SPECIAL ARTICLE

Regional Variations in Diagnostic Practices

Yunjie Song, Ph.D., Jonathan Skinner, Ph.D., Julie Bynum, M.D., M.P.H., Jason Sutherland, Ph.D., John E. Wennberg, M.D., M.P.H., and Elliott S. Fisher, M.D., M.P.H.

ABSTRACT

BACKGROUND

Current methods of risk adjustment rely on diagnoses recorded in clinical and administrative records. Differences among providers in diagnostic practices could lead to bias.

METHODS

We used Medicare claims data from 1999 through 2006 to measure trends in diagnostic practices for Medicare beneficiaries. Regions were grouped into five quintiles according to the intensity of hospital and physician services that beneficiaries in the region received. We compared trends with respect to diagnoses, laboratory testing, imaging, and the assignment of Hierarchical Condition Categories (HCCs) among beneficiaries who moved to regions with a higher or lower intensity of practice.

RESULTS

Beneficiaries within each quintile who moved during the study period to regions with a higher or lower intensity of practice had similar numbers of diagnoses and similar HCC risk scores (as derived from HCC coding algorithms) before their move. The number of diagnoses and the HCC measures increased as the cohort aged, but they increased to a greater extent among beneficiaries who moved to regions with a higher intensity of practice than among those who moved to regions with the same or lower intensity of practice. For example, among beneficiaries who lived initially in regions in the lowest quintile, there was a greater increase in the average number of diagnoses among those who moved to regions in a higher quintile than among those who moved to regions within the lowest quintile (increase of 100.8%; 95% confidence interval [CI], 89.6 to 112.1; vs. increase of 61.7%; 95% CI, 55.8 to 67.4). Moving to each higher quintile of intensity was associated with an additional 5.9% increase (95% CI, 5.2 to 6.7) in HCC scores, and results were similar with respect to laboratory testing and imaging.

CONCLUSIONS

Substantial differences in diagnostic practices that are unlikely to be related to patient characteristics are observed across U.S. regions. The use of clinical or claims-based diagnoses in risk adjustment may introduce important biases in comparative-effectiveness studies, public reporting, and payment reforms.

From the Dartmouth Institute for Health Policy and Clinical Practice (Y.S., J. Skinner, J.B., J. Sutherland, J.E.W., E.S.F.), and the Departments of Medicine and Community and Family Medicine (J.B.), Dartmouth Medical School; and the Department of Economics, Dartmouth College (J. Skinner) — all in Hanover, NH; the National Bureau of Economic Research, Cambridge, MA (J. Skinner), and the Veterans Affairs Outcomes Group, White River Junction, VT (E.S.F.). Address reprint requests to Dr. Fisher at the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, 35 Centerra Pkwy., Suite 205, Lebanon, NH 03766, or at elliott.fisher@dartmouth.edu.

This article (10.1056/NEJMsa0910881) was published on May 12, 2010, and updated on June 8, 2010, at NEJM.org.

N Engl J Med 2010;363:45-53.

Risk adjustment is an essential element of comparative-effectiveness studies, measurements of health care performance, and payment programs and is destined to become even more important as health care reform proceeds. Observational studies comparing the outcomes of various approaches to treatment[1] or the performance of specific providers often adjust for patients' preexisting diagnoses.[2] The Medicare payment systems for institutional providers and health plans include payment adjustments that take into account the beneficiaries' health or functional status.[3,4] As payers move toward more bundled and value-based payment systems, incentives to avoid providing care for patients who are difficult to treat or patients for whom the cost of treatment is high will only increase.[5] Inadequate risk adjustment could thus lead to flawed inferences, the "dumping" of high-risk patients, and distortions in insurance markets.

Risk adjustment is only as good as the information on which it is based. Current risk-adjustment methods depend on the diagnoses that are recorded by physicians in medical records or registries or are coded by medical-records personnel and billing staff in hospital discharge abstracts and physician claims. Concern about the accuracy of the underlying data for risk adjustment has focused largely on the problem of upcoding — that is, recording conditions on submitted claims data in such a way that risk scores, and thus payments, are higher.[6] Differences in diagnostic practice have received much less attention. Studies have highlighted the ways in which the interpretation of pathological and radiologic examinations varies among physicians and the ways in which these differences affect the proportion of test results that are identified as abnormal.[7-10] There is also variation across practices[11] and regions,[12] among physicians treating patients with similar conditions, in the propensity to order diagnostic tests or refer patients to subspecialists. If physicians have substantial and systematic differences in their diagnostic practices that are unrelated to the underlying health of their patients but are related to institutional or regional practice patterns, biases in risk adjustment will result.

