Relationship Between Patient Panel Characteristics and Primary Care Physician Clinical Performance Rankings

Clemens S. Hong, MD, MPH
Steven J. Atlas, MD, MPH
Yuchiao Chang, PhD
S. V. Subramanian, PhD
Jeffrey M. Ashburner, MPH
Michael J. Barry, MD
Richard W. Grant, MD, MPH

Physicians have increasingly become the focus of quality performance measurement. Many health care systems now use physician clinical performance assessment as part of their re-credentialing process—to support health care choices for consumers and to guide quality improvement and cost-containment efforts. Partly in response to significant variation in the quality of care,1-3 pay-for-performance and public reporting programs have become widely adopted approaches to influence clinician performance.4 These programs use performance incentives including cash payments and public reports to motivate clinicians, practice groups, and health care systems to achieve specific health care quality goals.5-8

An intrinsic assumption underlying physician clinical performance assessment is that the measures represent physician performance. However, the same physician may have higher or lower measured quality scores depending on the panel of patients he or she manages. The association of patient panel characteristics with physician quality scores could lead to inaccurate physician clinical performance rankings that could have implications on how physicians are rewarded and on how resources are allocated within health care systems.

Context  Physicians have increasingly become the focus of clinical performance measurement.

Objective  To investigate the relationship between patient panel characteristics and relative physician clinical performance rankings within a large academic primary care network.

Design, Setting, and Participants  Cohort study using data from 125,303 adult patients who had visited any of the 9 hospital-affiliated practices or 4 community health centers between January 1, 2003, and December 31, 2005, (162 primary care physicians in 1 physician organization linked by a common electronic medical record system in Eastern Massachusetts) to determine changes in physician quality ranking based on an aggregate of Health Plan Employer and Data Information Set (HEDIS) measures after adjusting for practice site, visit frequency, and patient panel characteristics.

Main Outcome Measures  Composite physician clinical performance score based on 9 HEDIS quality measures (reported by percentile, with lower scores indicating higher quality).

Results  Patients of primary care physicians in the top quality performance tertile compared with patients of primary care physicians in the bottom quality tertile were older (51.1 years [95% confidence interval {CI}, 49.6-52.6 years] vs 46.6 years [95% CI, 43.8-49.5 years], respectively; P < .001), had a higher number of comorbidities (0.91 [95% CI, 0.83-0.98] vs 0.80 [95% CI, 0.66-0.95]; P = .008), and made more frequent primary care practice visits (71.0% [95% CI, 68.5%-73.5%] vs 61.8% [95% CI, 57.3%-66.3%] with >3 visits/year; P = .003). Top tertile primary care physicians compared with the bottom tertile physicians had fewer minority patients (13.7% [95% CI, 10.6%-16.7%] vs 25.6% [95% CI, 20.2%-31.1%], respectively; P < .001), non–English-speaking patients (3.2% [95% CI, 0.7%-5.6%] vs 10.2% [95% CI, 5.5%-14.9%]; P < .001), and patients with Medicaid coverage or without insurance (9.6% [95% CI, 7.5%-11.7%] vs 17.2% [95% CI, 13.5%-21.0%]; P < .001). After accounting for practice site and visit frequency differences, adjusting for patient panel factors resulted in a relative mean change in physician rankings of 7.6 percentiles (95% CI, 6.6-8.7 percentiles) per primary care physician, with more than one-third (36%) of primary care physicians (59/162) reclassified into different quality tertiles.

Conclusion  Among primary care physicians practicing within the same large academic primary care system, patient panels with greater proportions of underinsured, minority, and non–English-speaking patients were associated with lower quality rankings for primary care physicians.

JAMA. 2010;304(10):1107-1113

www.jama.com

©2010 American Medical Association. All rights reserved.
We tested the hypothesis that a physician’s patient panel characteristics are independently associated with changes in his or her relative quality ranking. We ranked all primary care physicians within a single, academic primary care network according to a composite of commonly used Health Plan Employer and Data Information Set (HEDIS) measures and (1) compared patient panel characteristics of highest vs lowest ranked physicians, (2) examined changes in primary care physician rankings after adjusting for differences in patient panel characteristics, and (3) compared the patient panel characteristics of primary care physicians who moved up or down in their quality rankings.

