Higher Health Care Quality And Bigger Savings Found At Large Multispecialty Medical Groups

ABSTRACT The belief that integrated delivery systems offer better care at lower cost has contributed to growing interest in accountable care organizations. These provider-led delivery systems would accept responsibility for their primary care populations and would have financial incentives for improving care and reducing costs. We investigated this belief by comparing the costs and quality of care provided to Medicare beneficiaries in twenty-two health care markets by physicians who did and did not work within large multispecialty group practices affiliated with the Council of Accountable Physician Practices. In most markets, and after adjustment for patient factors, group physicians affiliated with the council provided higher-quality care at a 3.6 percent lower annual cost ($272 per patient).

A growing body of evidence points to substantial opportunities to improve the performance of the U.S. health care system. The Institute of Medicine (IOM) and others have documented serious deficiencies in health care quality. The quality and costs of health care vary dramatically across regions and health systems within the United States. Higher-spending systems, on average, provide lower quality of care and achieve equal or worse outcomes. These studies point to large potential opportunities to improve both the quality and the cost of care—that is, to increase the efficiency of the U.S. health care delivery system.

Many have argued that large multispecialty group practices provide higher-quality care at lower costs than small group practices. Some policy analysts suggest that U.S. policy should foster the growth of accountable care organizations—provider-led, local health care delivery systems that would accept responsibility for the continuum of care for their primary care patients and have financial incentives to improve quality and reduce the costs of care. Large multispecialty group practices would be well positioned to serve as accountable care organizations. Evidence, however, has been limited.

To provide insight into the potential advantages of such integrated group practices, we identified Medicare beneficiaries who received care from physicians in the large multispecialty group practices that are members of the Council of Accountable Physician Practices (CAPP). The council is an affiliate of the American Medical Group Association and constitutes a consortium of twenty-seven large U.S. multispecialty group practices. Members recognize accountability for cost and quality as a model for transforming the U.S. health care system. We compared the care given these Medicare patients on selected measures of ambulatory care quality and cost with that of Medicare beneficiaries receiving care from physicians in the same markets who were not affiliated with a CAPP group.

Study Data And Methods The self-selected medical groups that participate in CAPP are all multispecialty physician groups that range in size from approximately 250 to 13,000 physicians. Of the current twenty-seven...
council members, eleven are integrated with hospital systems; four are integrated with their own health plans; eight are integrated with both hospitals and a health plan; and four are not affiliated with a hospital or a health plan. Twenty groups, working in twenty-two markets, agreed to participate in this study and provided the names, specialties, and Unique Physician Identification Numbers (which Medicare assigns to physicians) of physicians with whom the practice had an affiliation for any part of calendar year 2005. In addition, the participating groups identified the date on which any physicians left or joined the practice in 2005.

In each of the twenty-two markets, we identified all physicians who were not affiliated with the participating CAPP groups. We assigned patients to physicians, then computed quality and cost measures for fee-for-service Medicare beneficiaries who received the predominance of their ambulatory care from CAPP-affiliated physicians in 2005. In each market, we computed these measures for beneficiaries assigned to all non-council group physicians. This study was approved by Dartmouth Medical School’s Committee for the Protection of Human Subjects.

**Study and Control Populations**

The study population was defined as fee-for-service Medicare beneficiaries age sixty-five and older who were in a 20 percent random sample of the Medicare population. Each beneficiary had at least one non-inpatient visit to a physician in 2005 or 2006, was continuously enrolled in Medicare Part B throughout 2005, and was not enrolled in a Medicare health maintenance organization (HMO) in 2005. To assign beneficiaries to their medical groups, we used a multistep process:

▸ **Assignment to a Physician:** Using methods developed in an earlier study, we assigned each beneficiary to a unique physician. To do so, we used all of patients’ ambulatory care claims over a two-year period beginning with their first visit in 2004 and continuing through December 2006. Beneficiaries were assigned first to a primary care physician, then a medical specialist, and finally to another physician, as follows. If possible, we first assigned patients to a primary care physician, defined as a general internist, a general practitioner, a family physician, or a geriatrician, who provided the plurality of visits during the two-year assignment period.

In the case of a tie, patients were assigned to the physician with the greatest number of days between billing dates, which more likely indicates a longitudinal relationship. If beneficiaries could not be assigned to a primary care physician, we then assigned them to the medical specialist (such as a cardiologist, pulmonologist, or gastroenterologist) who provided the plurality of visits during the two-year time period examined.

Finally, if beneficiaries had no primary care or medical specialist visits, we assigned them to the other physician who provided the plurality of visits during the two-year period.