We conducted a study to determine the magnitude of the differences in diagnostic practices across U.S. regions, using changes in Medicare beneficiaries' place of residence as a natural experiment.

Methods

Design Overview

Previous research has documented substantial regional differences in the intensity of health care services provided to Medicare beneficiaries — differences that are independent of the beneficiaries' health or socioeconomic status.[13] Because beneficiaries are unlikely to be aware of the relative intensity of practice in their current region of residence or that of the region to which they move, we hypothesized that the average health status of beneficiaries who move to regions with a higher intensity of practice is similar to the average health status of beneficiaries from the same region who move to lower-intensity regions. We tested this hypothesis by comparing beneficiaries with respect to the number of diagnoses they had received and their use of health services before their move. We then followed the beneficiaries for 3 years after their move, stratifying them according to the intensity of practice in the region to which they moved. We hypothesized that those moving to higher-intensity regions would undergo more diagnostic tests and imaging services, would receive more diagnoses, and would thus have higher risk scores over time than those moving to lower-intensity regions.

Geographic Regions and the Definition of Intensity

We used a previously derived Medicare spending measure, the End-of-Life Expenditure Index, to define the local intensity of practice. This measure is calculated as the average spending (according to standardized national prices) on hospital and physician services provided to Medicare enrollees 65 years of age or older who were in their last 6 months of life, adjusted for age, sex, and race. The End-of-Life Expenditure Index reflects the component of local Medicare spending that is most closely associated with physician practice, rather than with local differences in the prevalence and severity of illnesses or in prices.[14] In previous articles, we have shown that the large differences that exist across U.S. regions in health care spending at the end of life are unrelated to differences in case mix[14] or patients' preferences regarding their care.[15] We calculated the End-of-Life Expenditure Index for each of the 306 U.S. Hospital Referral Regions (HRRs) for the period