METHODS
We conducted a cohort analysis to study the relationship between patient panel characteristics and physician clinical performance. The Massachusetts General Hospital Practice–based research network includes 181 primary care physicians working in 9 hospital-affiliated practices and 4 community health centers. All practices use the same electronic billing and scheduling systems and share an advanced electronic medical record system. Physicians are hired and credentialed using similar criteria, share the same compensation plan, and have similar staffing resources.

We identified 164,283 adult patients who visited any Massachusetts General Hospital Practice–based research network primary care practice from January 1, 2003, to December 31, 2005, using electronic billing records and excluding those who died (n = 2,817). We calculated physician composite rankings at each stage of adjustment as 162 primary care physicians caring for 125,303 adult primary care patients.

We obtained data from an electronic record repository for Massachusetts General Hospital and affiliated institutions. Patient characteristics included date of birth, sex, self-identified race/ethnicity, primary language, and insurance status. We assessed race/ethnicity data because prior studies have demonstrated associations with patient quality outcomes. Based on a previously published method, we estimated patient medical complexity using the number of chronic medical conditions for each patient taken from International Classification of Diseases, Ninth Revision, billing codes for 9 common chronic conditions: atrial fibrillation, chronic obstructive pulmonary disease, coronary artery disease, depression, diabetes, heart failure, hypertension, osteoarthritis, and stroke.

We also approximated socioeconomic status in the absence of patient-level administrative data by matching patient addresses to the geographic information system coordinates to obtain median household income and high school graduation rates for each patient’s US Census block group. We matched addresses using StreetMap Premium software (Esri, Redlands, California), obtained geographic information system coordinates from successfully matched addresses using ArcGIS software (Esri), and input these geographic information system coordinates into the Demographic Update (Esri) to obtain 2007 US Census block group data. We matched 124,419 of the addresses (99.3%) with 80% sensitivity (requiring 80% of the characters to match identically) and an additional 838 using lower sensitivity limits (0.6%), leaving 46 unmatched (0.4%).

We created a composite quality measure based on 9 HEDIS measures including (1) mammography in the previous 2 years for eligible women aged 42 to 69 years; (2) Papanicolaou cervical screening in the previous 3 years for eligible women aged 21 to 64 years; (3) colonoscopy within 10 years, sigmoidoscopy or double-contrast barium enema within 5 years, or home fecal occult blood testing within 1 year for eligible patients aged 52 to 69 years; (4) hemoglobin A1c testing in the prior year and proportion with levels of 7.0% or less in patients with diabetes; (5) low-density lipoprotein cholesterol testing in the previous year and proportion with levels of 100 mg/dL or less (to convert to mmol/L, multiply by 0.0259) for patients with diabetes and coronary artery disease.

To create the composite physician quality score, we first ranked primary care physicians for each of the 9 measures based on the log odds of achieving the performance measure using multilevel logistic regression models. This accounted for both clustering of patients within primary care physicians and for heteroscedasticity produced by the varying number of quality measurements for each primary care physician. We ranked primary care physicians after each of the 3 following stages of adjustment. Model 1 was unadjusted, model 2 was adjusted for practice and visit frequency, and model 3 was adjusted for practice, visit frequency, and all patient panel variables (patient age, sex, number of comorbidities, race/ethnicity, primary language spoken, and insurance status). In model 2, we first adjusted for primary care physician practice site to account for unmeasured practice characteristics and number of practice visits to allow for fairer assessment of primary care physician quality by adjusting for differences in direct patient contact (and therefore differences in direct opportunities to order the indicated tests or make appropriate management changes to achieve optimal hemoglobin A1c or low-density lipoprotein cholesterol control). Geo-coded variables for median household income and high school graduation rates were excluded in the final model because of minimal additional effect after adjustment for patient-level variables (mean change in primary care physician ranking of 1.9 percentiles).

