plan B		Specialists are individually classified into two tiers on the basis of claims data from 2003 to 2005.
Tier 1	\$15	Tiering is accomplished with use of a combined quality-cost score. Specialists who meet a benchmark score are designated Tier 1. Specialists who do not meet the benchmark or who have insufficient claims data are designated Tier 2.
Tier 2	\$25	
Plan C		Specialists are tiered at the group level on the basis of claims data from 2003 to 2005. Tiering relies on a gated method: only specialist groups that meet the quality threshold, or for whom there is insufficient data for quality, are analyzed for cost-efficiency. Specialist groups that then meet the cost-efficiency threshold are designated Tier 1, while all other groups are Tier 2.
Tier 1	\$15	
Tier 2	\$25	

*No physician was placed in Tier 1 for this health plan.

Statistical Analysis

We computed descriptive univariate data on the surgeon cohort. For ease of interpretation, surgeons were classified by whether they practiced in high-income (top quartile of median income, >\$56,000) or low-poverty (bottom quartile, <6% poverty) zip codes. We also analyzed these variables as continuous variables and found no change in results (data not shown).

We then compared the concordance of top-tier classification of surgeons between health plans. We limited the analysis to surgeons who participated in multiple health plans and calculated the kappa statistic. Agreement was defined as poor ($\kappa < 0.20$), fair ($\kappa = 0.21$ to 0.40), moderate ($\kappa = 0.41$ to 0.60), substantial ($\kappa = 0.61$ to 0.80), or very good ($\kappa > 0.80$)¹¹. For Plan A, zero physicians were in Tier 1, so we analyzed Tier 2 as the effective top tier.

For each health plan, we screened for associations between surgeon demographic factors and top-tier status. Factors that were found to be significantly associated ($p < 0.05$) were included in multivariate models. We used binary logistic regression, with tier as the dependent variable, and calculated odds ratios and 95% confidence intervals for each factor. We repeated the process for each of the three health plans.

Geographic Analysis

In order to understand the geographic distribution of top-tier orthopaedic surgeons, we used practice zip codes to calculate the number of surgeons practicing in each city or town in Massachusetts. We then stratified the number of surgeons per city/town into quartiles and created a surgeon-distribution map. We proceeded to calculate the percentage of orthopaedic surgeons in a city/town who were rated as top-tier by at least one health plan and then created a map of the results.

Source of Funding

This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. The funding was used to pay salaries.

Results

Surgeon Cohort

The majority (84%) of orthopaedic surgeons were board certified, and 96% reported one or fewer malpractice claims in the last ten years (Table II). Four hundred and eight surgeons (66%) participated in all three health plans.

Health Plan Tier Agreement

The percentage of surgeons who were rated as top-tier varied widely among the health plans, from 21% to 62% (Table III). None of the surgeons participating in Plan A met both of the health plan's cost-efficiency and quality benchmarks, and therefore, none was rated in the top tier.

TABLE II Cohort of Surgeons (N = 615)

	Number of Surgeons or Mean Value*
Sex	
Male	585 (95%)
Female	30 (5%)
No. of years in practice	18 ± 12
Board certification	519 (84%)
Orthopaedic subspecialty†	
None reported	380 (62%)
Sports medicine	90 (15%)
Hand surgery	78 (13%)
Arthroplasty	20 (3.3%)
Pediatric orthopaedics	20 (3.3%)
Spine	14 (2.3%)
Trauma	9 (1.5%)
Foot and ankle	5 (0.81%)
Other	8 (1.3%)
Reported practice type‡	
Private office	220 (36%)
Hospital	178 (29%)
Partnership or group practice	176 (29%)
Educational institution	16 (2.6%)
Clinic	13 (2.1%)
HMO§	2 (0.33%)
Accepts Medicaid	517 (84%)
No. of malpractice claims since 1997	0.20 ± 0.54
Malpractice claim(s) since 1997	
0 or 1	591 (96%)
>1	24 (3.9%)
Practice setting#	
Urban	344 (56%)
Suburban	201 (33%)
Rural	67 (11%)
Annual joint replacement volume**	73 ± 75
Health plan tier status	
Rated top-tier by all 3 health plans	32 (5.2%)
Rated top-tier by 2 health plans	130 (21%)
Rated top-tier by 1 health plan	219 (36%)
Rated top-tier by no health plan	234 (38%)

*The values are given as the number of surgeons, with the percentage in parentheses, or as the mean and the standard deviation. †Some surgeons reported two specialties. ‡Ten surgeons did not report a practice type. §Health-maintenance organization. #Data were missing for three surgeons. **For surgeons performing ten or more hip or knee replacements per year.