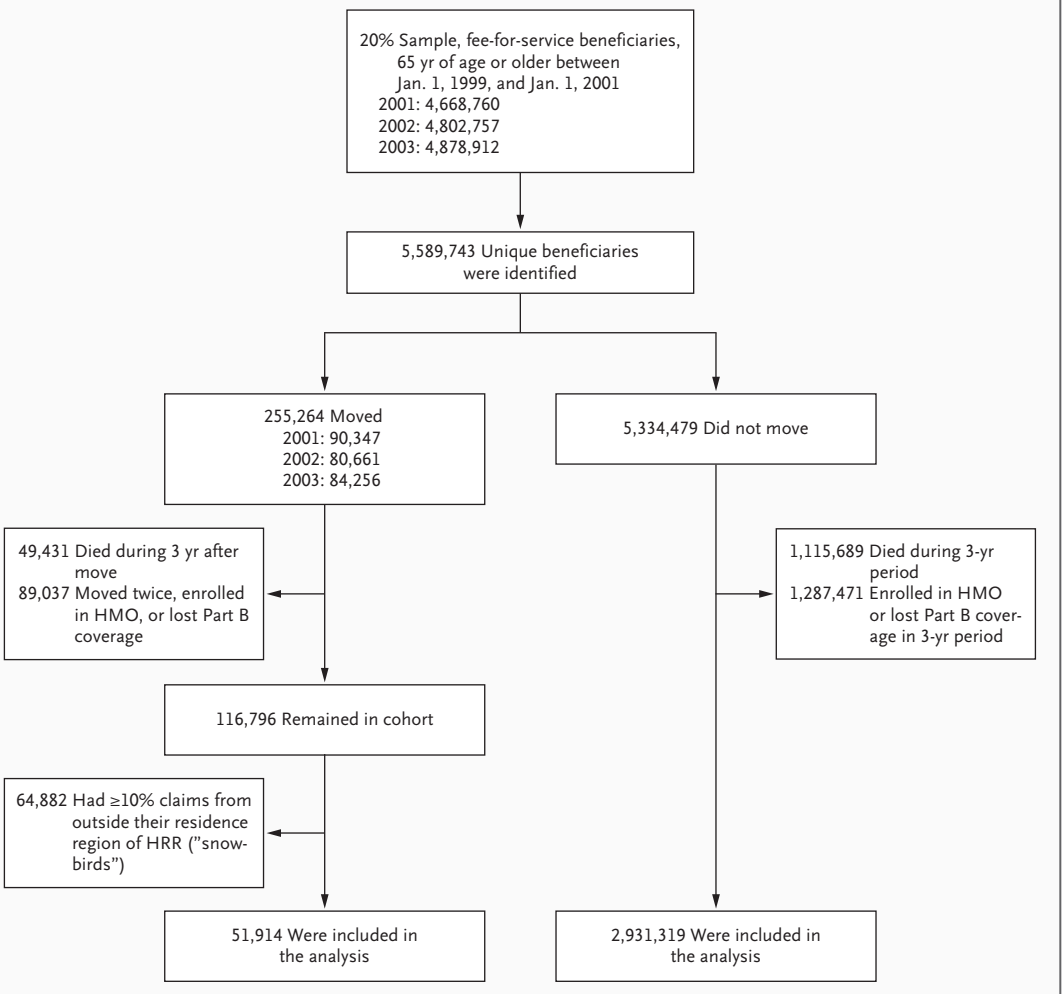

Figure 1. Selection of the Study Population.
HMO denotes health maintenance organization, and HRR Hospital Referral Region.

from 2001 through 2003 and grouped HRRs into quintiles of increasing intensity of practice.

STUDY POPULATION

Figure 1 shows the way in which we defined the study population. We used Medicare enrollment and claims files to identify all persons who were enrolled in Medicare Part A and Part B and who were at least 65 years of age as of January 1, 1999, resided in 1 of the 50 U.S. states or Washington, DC, and changed their place of residence in 2001, 2002, or 2003. A total of 255,264 beneficiaries were identified. To ensure that we had complete follow-up and that the exposure to the health system in each region was consistent over the entire 3-year period after the beneficiary moved, we excluded beneficiaries who died, who enrolled in a health maintenance organization (HMO) or lost their Part B coverage, or who moved more than once. We also excluded beneficiaries who received more than 10% of their claims from outside their place of residence ("snowbirds"); however, including these snowbirds in the analysis yielded similar results (see Tables 1 through 3 in the Supplementary Appendix, available with the full text of this article at NEJM.org).

Medicare beneficiaries were further stratified according to the quintile of the intensity of services provided in the region in which they originally resided and the quintile of the intensity of

services provided in the region to which they moved. This classification resulted in 25 subgroups, 1 for each possible combination of the quintiles before and after the move. For some analyses, we simply compared beneficiaries who moved to any HRR with a lower intensity of practice, those who moved to a region within their original quintile of intensity, and those who moved to an HRR in any higher-intensity quintile. As a comparison group, we also included those who did not move outside their HRR.

DIAGNOSTIC PRACTICES AND RISK SCORES

Our measures include both diagnostic practices (rates of diagnostic testing, imaging rates, and numbers of major chronic conditions) and Hierarchical Condition Categories (HCC) risk scores, which are currently used by Medicare for program payment. We measured rates of diagnostic testing and imaging services by first using the Berenson–Eggers Type of Service Codes (BETOS) to classify physician claims as laboratory tests or imaging services. We then counted the frequency of each type of claim for each beneficiary.