We calculated physician composite rankings at each stage of adjustment as
the average of all individual scores for each of the 9 available HEDIS measures. The final composite score was converted to percentiles (range 1-100) with a lower percentile indicating higher primary care physician quality ranking. Each primary care physician thus had a composite quality score after each of the 3 stages of adjustment.

Using unadjusted composite quality scores (model 1), we compared physician and patient panel characteristics between physicians ranked in the top vs bottom quality tertile using the χ² test, the Wilcoxon rank sum test, or the t test. Then, we examined changes in relative physician rankings from the unadjusted base model (model 1) attributable to patient panel differences (model 3). All variables except for the intercept were considered as fixed effects in the multilevel regression analysis.

At each stage of adjustment, we identified physicians with a greater than 5 and 10 percentile absolute change in their relative composite quality ranking. Because many performance incentive programs focus on tiers of physician quality, we also divided the 162 primary care physicians into tertiles based on their unadjusted composite quality rankings and assessed for reclassification into different postadjustment tertiles. We chose to classify physicians by tertiles because division into 3 quality categories is more intuitive than 2, and because it decreases the opportunity for misclassification, thus serving as a conservative classification approach. Finally, among the subset of physicians with a greater than 10 percentile change in their relative composite quality rankings from model 1 (unadjusted) to model 3 (fully adjusted), we used χ², Wilcoxon rank sum, and t tests to compare the physician and patient panel characteristics of physicians who went up vs down in their composite quality rankings.

We used a threshold P value of less than .05 to determine statistical significance. Data were missing for less than 1.8% of all variables, and we performed a complete case analysis. We used SAS statistical software version 9.1.3 (SAS Institute Inc, Cary, North Carolina) for all analyses except for multilevel models that were estimated using MLwiN multilevel modeling software version 2.11 (Centre for Multilevel Modelling, London, England). The Massachusetts General Hospital institutional review board approved the study.

**RESULTS**

The 162 primary care physicians eligible for analysis were a mean of 18.6 years (95% confidence interval [CI], 16.9-20.2 years) from medical school graduation (TABLE 1). Fewer top tertile compared with bottom tertile primary care physicians practiced in community health centers (17.0% [95% CI, 6.9%-27.1%] vs 47.2% [95% CI, 33.7%-60.6%], respectively; P = .001). A greater proportion of female compared with male primary care physicians were in the top tertile of unadjusted primary care physician composite ranking (62.3% [95% CI, 49.2%-75.3%] vs 34.0% [95% CI, 21.2%-46.7%], respectively; P = .001).