Of the 741,448 patients included in this study, 91.4 percent were assigned to a primary care physician, 7.4 percent to a medical specialist, and 1.2 percent to another specialist.

▸ **Assignment to CAPP Group Practices:** We assigned physicians and their patients to CAPP groups only if their assigned physicians were affiliated with that group throughout 2005—that is, the physician neither joined nor left the group with which he or she had an affiliation between 1 January and 31 December 2005. The patients of physicians who joined or left a CAPP group during 2005 were dropped from the analysis, because we could not be sure whether their care reflected group performance.

▸ **Assignment of CAPP Groups to Markets:** CAPP groups were assigned to Hospital Referral Regions, which are health care markets for fee-for-service Medicare beneficiaries defined by the Dartmouth Atlas Project. We included in the analysis only those regions in which there were at least 100 beneficiaries assigned to a CAPP group’s physicians.

Two of the twenty groups that provided Unique Physician Identification Numbers for their affiliated physicians operated in multiple contiguous regions. For these groups, we simply aggregated those regions for analytic purposes.

For the single group with a presence in three geographically disparate regions, we analyzed each of the subgroups in those three regions as a separate group. Therefore, although twenty CAPP groups provided information on their affiliated physicians, our analysis treated these groups as though they were twenty-two groups in twenty-two markets. Eight of these twenty-two groups paid their physicians primarily on a salary basis, as opposed to a production basis.

▸ **Comparison Population:** Medicare beneficiaries assigned to non-CAPP physicians practicing in the same Hospital Referral Regions constituted the comparison group for each of the twenty-two markets under study. Because physicians could have multiple practice addresses, we defined non-CAPP physicians’ practice locations by the location of the hospital where they did a preponderance of their inpatient work or, in the instances where they did no inpatient work, where most of the beneficiaries to whom they provided care were admitted during the two-year assignment period, as in previous work.

**Measures** From the denominator file, we
obtained several patient-level demographic variables: age, sex, race, ZIP code of residence, whether the beneficiary was concurrently enrolled in Medicaid, and whether the beneficiary was originally eligible for Medicare by virtue of disability, as opposed to age. We used diagnoses recorded on hospital and physician evaluation and management claims data from the previous year, 2004, to identify beneficiaries with diabetes and to calculate a Charlson comorbidity score, a measure of disease burden that increases as patients have more comorbidities.

We used data from the U.S. census to determine the mean annual family income within a ZIP code. We categorized low-income ZIP codes as those with a mean annual family income of less than $24,000; high-income ones had mean income greater than $55,000.

We used data from the Dartmouth Atlas to calculate Medicare HMO penetration of the Hospital Referral Regions in which the CAPP groups had a market presence. Finally, from the groups, we obtained binary information on whether, as a whole, their affiliated physicians were paid primarily on a salary basis or a production basis.

We calculated five measures of the quality of ambulatory care from claims for relevant services submitted by any provider during the period examined, drawn from measures approved by the Ambulatory Quality Alliance. These include mammographic screening for breast cancer for women ages 65–69; annual hemoglobin A1c testing among diabetics ages 65–75; annual lipid testing among diabetics ages 65–75; annual fundoscopy (examination of the retina) by an ophthalmologist among diabetics ages 65–75; and the proportion of diabetics ages 65–75 who received all three of the diabetic quality measures. We calculated crude and risk-adjusted measures.

As a sixth indicator of quality, we calculated the crude and risk-adjusted numbers of ambulatory care–sensitive admissions, such as admission for hypertension or for chronic obstructive pulmonary disease, as defined by the Agency for Healthcare Research and Quality. With ambulatory care–sensitive conditions, a higher number suggests a higher number of potentially preventable admissions.

We calculated four measures of costs: mean standardized physician spending per Medicare patient; mean standardized inpatient spending per patient, which includes spending for acute hospitalizations, long-term care facilities, and skilled nursing facilities; mean standardized home health care costs per patient; and mean total standardized Centers for Medicare and Medicaid Services (CMS) program payments per patient. Standardization of cost measures for hospital and physician services was carried out by assigning a single price to each specific Medicare-financed encounter, thereby eliminating cost differences attributable to geographic differences in wages, insurance rates, and disproportionate payments for teaching hospitals, as in other studies.

For other services, Medicare payments were adjusted by the Medicare Geographic Practice Cost Index, which takes into account the fact that labor costs vary in different markets. To reduce undue influence of outliers, we truncated cases where costs were above the ninety-ninth percentile. Cost measures were adjusted for differences across systems in patient age, race, sex, imputed income, Medicaid status, and comorbidity score.