We counted the number of major chronic conditions that were documented in the Medicare physician and hospital claims data for each beneficiary during each of the 2 years before and the 3 years after a move. To reduce the likelihood of including "rule-out" diagnoses, we recorded a diagnosis as present if it was coded on an inpatient discharge abstract or on two physician claims submitted at least 7 days apart. We restricted this analysis to nine serious chronic conditions, on the basis of the work of Iezzoni and colleagues,[16] as adapted for use in the 2008 Dartmouth Atlas of Health Care (see Table 7 in the Supplementary Appendix). We calculated HCC risk scores with the use of the HCC coding algorithms that are used by the Centers for Medicare and Medicaid Services (CMS) to adjust payments for Medicare Advantage plans.[17]

STATISTICAL ANALYSIS

Unless otherwise specified, all the data presented in the tables are simple counts, means, and proportions. We estimated the average effect of a move to a region that was one quintile higher in intensity, using regression models in which the dependent variable was the beneficiary's final number of diagnoses or HCC risk score and the independent variables were age, race, sex, the original number of diagnoses or HCC risk score, and the change in the number of quintiles up or down they moved, which could range from −4 to +4. We carried out sensitivity analyses in which we included and excluded snowbirds and stratified the moves across all five levels of intensity. Results were similar in each case. Finally, to further determine potential differences in health status at the time of the move, we compared 1-year and 3-year rates of death among beneficiaries who moved to regions with a higher intensity or a lower intensity of practice. Confidence intervals were calculated with the use of the bootstrap method (Stata software).

RESULTS

STUDY POPULATION

Table 1 shows the baseline characteristics of all the Medicare beneficiaries who were eligible for inclusion and who had complete follow-up, stratified according to the intensity of services provided in their original HRR of residence and, if they moved, the intensity of services in the region to which they moved. Residents of higher-intensity regions generally had more office visits, underwent more diagnostic tests, had a higher number of diagnoses, and had higher risk scores than did residents in lower-intensity regions. The key analysis, however, involved determining whether there was a difference in these factors when persons who originally resided in a region with a given quintile of intensity moved to a region with a higher or lower quintile of intensity or to another region with the same quintile of intensity. With few exceptions, notably in quintile 5, among the beneficiaries who moved, the number of diagnoses, the risk scores, and the number of diagnostic tests and imaging services were similar among persons within each quintile during the period before their move.

DIAGNOSTIC PRACTICES AND HCC RISK SCORES

Table 2 shows the percent increases in laboratory tests, imaging services, and number of diagnoses among beneficiaries who did not move and among those who moved to a new HRR, stratified according to the intensity of services in the region to which they moved. Among both those who moved and those who did not move, there was a strong secular trend toward increased rates of diagnostic testing and a greater prevalence of