---

**Table 1. Primary Care Physician (PCP) and Patient Characteristics**

<table>
<thead>
<tr>
<th>PCPs by Unadjusted Composite Quality Score Tertile</th>
<th>Top (n=53)</th>
<th>Bottom (n=53)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCPs Overall (N=162)</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite score percentile</td>
<td>50.5 (46.0-55.0)</td>
<td>17.1 (14.4-19.7)</td>
<td>84.0 (81.4-86.6)</td>
</tr>
<tr>
<td>Period since medical school graduation, y</td>
<td>18.6 (16.9-20.2)</td>
<td>18.8 (16.3-21.3)</td>
<td>17.3 (14.2-20.5)</td>
</tr>
<tr>
<td>Community health center practice</td>
<td>48 (29.6) (22.6-36.7)</td>
<td>9 (17.0) (6.9-27.1)</td>
<td>25 (47.2) (33.7-60.6)</td>
</tr>
<tr>
<td>Female sex</td>
<td>79 (48.8) (41.1-56.5)</td>
<td>33 (62.3) (49.2-75.3)</td>
<td>18 (34.0) (21.2-46.7)</td>
</tr>
<tr>
<td>Panel age, y</td>
<td>773 (706-841)</td>
<td>746 (631-861)</td>
<td>832 (700-965)</td>
</tr>
<tr>
<td>Patients (n = 125-303)</td>
<td>49.2 (47.9-50.4)</td>
<td>51.1 (49.6-52.6)</td>
<td>46.6 (43.8-49.5)</td>
</tr>
<tr>
<td>No. of comorbidities</td>
<td>0.85 (0.79-0.91)</td>
<td>0.91 (0.83-0.98)</td>
<td>0.80 (0.66-0.95)</td>
</tr>
<tr>
<td>Female sex, %</td>
<td>40.3 (36.6-44.0)</td>
<td>34.2 (27.6-40.8)</td>
<td>47.5 (42.0-53.0)</td>
</tr>
<tr>
<td>Minority race/ethnicity, %</td>
<td>19.2 (16.7-21.7)</td>
<td>13.7 (10.6-16.7)</td>
<td>25.6 (20.2-31.1)</td>
</tr>
<tr>
<td>Non-English-speaking, %</td>
<td>6.3 (4.3-8.3)</td>
<td>3.2 (0.7-5.6)</td>
<td>10.2 (5.5-14.9)</td>
</tr>
<tr>
<td>Pvtately insured, %</td>
<td>69.7 (67.2-72.3)</td>
<td>72.4 (69.5-75.3)</td>
<td>66.0 (60.4-71.6)</td>
</tr>
<tr>
<td>Medicare, %</td>
<td>17.5 (15.4-19.6)</td>
<td>18.0 (16.0-20.1)</td>
<td>16.7 (11.7-21.7)</td>
</tr>
<tr>
<td>Medicaid or uninsured, %</td>
<td>12.8 (11.0-14.5)</td>
<td>9.6 (7.5-11.7)</td>
<td>17.2 (13.5-21.0)</td>
</tr>
<tr>
<td>High school graduate, %</td>
<td>85.7 (84.5-86.9)</td>
<td>87.9 (86.5-89.4)</td>
<td>82.7 (80.3-85.1)</td>
</tr>
<tr>
<td>Family household income in US $</td>
<td>60.155 (58.178-62.131)</td>
<td>63.901 (61.118-66.684)</td>
<td>53.890 (50.518-57.261)</td>
</tr>
</tbody>
</table>

> 3 Primary care practice visits⁎, %                     | 67.3 (65.3-69.3) | 71.0 (68.5-73.5) | 61.8 (57.3-66.3) | .003 |

Abbreviation: CI, confidence interval.
⁎ Unless otherwise indicated.
*Calculated using the y² test, the Wilcoxon rank sum test, or the t test.
**Based on 9 Healthcare Effectiveness Data and Information Set measures. All PCPs were ranked and given percentile scores (a lower score indicates higher overall quality).
*Values are expressed as number (percentage) [95% CI].
*Based on 2007 US Census block group data.

©2010 American Medical Association. All rights reserved.
physician clinical performance rankings

Table 2. Changes in Primary Care Physician (PCP) Rankings on Composite Quality Scores

<table>
<thead>
<tr>
<th>Change in PCP rankings, mean percentile (95% CI)</th>
<th>Model 2‡</th>
<th>Model 3‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in PCP rankings, mean percentile (95% CI)</td>
<td>15.1 (12.9-17.3)</td>
<td>7.6 (6.6-8.7)</td>
</tr>
<tr>
<td>PCPs with change in rankings, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5 percentiles</td>
<td>122 (75.3)</td>
<td>95 (58.6)</td>
</tr>
<tr>
<td>&gt;10 percentiles</td>
<td>86 (53.1)</td>
<td>52 (32.1)</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.
‡Adjusted for practice and visit frequency.
§Additionally adjusted for age, sex, comorbidity, race/ethnicity, language, and insurance status.