**Analysis** We performed several analyses. First, within each market, we compared results for Medicare beneficiaries who were assigned to a CAPP physician with results for those who were not assigned to such a physician. For each referral region (or aggregation, in the instances where a council group spanned more than one contiguous region), we used chi-square tests to compare quality measures and two-sample t-tests to compare costs of CAPP to non-CAPP group patients in the same market.

When a council group had a market presence in multiple Hospital Referral Regions, we weighted the control population by region such that they would be proportionally represented. For ambulatory care–sensitive admissions, we used Poisson regression to calculate differences in admission rates for CAPP-assigned beneficiaries compared to their control counterparts. We defined differences as statistically significant when \( p < 0.01 \) (two-tailed).

In addition, we conducted five multivariate regression analyses, where each quality or standardized cost variable was the dependent variable. For the binary quality measures, we used logistic
regression. Because the likelihood of obtaining a quality measure in our study population was high, we converted odds ratios to relative risk ratios using standard methods. For costs, we used analysis of variance methods. For ambulatory care-sensitive hospitalizations, we used Poisson regression.

To compare overall costs and quality among all CAPP groups, we used a fixed-effects model in which the referral region was the fixed effect. Although we believe that the adjusted model is most appropriate for costs and the crude model for comparing quality, we also evaluated a crude cost model and an adjusted quality model as sensitivity analyses.

We conducted another analysis in which the level of Medicare HMO penetration was treated as a continuous fixed effect, replacing the referral region fixed effect. In a final analysis, we treated the primary form of compensation of physicians as a binary fixed effect. We used the software program SAS, version 9.1, for all statistical analyses.

Study Results
Medicare beneficiaries receiving care from physicians in CAPP groups differed from those cared for in non-CAPP groups in terms of their demographics, socioeconomic status, and levels of comorbidity. Overall, CAPP-group patients were slightly but significantly more likely to be younger and male, and slightly less likely to be black, to be concurrently enrolled in Medicaid, and to live in a low-income ZIP code.

In addition, CAPP patients were significantly healthier, with fewer patients having at least one comorbidity, as demonstrated by having a Charlson score of one or higher (Exhibit 1). However, we found much variation in demographics across markets (Online Appendix Exhibit 1).

### EXHIBIT 1

**Characteristics Of Medicare Beneficiaries Assigned To Council Of Accountable Physician Practices (CAPP) Group Physicians, Those Assigned To Control Physicians, And U.S. Medicare Beneficiaries**

<table>
<thead>
<tr>
<th></th>
<th>CAPP 63,647</th>
<th>Control 677,901</th>
<th>U.S. beneficiaries 5,124,162</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of fee-for-service beneficiaries</td>
<td>63,647</td>
<td>677,901</td>
<td>5,124,162</td>
</tr>
<tr>
<td>Average age (years)</td>
<td>77.0</td>
<td>77.3</td>
<td>76.3</td>
</tr>
<tr>
<td>Percent female</td>
<td>57.9</td>
<td>60.1</td>
<td>59.7</td>
</tr>
<tr>
<td>Percent black</td>
<td>5.9</td>
<td>5.4</td>
<td>7.3</td>
</tr>
<tr>
<td>Percent concurrently enrolled in Medicaid</td>
<td>12.6</td>
<td>15.4</td>
<td>13.9</td>
</tr>
<tr>
<td>Percent living in ZIP code with annual family income less than $24,000</td>
<td>2.1</td>
<td>2.8</td>
<td>4.1</td>
</tr>
<tr>
<td>Percent with Charlson score of 1 or higher</td>
<td>31.4</td>
<td>34.1</td>
<td>34.1</td>
</tr>
</tbody>
</table>

**Source** Authors’ analysis of Medicare data. **Note** All CAPP-to-control comparisons are statistically significant at $p < 0.001$ (two-tailed).

**VARIATION IN QUALITY MEASURES** Across CAPP groups, we found wide variation in all measures of quality examined: the hemoglobin A1c measure had the narrowest variation (73.1–93.5 percent), and comprehensive diabetic testing had the widest variation (38.7–80.0 percent; Online Appendix Exhibit 2). In general, fee-for-service Medicare beneficiaries who were assigned to Council of Accountable Physician Practices group physicians received more evidence-based care than those in the same markets who were not assigned to such physicians. Similar results were found for annual lipid testing and annual fundoscopy among diabetics (data not shown).