Table 1. Baseline Characteristics of the Study Population.*

Beneficiary	No. of Beneficiaries	Demographic Characteristics			Evaluative and Diagnostic Practices			Diagnoses	HCC Risk Score†
		Age	Female Sex	Black Race	Doctor Visits	Laboratory Tests	Imaging Services		
		yr (95% CI)	*% (95% CI)*		*no. (95% CI)*				*score (95% CI)*
Resident of quintile 1, region with lowest intensity of practice									
Did not change HRR of residence	600,490	75.5 (75.5–75.5)	59.4 (59.3–59.5)	1.5 (1.5–1.6)	9.7 (9.7–9.7)	7.3 (7.3–7.4)	2.9 (2.9–2.9)	0.51 (0.51–0.51)	0.95 (0.94–0.95)
Moved within quintile 1	7,984	76.1 (76.0–76.3)	64.3 (63.3–65.4)	0.5 (0.3–0.6)	10.8 (10.6–11.1)	7.5 (7.3–7.7)	3.1 (3.0–3.2)	0.59 (0.57–0.61)	1.03 (1.01–1.05)
Moved to higher-intensity region	3,811	75.5 (75.3–75.7)	64.0 (62.5–65.5)	3.0 (2.4–3.5)	10.8 (10.5–11.2)	7.5 (7.1–7.8)	3.1 (2.9–3.2)	0.55 (0.53–0.58)	0.97 (0.95–0.99)
Resident of quintile 2									
Did not change HRR of residence	583,865	75.4 (75.4–75.4)	61.0 (60.8–61.1)	6.3 (6.3–6.4)	11.0 (10.9–11.0)	8.5 (8.4–8.5)	3.2 (3.2–3.2)	0.59 (0.59–0.59)	0.99 (0.99–0.99)
Moved to lower-intensity region	1,589	76.1 (75.8–76.4)	67.1 (64.8–69.4)	2.1 (1.4–2.9)	12.4 (11.8–13.1)	8.6 (8.1–9.2)	3.5 (3.3–3.7)	0.64 (0.60–0.69)	1.05 (1.01–1.09)
Moved within quintile 2	2,302	76.3 (76.0–76.6)	67.0 (65.1–69.0)	4.0 (3.2–4.8)	12.4 (11.8–12.9)	8.3 (7.8–8.8)	3.6 (3.4–3.8)	0.68 (0.65–0.72)	1.08 (1.04–1.11)
Moved to higher-intensity region	3,206	75.7 (75.5–75.9)	67.1 (65.5–68.8)	7.3 (6.4–8.2)	11.9 (11.4–12.3)	8.5 (8.1–8.9)	3.4 (3.2–3.5)	0.60 (0.57–0.63)	1.00 (0.98–1.03)
Resident of quintile 3									
Did not change HRR of residence	593,265	75.3 (75.3–75.3)	61.5 (61.4–61.7)	8.0 (8.0–8.1)	11.5 (11.5–11.5)	9.0 (9.0–9.0)	3.4 (3.3–3.4)	0.61 (0.61–0.61)	1.00 (1.00–1.01)
Moved to lower-intensity region	2,205	76.0 (75.7–76.2)	67.9 (65.9–69.8)	4.5 (3.7–5.4)	12.5 (11.9–13.1)	8.8 (8.3–9.3)	3.5 (3.3–3.7)	0.64 (0.60–0.68)	1.04 (1.01–1.08)
Moved within quintile 3	2,862	75.9 (75.7–76.1)	67.3 (65.5–69.0)	4.9 (4.1–5.7)	13.4 (12.9–13.9)	9.3 (8.9–9.7)	3.9 (3.7–4.0)	0.74 (0.70–0.77)	1.11 (1.08–1.14)
Moved to higher-intensity region	2,448	75.6 (75.4–75.9)	66.6 (64.8–68.5)	7.5 (6.4–8.5)	11.8 (11.4–12.3)	8.3 (7.9–8.7)	3.3 (3.1–3.5)	0.61 (0.58–0.65)	1.01 (0.98–1.04)
Resident of quintile 4									
Did not change HRR of residence	624,235	75.3 (75.3–75.4)	61.0 (60.9–61.1)	8.3 (8.3–8.4)	12.3 (12.2–12.3)	9.7 (9.7–9.8)	3.6 (3.5–3.6)	0.65 (0.64–0.65)	1.02 (1.02–1.02)
Moved to lower-intensity region	4,509	76.1 (76.0–76.3)	66.0 (64.6–67.4)	4.9 (4.3–5.6)	13.4 (12.9–13.8)	10.2 (9.8–10.5)	3.8 (3.6–3.9)	0.68 (0.66–0.71)	1.07 (1.05–1.10)
Moved within quintile 4	5,188	75.6 (75.4–75.8)	64.1 (62.8–65.4)	4.3 (3.8–4.9)	13.8 (13.4–14.2)	10.2 (9.9–10.5)	4.0 (3.9–4.2)	0.72 (0.69–0.75)	1.09 (1.07–1.11)
Moved to higher-intensity region	1,211	76.7 (76.4–77.1)	67.1 (64.5–69.8)	7.9 (6.4–9.5)	15.0 (14.1–16.0)	10.8 (10.1–11.6)	4.2 (3.9–4.6)	0.74 (0.68–0.79)	1.16 (1.11–1.22)
Resident of quintile 5, region with highest intensity of practice									
Did not change HRR of residence	529,464	75.9 (75.9–75.9)	62.0 (61.8–62.1)	9.3 (9.2–9.3)	14.8 (14.8–14.9)	12.7 (12.6–12.7)	3.8 (3.8–3.8)	0.76 (0.76–0.77)	1.11 (1.11–1.11)
Moved to lower-intensity region	5,485	76.5 (76.3–76.7)	66.0 (64.7–67.2)	8.9 (8.1–9.7)	15.3 (14.9–15.8)	12.0 (11.7–12.4)	4.0 (3.9–4.2)	0.79 (0.76–0.81)	1.15 (1.13–1.18)
Moved within quintile 5	9,114	76.7 (76.5–76.8)	66.1 (65.1–67.1)	7.1 (6.6–7.6)	18.3 (17.9–18.7)	14.7 (14.4–15.1)	4.6 (4.5–4.7)	0.96 (0.93–0.98)	1.28 (1.26–1.30)