Figure. Absolute Ranking Change for Primary Care Physicians From Practice- and Visit Frequency-Adjusted Model After Additional Adjustment for Patient Panel Characteristics

The proportion of minority patients (13.7% [95% CI, 10.6%-16.7%] vs 25.6% [95% CI, 20.2%-31.1%; P < .001), non–English-speaking patients (3.2% [95% CI, 0.7%-5.6%] vs 10.2% [95% CI, 5.5%-14.9%; P < .001), and patients with Medicaid coverage or without insurance (9.6% [95% CI, 7.5%-11.7%] vs 17.2% [95% CI, 13.5%-21.0%; P < .001) was significantly lower in the top vs bottom tertile, respectively, of primary care physicians. Patients of top vs bottom tertile primary care physicians also lived in neighborhoods with higher median household incomes ($63 901 [95% CI, $61 118-$66 684] vs $53 890 [95% CI, $50 518-$57 261], respectively; P < .001) and higher high school graduation rates (87.9% [95% CI, 86.5%-89.4%] vs 82.7 [95% CI, 80.3-85.1]; P < .001).

Adjustment for practice and visit frequency (model 2) led to marked changes in relative physician quality rankings with more than 75% of primary care physicians changing by more than 5 percentiles and more than half of primary care physicians changing by more than 10 percentiles following adjustment (Table 2). Physician rankings changed by an average of 15.1 percentiles (95% CI, 12.9-17.3 percentiles) following adjustment for practice and visit frequency.

Additional adjustment for patient panel characteristics led to further re-ranking of primary care physicians, with 58.6% of primary care physicians changing by more than 5 percentiles and 32.1% changing more than 10 percentiles. The mean change in physician rankings attributable to further adjustment for patient characteristics was 7.6 percentiles (95% CI 6.6-8.7 percentiles). The Figure shows the distribution of absolute changes in rank among primary care physicians following this adjustment.

When comparing the fully adjusted model (model 3) with the model adjusted for practice and visit frequency (model 2), 11.3% of primary care physicians originally in the top composite score tertile decreased in ranking to the middle tertile, 14.3% of bottom tertile primary care physicians increased in ranking to the middle tertile, and 25.0% of middle tertile primary care physicians moved into the top or bottom tertile (Table 3).

The 34 primary care physicians whose relative quality rankings increased by more than 10 percentiles after adjustment were more likely to be physicians practicing at community health centers and have larger overall panel sizes. These physicians were also more likely to have panels with a higher proportion of minority, non–English-speaking, and younger patients with fewer comorbidities compared with the 44 primary care physicians whose relative quality rankings decreased after adjustment (Table 4).

COMMENT

We examined relative physician quality rankings using an aggregate of commonly used HEDIS measures in a cohort of primary care physicians working within a single integrated health system. Primary care physicians in the top tertile of measured quality were more likely to care for older patients with greater comorbidity who made more frequent visits to see a primary care physician. This finding is consistent with prior studies. Because older patients with more comorbidities are often seen more frequently, they may have stronger relationships with their physicians, and physicians caring for such patients may have more opportunities to complete process measures. Also, in concordance with literature in health care disparities, top tertile primary care physicians were less likely to care for minority, non–English-speaking, Medicaid, and uninsured patients compared with bottom tertile primary care physicians.

After accounting for practice site and visit frequency differences, adjusting for patient panel factors resulted in an additional relative mean rank change of...
These results have considerable policy implications given the increasing role of physician performance incentive programs. Prior studies have posited numerous potential unintended consequences of these programs. Our study demonstrates a potential mechanism through which performance incentive programs might worsen health care disparities. Because physicians and practices with higher quality scores receive higher payments and recognition under these programs, these incentive approaches could erroneously distribute resources away from high-quality physicians caring for more vulnerable patients and worsen health care disparities.

Additionally, prior studies by our group demonstrate that socioeconomically deprived patients are less likely to be connected to their physicians and make fewer visits to their primary care physicians. These practice-connected rather than primary care physician–connected patients were more likely to be found in community health centers. Beyond issues surrounding patient attribution, decreased connectedness to physicians could exacerbate the unintended consequences of these programs. These consequences are more likely to be found in community health centers.9 Beyond issues surrounding patient attribution, decreased connectedness to physicians could exacerbate the unintended consequences of these programs.