**VARIATION IN COSTS** Similarly, we found substantial variation in standardized costs of care across CAPP groups (Online Appendix Exhibit 3). Annual physician spending varied from $2,029 to $4,063; annual hospital spending varied from $1,795 to $2,206; and total Medicare spending varied from $5,591 to $7,394. When compared with patients cared for by physicians not affiliated with CAPP groups, fee-for-service Medicare beneficiaries cared for by CAPP groups had lower physician spending in twelve of the twenty-two markets (nine statistically so), lower inpatient spending in fifteen markets (four statistically so), lower home health spending in seventeen markets (eight statistically so), and lower total Medicare spending in fifteen markets (seven statistically so).

Unadjusted analyses showed that fee-for-service Medicare beneficiary assignment to a CAPP-group physician was associated with higher quality and lower costs (Exhibit 2) on all examined measures, whether compared to the U.S. mean or to non-CAPP physicians in their same markets. Using crude measures, standardized physician spending was $239 (8.0 percent) lower; standardized hospital spending was $235
(9.7 percent) lower; and total standardized Medicare payments were $540 (7.1 percent) lower for patients who saw CAPP-group physicians as compared to those who did not.

**Patients’ Characteristics** Adjustment for patient characteristics had little effect on the council versus noncouncil differences for most quality measure scores, but it reduced the cost savings associated with council groups by half. The adjusted models showed that Medicare beneficiaries cared for by CAPP physicians received 5–15 percent higher quality of care, depending on the measure used, at a cost that was $272 (3.6 percent) lower than provided by non-council physicians (95 percent confidence interval: $189–$533).

These findings were consistent across all specifications of the regression models, including those that adjusted for market-specific Medicare HMO penetration and for the primary form of physician compensation (data not shown). Method of physician compensation was not independently associated with different outcomes.

**Discussion**

Our findings provide some support for advocates of large multispecialty group practice and of accountable care organizations. Although the cost and quality differences between Council of Accountable Physician Practices groups and all other physicians were not large, they were consistent across all measures and across most markets. For a patient, an 8 percent lower risk of having an avoidable admission is not trivial, nor is a 15 percent better chance of having all three diabetic tests performed.

For Medicare, although a 3.6 percent cost savings is relatively small, if all physicians could perform at this level, about $15 billion a year in savings to the Medicare program would be generated. This would amount to $150 billion over ten years—enough to make a substantial contribution to the $940 billion estimated cost of the health care legislation enacted in March 2010.

**Study Limitations** This study has several important limitations. First, our study focused on a sample of twenty large multispecialty groups, all of which had self-selected to participate in the

---

**EXHIBIT 2**

Comparison Of Crude And Adjusted Quality And Cost Measures For Council Of Accountable Physician Practices (CAPP) Physicians And Non-CAPP Physicians

<table>
<thead>
<tr>
<th>Quality measures</th>
<th>Crude measures</th>
<th>Adjusted measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U.S.</td>
<td>Non-CAPP</td>
</tr>
<tr>
<td>Mammography in women ages 65–69</td>
<td>50.4%</td>
<td>57.9%</td>
</tr>
<tr>
<td>Annual HbA1c testing for diabetics ages 65–75</td>
<td>80.3%</td>
<td>85.2%</td>
</tr>
<tr>
<td>Annual lipid testing for diabetics ages 65–75</td>
<td>86.7%</td>
<td>91.1%</td>
</tr>
<tr>
<td>Annual fundoscopy for diabetics ages 65–75</td>
<td>67.8%</td>
<td>75.3%</td>
</tr>
<tr>
<td>Completion of all three diabetic tests</td>
<td>53.9%</td>
<td>63.4%</td>
</tr>
<tr>
<td>ACS admission rate rate per 100</td>
<td>8.3</td>
<td>6.9</td>
</tr>
</tbody>
</table>

**Cost measures**

<table>
<thead>
<tr>
<th>Cost measures</th>
<th>Crude measures</th>
<th>Adjusted measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U.S.</td>
<td>Non-CAPP</td>
</tr>
<tr>
<td>Standardized MD in 2005</td>
<td>$2,881</td>
<td>$2,764</td>
</tr>
<tr>
<td>Standardized hospital spending in 2005</td>
<td>$2,405</td>
<td>$2,193</td>
</tr>
<tr>
<td>Total standardized CMS payments in 2005</td>
<td>$7,406</td>
<td>$7,053</td>
</tr>
</tbody>
</table>