Agreement between the health plans as to which surgeons were top-tier was low, ranging from 8% (between Plans B and C) to 28% (between Plans A and B). Only 5.2% of

TABLE III Concordance of Orthopaedic Surgeon Top-Tier Status Among Health Plans

	Plan A, Tier 2 (455 participating surgeons; 0 designated Tier 1, 281 [62%] designated Tier 2)	Plan B, Tier 1 (571 participating surgeons; 176 [31%] designated Tier 1)	Plan C, Tier 1 (569 participating surgeons; 118 [21%] designated Tier 1)
Plan A, Tier 2			
No. of surgeons who accepted both health plans	—	421	413
No. of surgeons rated as top-tier in both health plans	—	118 (28%)	64 (15%)
Kappa value	—	0.25	0.10
Plan B, Tier 1			
No. of surgeons who accepted both health plans	421	—	554
No. of surgeons rated as top-tier in both health plans	118 (28%)	—	44 (8%)
Kappa value	0.25	—	0.04
Plan C, Tier 1			
No. of surgeons who accepted both health plans	413	554	—
No. of surgeons rated as top-tier in both health plans	64 (15%)	44 (8%)	—
Kappa value	0.10	0.04	—

surgeons were rated top-tier by all three health plans. The kappa statistic ranged from 0.04 (Plan B to Plan C) to 0.25 (Plan A to Plan B).

Surgeon Demographics and Top-Tier Status

On univariate analysis, board certification, longer time in practice, more malpractice claims, practicing in a low-poverty area, and Medicaid acceptance were all associated with top-tier status for Plans A and B ($p < 0.02$ for all). For Plan C, only Medicaid acceptance and suburban practice location were associated with top-tier status ($p < 0.029$ and $p < 0.0001$, respectively).

Table IV shows physician factors that were independently associated with top-tier status for each health plan.

Practicing in a suburban location was significantly associated with being ranked as top-tier for all three health plans. For Plans A and B, board certification and Medicaid acceptance were also significantly associated with top-tier status. On multivariate analysis, the number of malpractice claims filed against a physician in the last ten years and the number of years in practice were not significantly associated with top-tier ratings for any health plan.

Geographic Distribution of Top-Tier Surgeons

The geographic distribution of orthopaedic surgeons generally reflects the population distribution of the state (see Appendix). However, health plans tended to assign top-tier rankings to physicians who practiced in suburban locations (see Appendix).

TABLE IV Physician Factors Associated with Top-Tier Status

Factor	Health Plan		
	Plan A, Highest Tier*	Plan B, Highest Tier*	Plan C, Highest Tier*
Board certified	8.29 (3.4 to 20.1)	3.61 (1.46 to 8.94)	
Accepts Medicaid	3.67 (1.74 to 7.7)	2.53 (1.19 to 5.34)	1.83 (0.91 to 3.71)
Number of years in practice	1.02 (1.00 to 1.04)	1.01 (0.99 to 1.03)	
Number of malpractice claims since 1997	1.20 (0.75 to 1.91)	1.31 (0.93 to 1.85)	
Suburban practice location	1.94 (1.19 to 3.17)	1.69 (1.14 to 2.52)	2.18 (1.44 to 3.32)

*The data are given as the odds ratio, with the 95% confidence interval in parentheses.

The Boston urban area had a lower percentage of top-tier surgeons compared with the surrounding suburban areas.

Surgeon Volume and Top-Tier Status

Surgeon volume was not correlated with tiering status for Plans A and B ($p = 0.75$ and 0.97 , respectively). For Plan C, a top-tier status was associated with a lower annual volume of total joint replacements (average, eighty-one compared with fifty-five cases per year; $p = 0.01$).