* The study population comprised Medicare beneficiaries, 65 years of age or older at the beginning of the study, who remained in a fee-for-service practice during the study period. Beneficiaries who moved were stratified according to whether their Hospital Referral Region (HRR) of residence changed between 2001 and 2003 and, if it did, whether the move was to a region with a lower intensity of practice, a region with a higher intensity of practice, or a region with the same intensity of practice as the region from which they had moved. Among beneficiaries who moved, the demographic characteristics and numbers of diagnoses and laboratory tests were based on information from the year before the move. Beneficiaries who did not change their HRR of residence were included in the analysis if they were at least 65 years of age between January 1, 1999, and January 1, 2001. All values, except for percentages, are means. HCC denotes Hierarchical Condition Category.

† HCC risk scores are derived from Medicare's enrollment and claims files with the use of demographic and diagnostic data. Higher scores indicate greater risk. In the 20% sample of beneficiaries in 2004, the scores ranged from 0.30 to 15.26.

Table 2. Percent Increase in the Number of Major Diagnoses, Diagnostic Tests, and Imaging Services among Medicare Beneficiaries, According to the Intensity of Practice in the Hospital Referral Region (HRR) in Which They Lived.

Beneficiary	No. of Beneficiaries	Increase in Diagnoses per Person	Increase in Imaging Services per Person	Increase in Laboratory Tests per Person
		percent (95 percent confidence interval)		
Resident of quintile 1, region with lowest intensity of practice				
Did not change HRR of residence	600,490	65.6 (64.9 to 66.3)	43.9 (43.2 to 44.5)	31.6 (31.1 to 32.1)
Moved within quintile 1	7,984	61.7 (55.8 to 67.4)	24.1 (19.1 to 29.0)	23.7 (19.1 to 28.2)
Moved to higher-intensity region	3,811	100.8 (89.6 to 112.1)	66.1 (57.5 to 74.7)	74.2 (64.7 to 83.6)
Resident of quintile 2				
Did not change HRR of residence	583,865	63.9 (63.3 to 64.5)	49.6 (48.9 to 50.3)	37.4 (36.9 to 37.9)
Moved to lower-intensity region	1,589	52.7 (40.5 to 64.9)	26.4 (15.5 to 37.4)	6.0 (–2.9 to 14.9)
Moved within quintile 2	2,302	62.6 (53.4 to 72.6)	26.6 (17.8 to 35.3)	36.4 (27.1 to 45.7)
Moved to higher-intensity region	3,206	91.3 (80.3 to 102.4)	49.7 (41.9 to 57.5)	59.0 (50.1 to 67.9)
Resident of quintile 3				
Did not change HRR of residence	593,265	62.6 (62.0 to 63.2)	47.8 (47.1 to 48.5)	36.0 (35.5 to 36.4)
Moved to lower-intensity region	2,205	65.4 (54.1 to 76.7)	35.3 (25.0 to 45.5)	33.5 (24.9 to 42.1)
Moved within quintile 3	2,862	60.0 (50.4 to 68.7)	28.7 (20.9 to 36.6)	36.7 (29.3 to 44.1)
Moved to higher-intensity region	2,448	84.6 (72.6 to 96.7)	62.0 (50.8 to 73.2)	76.3 (64.5 to 88.2)
Resident of quintile 4				
Did not change HRR of residence	624,235	62.5 (61.9 to 63.1)	49.7 (49.1 to 50.3)	37.6 (37.1 to 38.0)
Moved to lower-intensity region	4,509	61.8 (54.4 to 69.2)	25.9 (20.1 to 31.6)	15.5 (10.3 to 20.7)
Moved within quintile 4	5,188	68.4 (62.4 to 74.4)	33.1 (27.2 to 39.1)	40.2 (34.7 to 45.6)
Moved to higher-intensity region	1,211	84.3 (70.0 to 98.6)	50.2 (35.8 to 64.5)	48.8 (37.4 to 60.1)
Resident of quintile 5, region with highest intensity of practice				
Did not change HRR of residence	529,464	54.8 (54.3 to 55.3)	50.4 (49.7 to 51.1)	34.1 (33.6 to 34.6)
Moved to lower-intensity region	5,485	42.6 (37.0 to 48.3)	16.6 (11.4 to 21.9)	6.6 (2.2 to 11.0)
Moved within quintile 5	9,114	58.5 (54.4 to 62.5)	38.6 (33.9 to 43.4)	32.8 (29.2 to 36.4)