**SOURCE** Authors’ analysis of Medicare data. **NOTES** Adjusted for Hospital Referral Region fixed effects and patient age, sex, race, imputed income, Medicare enrollment through disability, concurrent Medicaid enrollment, and Charlson score. For all quality measures save ambulatory care–sensitive (ACS) admission rates, when relative risk ratios are greater than 1, CAPP groups perform better as a whole than physicians in the same markets who are not associated with such a group; for ACS admission rates, a relative risk ratio that is less than 1 indicates that CAPP groups perform better. A negative cost measure coefficient indicates that CAPP groups provide services at a lower standardized Medicare cost as a whole than physicians practicing in the same markets who are not associated with such a group. CI is confidence interval.
Second, our quality measures were relatively few and are likely to be sensitive to automated processes that large groups are more likely to have resources to implement, such as reminders for test scheduling. We did not provide evidence for other measures of quality—for example, the quality of diagnosis or the degree to which care was responsive to patients’ preferences. These are important but are much harder to measure. Nonetheless, the measures we used include important dimensions of care for the Medicare population and have been used in prior studies of the quality of beneficiaries’ care.24

Third, we were not able to determine why or how CAPP groups provided higher-quality care at lower costs. Fourth, our analysis was limited to fee-for-service Medicare Part A and Part B enrollees; findings may be different when examining performance for Medicare Advantage patients. Fifth, although we used Charlson scores to account for illness burden, others might find alternative methods more robust. Using those methods might result in somewhat different findings.

Finally, although we corrected for patient characteristics, the CAPP groups tended to care for patients who were younger, lived in more affluent neighborhoods, had less complex illnesses, and relied less on Medicaid to augment financial access to care. These factors could influence the relative ease with which patients can obtain, and practices can provide, preventive services. If our adjustments failed to fully correct for these factors, our findings might be explained by differences in the populations who did and did not obtain care from CAPP-group physicians.

From the perspective of current health reform discussions, the most important limitation of the study is that all of the groups were reimbursed under fee-for-service payment, which provides strong financial incentives to increase the amount and complexity of the services provided. Proposals to foster the development of accountable care organizations include modifications of fee-for-service payment (through shared savings or partial capitation) that would allow such organizations to be financially rewarded for reducing overall costs if quality benchmarks are achieved.25 The current analysis thus provides very little insight into the magnitude of potential savings that might be achieved through broad implementation of such new payment models.

### Feasibility of Moving Forward

At the same time, our findings underscore the feasibility of moving forward with the development and further testing of payment and performance measurement innovations to foster the development of accountable care organizations. Existing integrated delivery systems, multispecialty group practices, hospitals that employ their own primary care physicians, and well-established physician networks (such as some independent practice associations) are already structured in ways that would allow them to participate in these payment models. Mechanisms would need to be established to help the many physicians who still remain in small office-based practices establish more integrated networks. These networks would be capable of the kinds of practice support seen in the Council of Accountable Physician Practices groups.

### Initial Reliance on Administrative Data

Our analysis suggests that initial implementation and evaluation could rely largely on administrative data, such as those used in this analysis. These data could be supplemented with an easily collected and parsimonious set of clinically augmented quality measures that are already used to evaluate health plan performance, such as control of blood sugar levels.

### Advantages of Multispecialty Group Practices

Our findings are at least consistent with the view that large multispecialty group practices, which are more likely to have the infrastructure to support improved clinical care,26 would be able to provide higher-quality care at lower costs than other types of practices can provide. To further explore their potential, it would be wise to enable both fully integrated and networked health care systems to participate in programs to implement accountable care organizations and continually evaluate and adapt the models as experience is gained. ■
This work was funded by the Kaiser Family Foundation, the California HealthCare Foundation, and the Commonwealth Fund. The role of the funding organizations was to provide funding for study design, data collection, management, and analysis; interpretation of the data; and preparation of the manuscript. No funders had approval authority over the manuscript, although they were given opportunities to comment on it. The authors thank the Council of Accountable Physician Practices groups that participated in the study: Austin Regional Clinic, Billings Clinic, Cleveland Clinic Foundation, Everett Clinic, Fallon Clinic, Geisinger Clinic, Harvard-Vanguard Health Care System, HealthCare Partners, HealthPartners, Henry Ford Health Care System, Intermountain Clinic, Jackson Clinic, Lahey Clinic, Marshfield Clinic, Mayo Clinic, Oschner Medical Clinic, Palo Alto Medical Foundation, Sharp Rees-Stealy Clinic, Virginia Mason Medical Group, and Wenatchee Valley Medical Group. The authors also acknowledge the considerable work performed by Karen Swenson, Courtney Cote, Lorraine Bauman, Ellen Westbrook, and Tom DiGiacinto of IMS Health in Manchester, New Hampshire. Without their work, the research presented herein could not have been completed.

NOTES

23. The Online Appendix can be accessed by clicking the Online Appendix link in the box to the right of the article online.

MAY 2010 29:5 HEALTH AFFAIRS 997