Discussion

Our review of physician clinical performance assessment by three health plans in Massachusetts shows only modest physician tiering agreement between health plans. Top-tier physicians tended to be board certified, to practice in suburban locations, and to accept Medicaid. More years in practice or fewer malpractice claims were not associated with top-tier status.

Physician clinical performance assessment, or physician tiering, posits that physicians may be graded on their quality of care on the basis of administrative and claims data. Physicians who are higher-quality and are highly cost-efficient are assigned to a lower copayment. The lower copayment fee encourages patients to visit more cost-efficient “top-tier” practitioners, thus lowering the cost of health care for the community. The field of physician stratification, however, is in its infancy¹². Given that the assessment of physician clinical performance and rewarding for better outcomes is part of many health-care reform strategies, rigorous analysis of current programs is necessary.

Because “high performance” is a quality that is specific to individual surgeons, we hypothesized that, while there would be some variability, the same surgeons would be rated top-tier by multiple health plans. However, agreement between health plans on top-tier status was modest at best. There are three reasonable inferences: (1) the plans are not measuring physician quality, (2) they are measuring very different aspects of quality, or (3) they are giving markedly different weights to the different aspects of quality. For example, two plans may in fact be measuring accurately but placing much more weight on cost effectiveness versus adherence to practice guidelines. The measurement may be further distorted by the limitations of using billing data to make clinical judgments.

Why would practicing in a suburban location, or accepting Medicaid, be associated with top-tier status? There are no data to indicate that higher-quality care is delivered in suburban areas or that lower-quality care is delivered in urban areas. While it is possible that high-quality surgeons may tend to self-cluster into suburban locations, physician tiering methods are represented to be based on adherence to practice guidelines and cost efficiency, which should have little relation to geography. We hypothesize that the clustering of top-tier physicians into suburban areas reflects the failure of tiering methodologies to adequately adjust for the increased severity and complexity of patients who seek care at urban academic medical centers, and we suggest that practice location be used as a controlling factor for risk adjustment^{13,14}. Furthermore, we

noted that Plan C would tend to steer patients to lower-volume surgeons, although higher surgeon volume has been associated with better outcomes for patients^{15,16}. However, our data can only document associations and cannot definitively exclude the possibility that high-quality surgeons tend to choose to practice in suburban locations.

Implications for Physicians

Individual physicians receive their tiering report cards from insurance companies in isolation, without access to relevant background data. Most surgeons are discouraged to find that they have not been classified into the top tier. However, in Plan A, for example, no orthopaedic surgeons were classified in the top tier. We strongly encourage health plans to supply data concerning how many physician peers and physicians in their practice area were ranked as top-tier and to clarify the measures used to stratify physicians.

Changing practice patterns seems to be unwarranted because the system is not sufficiently developed. The lack of agreement between health plans demonstrates that many physicians will be top-tier for one plan and lower-tier for another. For now, the difference in copayment level is a nominal \$10.00 and is probably unlikely to impact patient choices.

Implications for Patients

There is a true need for greater patient education about the current status of physician tiering. While health plans prominently display a physician's tier in online and printed directories, information about stratification methodology and its relation to quality of care is limited. For example, a patient may be frustrated to find that his or her surgeon is in the second tier. However, there is no evidence to suggest that second-tier physicians provide lower-quality care, and there is a good possibility that the surgeon may be ranked as top-tier by another health plan. Tiering programs may also disrupt existing referring patterns between physicians that have been developed on the basis of years of trust and good results.

Implications for the Health-Care System

With physician stratification increasing in popularity, payors must be attuned to its limitations. The systems for identifying high and low-quality physicians are nascent and have yet to be validated. If the methodologies for stratifying physicians are unreliable, health plans may not realize substantial cost savings with these programs because patients are not truly steered to high-quality physicians. Finally, greater transparency of tiering methodology will benefit both physicians and health plans.