diagnosed conditions. For example, among persons in quintile 2 who did not move, the number of diagnoses rose by 63.9% (95% confidence interval, [CI], 63.3 to 64.5); among those who moved to the less intensive quintile 1, the number of diagnoses increased by 52.7% (95% CI, 40.5 to 64.9), whereas among those who moved to a more intensive region (quintiles 3 through 5), the number of diagnoses increased by 91.3% (95% CI, 80.3 to 102.4). The results were similar for imaging and laboratory tests.

The percent increase in HCC risk scores, stratified according to the original quintile of intensity and the quintile of intensity of the region to which the beneficiary moved, is shown in Figure 2. Moving to a region with a higher intensity of practice was consistently associated with a greater increase in risk scores. Similar patterns were observed in each of the five quintiles.

We also estimated the average effect on the number of diagnoses of moving to a region that was one quintile of intensity higher. Receiving care in a quintile of intensity that was one step higher was associated with a 5.9% increase in the HCC risk score (95% CI, 5.2 to 6.7). We restricted the analysis to beneficiaries who moved within each of the nine major U.S. census regions and found the same pattern of greater increases in the number of diagnoses and in risk scores among those who moved to higher-intensity regions (see Tables 8 through 10 in the Supplementary Appendix). Finally, we compared 1-year and 3-year rates of death after the move, adjusting for age, sex, and race. The relative risk

of death at 1 year among beneficiaries who moved to a region with a higher intensity of practice and among those who moved to a region with a lower intensity of practice, was identical: 0.94 (95% CI, 0.90 to 0.98). At 3 years, there was no evidence of a survival benefit among those who moved to higher-intensity regions as compared with those who moved to lower-intensity regions (Table 11 in the Supplementary Appendix).

Beneficiaries who moved to quintile 5 regions had risk scores that were, on average, 19% higher than those of beneficiaries who moved to quintile 1 regions (see Table 6 in the Supplementary Appendix). Adjusting for such an increase in HCC scores would reduce the 1-year rate of death by 15%, as compared with the unadjusted rate of death.

DISCUSSION

Among all Medicare beneficiaries, residence in regions of the United States that have a higher intensity of services is associated with a higher reported prevalence of common chronic illnesses. Whether this is due to a higher disease burden or to differences in diagnostic practices in high-intensity regions (or both) was unknown. To address this question, we followed Medicare beneficiaries for 2 years before and 3 years after a move and found that a move to a region with a higher intensity of practice as compared with a move to a region with a lower intensity of practice was associated with greater increases in diagnostic testing, the number of recorded chronic conditions, and HCC risk scores, with no apparent survival benefit.

This study extends previous research on variations in diagnostic practices. Earlier studies have documented variations in the use of diagnostic tests among individual physicians and office practices[11] and variations across regions in physicians' reports of their likelihood of ordering tests.[12] Higher rates of diagnostic testing lead to an increased number of diagnoses of specific clinical conditions, such as prostate cancer, thyroid disease, or vascular disease.[18-20] The biases introduced by greater diagnostic intensity have been well documented in the context of cancer.[21,22] Our study builds on this earlier work.