Table 3. Changes in Primary Care Physician (PCP) Composite Score Tertile

<table>
<thead>
<tr>
<th>Initial Tertile</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (n=53)</td>
<td>47 (78.7)</td>
<td>6 (11.3)</td>
<td>0</td>
</tr>
<tr>
<td>2 (n=56)</td>
<td>6 (10.7)</td>
<td>42 (75.0)</td>
<td>8 (14.3)</td>
</tr>
<tr>
<td>3 (n=53)</td>
<td>0</td>
<td>8 (14.3)</td>
<td>45 (84.9)</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.
aAdjusted for practice and visit frequency (model 2).
bAdditionally adjusted for patient age, sex, comorbidity, race/ethnicity, language, and insurance status (model 3).

Table 4. Primary Care Physician (PCP) and Patient Characteristics From Unadjusted Baseline Model (Model 1) to Fully Adjusted Model (Model 3)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians Composite score percentile</td>
<td>64.3 (55.8-72.8)</td>
<td>.0001</td>
</tr>
<tr>
<td>Period since medical school graduation, y</td>
<td>16.1 (12.5-19.7)</td>
<td>.08</td>
</tr>
<tr>
<td>Community health center practice</td>
<td>16 (47.1) [30.3-63.8]</td>
<td>.01</td>
</tr>
<tr>
<td>Female sex</td>
<td>20 (58.8) [42.3-75.4]</td>
<td>.03</td>
</tr>
<tr>
<td>Patient panel size</td>
<td>946 (804-1088)</td>
<td>.02</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.
aUnless otherwise indicated.
bCalculated using the χ² test, the Wilcoxon rank sum test, or the t test.
cCalculated using the t test. The Wilcoxon rank sum test, or the t test.
dBased on 9 Healthcare Effectiveness Data and Information Set measures. All PCPs were ranked and given percentile scores (a lower score indicates higher overall quality).

©2010 American Medical Association. All rights reserved.
effect of inadequate adjustment for patient characteristics, leading to even lower incentive-based reimbursement for physicians working in community health centers and other safety-net sites. This reduction in income could lead to a reduction in both the number of physicians working in such communities and practice-level resources available to invest in structures and processes to improve quality. Physicians in these areas would be at an additional disadvantage due to traditionally decreased access to resources and a payer mix that is likely to include a higher proportion of uninsured patients and Medicaid recipients. Further studies using real-world performance incentive models may help us understand the effect on redistribution of resources within and between health care systems.

Our aim was not to develop a method for case-mix adjustment, but to study the effect of patient characteristics on physician clinical performance ranking. Further work in designing appropriate case-mix adjustment methods for physician clinical performance assessment is necessary. Case-mix adjustment may help improve physician acceptance of physician-level profiles by leveling the playing field and attenuating the effect of potential unintended consequences, and decrease the chance that physicians will seek to exclude patients likely to worsen their measured quality. These adjustment methods should consider the effect of the practice characteristics, including resources and infrastructure. However, because visit frequency is at least partially due to physician practice style, it should be excluded and further work is warranted to identify best practice follow-up appointment intervals for different clinical scenarios. Lastly, novel methods and performance incentive schemes must be paired with ongoing stratification of quality performance by sociodemographic characteristics to avoid giving undue credit to physicians and practices for lower performance with vulnerable patients and ensure that disparities in care are made the target of performance incentive programs rather than obscured by case-mix adjustment practices.

Our study strengths included the use of easily obtained patient characteristics and widely used clinical quality measures (drawn from system databases) that represent actual receipt of care rather than reported data. Furthermore, comparing physicians using an aggregate quality measure represents an overall quality construct, allows a greater number of patients to be included, and minimizes the influence of any single measure-specific outcome. The compositing method used in our study allows for easy addition of measures as well as transparent weighting when necessary. However, despite the advantages of such composite quality measures, few quality measurement tools incorporate other dimensions of clinical quality (eg, physician empathy and communication skills), and it remains important to continue evaluating physician clinical performance on individual measures because quality improvement is more easily targeted at specific outcomes and because one measure may not predict performance on others.