Limitations

The present study had several limitations. First, we explored tiering data for just one specialty, and thus our conclusions may not be applicable to other medical specialties. Second, it may be that the lack of agreement between health plans may be a result of measuring different aspects of health-care quality. For example, Plan A may place much greater weight on cost-efficient care, whereas Plan B may emphasize completion of all

recommended tests. In such a scenario, the payors are unlikely to agree on physician-tiering status. However, from the patient's perspective, "high quality" is a global characteristic, and the goal should be to measure and classify physicians on all aspects of quality. We analyzed the first year of this new system, and the methodologies have changed over time. Finally, there is no consensus gold-standard for high physician clinical performance that can be used as a benchmark to validate the physician-tiering systems.

Further research is needed by both physicians and health plans in the area of physician tiering. Ongoing efforts to validate the tiering methods against clinical data are needed. The development of a robust technique for identifying high-quality physicians would certainly prove beneficial for all parties.

Appendix

Maps depicting orthopaedic surgeon density and distribution by tier are available with the electronic version of

this article on our web site at jbjs.org (go to the article citation and click on "Supporting Data"). ■

Ajay D. Wadgaonkar, BS
950 25th Street N.W., Apartment 707 South,
Washington, DC 20037

Eric C. Schneider, MD, MSc
RAND Boston, 20 Park Plaza 7th Floor,
Suite 720, Boston, MA 02116

Timothy Bhattacharyya, MD
Clinical and Investigative Orthopaedic Surgery Section,
Office of Clinical Director, NIAMS,
10 Center Drive, Bldg 10-CRC Room 4-2339,
Bethesda, MD 20892.
E-mail address: Timothy.bhattacharyya@nih.gov

References

- Powell JH. Docs wary of insurance company plan to grade them. *Boston Herald*. 2005 Dec 1;59.
- Draper DA, Liebhaber A, Ginsburg PB. High-performance health plan networks: early experiences. *Issue Brief Cent Stud Health Syst Change*. 2007;111:1-6.
- Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA*. 2006;296:2694-702.
- Gornick ME, Eggers PW, Reilly TW, Mentnech RM, Fitterman LK, Kucken LE, Vladeck BC. Effects of race and income on mortality and use of services among Medicare beneficiaries. *N Engl J Med*. 1996;335:791-9.
- Santry HP, Gillen DL, Lauderdale DS. Trends in bariatric surgical procedures. *JAMA*. 2005;294:1909-17.
- Zingmond DS, Soohoo NF, Silverman SL. The role of socioeconomic status on hip fracture. *Osteoporos Int*. 2006;17:1562-8.
- Neuner JM, Zhang X, Sparapani R, Laud PW, Nattinger AB. Racial and socioeconomic disparities in bone density testing before and after hip fracture. *J Gen Intern Med*. 2007;22:1239-45.
- DeLia D. Distributional issues in the analysis of preventable hospitalizations. *Health Serv Res*. 2003;38(6 Pt 2):1761-79.
- Ingenix. Symmetry Episode Treatment Groups. <http://www.ingenix.com/Products/Payers/CareHealthManagementPAY/EnterpriseWideDecisionSupport/EpisodeTreatmentGroups/>. Accessed 2009 July 2.
- Resolution Health. Technical specifications posting booklet: Physician Quality Measures Group Insurance Commission (GIC). Version 1.0. 2008 Aug 18. <http://unicare-cip.com/pdf/RHI%20Quality%20Measures%20ITEM%208.pdf>. Accessed 2009 July 1.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-74.
- Landon BE, Normand SL, Blumenthal D, Daley J. Physician clinical performance assessment: prospects and barriers. *JAMA*. 2003;290:1183-9.
- Hill LD, Madara JL. Role of the urban academic medical center in US health care. *JAMA*. 2005;294:2219-20.
- The Commonwealth Fund. Envisioning the future of academic health centers. Final report of The Commonwealth Fund Task Force on Academic Health Centers. 2003 Feb. http://www.commonwealthfund.org/~media/Files/Publications/Fund%20Report/2003/Feb/Envisioning%20the%20Future%20of%20Academic%20Health%20Centers/ahc_envisioningfuture_600%20pdf.pdf. Accessed 2010 July 9.
- Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E. Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. *J Bone Joint Surg Am*. 2004;86:1909-16.
- Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States Medicare population. *J Bone Joint Surg Am*. 2001;83:1622-9.