We must interpret our results within the context of the study design. While it is a strength to compare the association of patient characteristics on clinical performance measures among a relatively homogenous physician cohort, our results may not be generalizable to other health care systems such as community health systems or smaller private practice networks that makeup the majority of health care delivery in the United States. In addition, use of currently available quality measures that are not comprehensive, are process-oriented, and may not measure actual physician quality has limitations. Nonetheless, because these HEDIS-based measures are widely used, our findings remain directly applicable to current quality measurement practices. Another consideration is that quality estimates for some physicians may be more reliable for one measure than another. Although we addressed heterogeneity in sample sizes between physicians using multilevel modeling techniques, this continues to be a central hurdle for physician-level clinical performance measures. Finally, although adjustment for patient variables in this study may have contributed to a more accurate estimation of quality outcomes, they only account for a fraction of the total variability in physician-level quality scores. Additional work is necessary to develop a wider range of reliable patient-, physician-, and practice-level variables.

In summary, our study demonstrates that patient panel characteristics are associated with the relative measured quality of physicians within a large academic primary care network. Adjustment for differences in patient panel characteristics resulted in significant reclassification of top tier vs bottom tier physicians. To the extent that health systems reward physicians for higher measured quality of care, lack of adjustment for patient panel characteristics may penalize physicians for taking care of more vulnerable patients, incentivize physicians to select patients to improve their quality scores, and result in the misallocation of resources away from physicians taking care of more vulnerable populations. Conversely, adjustment for patient panel characteristics may remove the incentive to improve care or may inappropriately reward lower-quality physicians caring for more vulnerable patients. Efforts to improve quality of care must address both fairness of physician clinical performance assessment and the design of incentive schemes to both provide equitable distribution of resources and reduce disparities in care for vulnerable patients.

Author Affiliations: Schools of Medicine (Drs Hong, Atlas, Chang, Barry, and Grant) and Public Health (Dr Subramanian), Harvard University, Boston, Massachusetts; and General Medicine Division, Massachusetts General Hospital, Boston (Drs Hong, Atlas, Chang, Barry, and Grant, and Mr Ashburner).

Author Contributions: Dr Hong had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Hong, Atlas, Chang, Subramanian, Barry, Grant.

Acquisition of data: Hong, Atlas, Ashburner.
Analysis and interpretation of data: Hong, Atlas, Chang, Subramanian, Ashburner, Grant. Drafting of the manuscript: Hong, Grant.

Critical revision of the manuscript for important intellectual content: Hong, Atlas, Chang, Subramanian, Ashburner, Barry, Grant.

Statistical analysis: Hong, Chang, Subramanian, Ashburner.

Obtained funding: Grant.

Administrative, technical, or material support: Hong, Atlas, Subramanian, Ashburner, Grant.

Study supervision: Atlas, Subramanian, Grant.

Financial Disclosures: None reported.

REFERENCES


©2010 American Medical Association. All rights reserved.

(Published) JAMA, September 8, 2010—Vol 304, No. 10 1113

Downloaded from jama.ama-assn.org at Harvard University on February 7, 2011

PHYSICIAN CLINICAL PERFORMANCE RANKINGS

1113

Funding/Support: Dr Hong is supported by National Research Service Award grant T32 HP12706 from the National Institutes of Health. Dr Grant is supported by career development grant NIDDK K23 DK067452 from the National Institutes of Health. Dr Atlas is supported by grant R18 HS181611 from the Agency for Healthcare Research and Quality. Dr Subramanian is supported by career development grant NHLBI K25 HL081275 from the National Institutes of Health.

Role of the Sponsor: The National Institutes of Health did not participate in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

Additional Contributions: Timothy Ferris, MD, MPH, the medical director of the Massachusetts General Physicians Organization and associate professor of medicine at Harvard Medical School, provided helpful discussions on the manuscript. Dr Ferris did not receive any compensation for his contribution. Guoping Huang from the Harvard University Center for Geographic Analysis helped us create the US Census block group data on socioeconomic status using geo-codes. The Harvard University Center for Geographic Analysis was paid a small fee for these services.

None reported.

September 8, 2010—Vol 304, No. 10 1113

Small fee for these services.