This study has several limitations. We cannot be certain that the beneficiaries who were re-

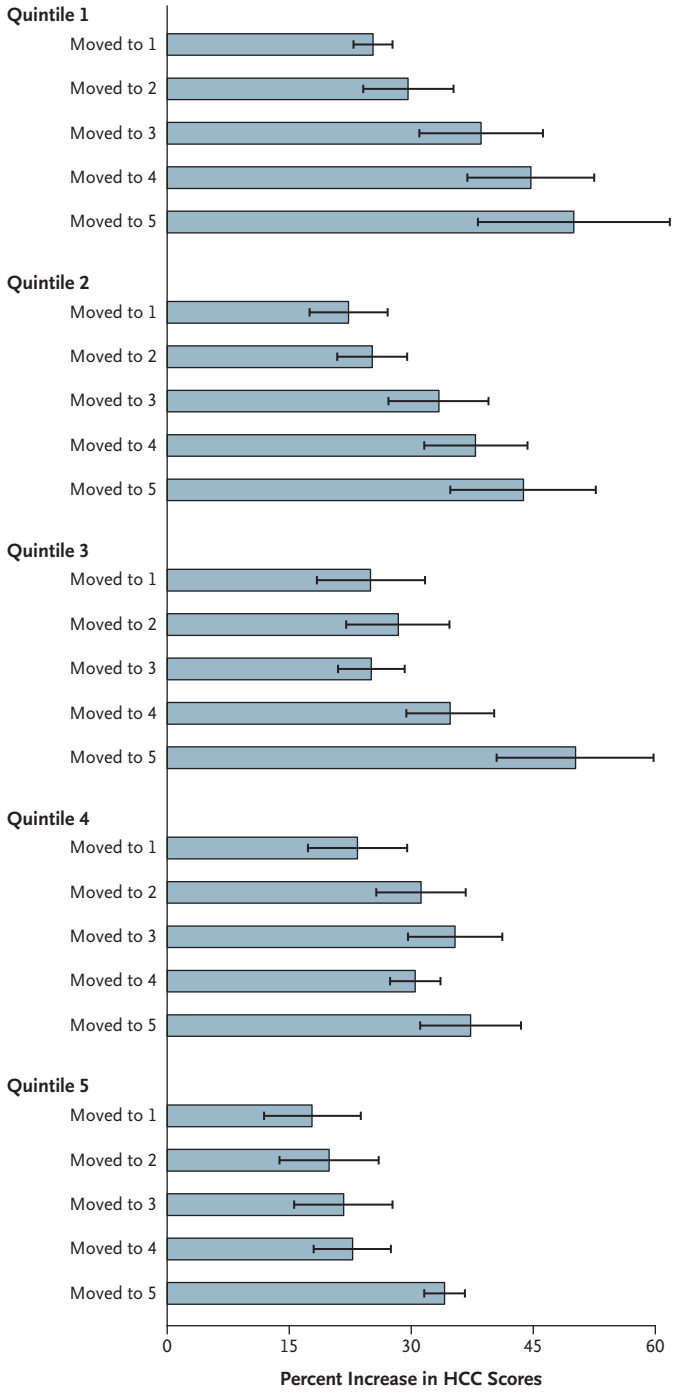

Figure 2. Percent Increase in Hierarchical Condition Category (HCC) Scores among Medicare Beneficiaries Who Changed Their Place of Residence.

Higher quintiles indicate a higher intensity of practice. The results were stratified according to the intensity of practice in the Hospital Referral Region in which the beneficiary lived before the move and the one in which he or she lived after the move. Horizontal I bars indicate 95% confidence intervals.

corded as having received a new diagnosis actually had the disease. Given the increased rates of imaging and laboratory testing, however, at least some of the additional diagnoses are likely to have indicated newly detected conditions. We also cannot rule out the possibility that there were differences in the beneficiaries' underlying health status at the time of the move. However, the similarities in the number of diagnoses and in rates of testing before the move, as well as in the relative risk of death 1 year after the move, are consistent with the view that the underlying health status of the beneficiaries was similar for those who moved to regions with more intensive practice and those who moved to regions with less intensive practice.

Our study has not adequately assessed the effect of the regional differences in diagnostic intensity on the health outcomes for beneficiaries. Several previous studies involving specific clinical cohorts have shown no evidence of a survival benefit when care is provided in regions[13,23] or hospitals[24] with a higher intensity of services. One study showed that there was a small benefit with respect to survival when the intensity of life-sustaining treatments was greater,[25] and a study involving six hospitals showed that among patients with heart failure, the rate of death was lower in the hospitals that used more resources in caring for patients than in those that used less.[2] The analyses in all these studies adjusted the data for patients' diagnoses as coded from provider data and thus risked overadjustment in higher-intensity regions and hospitals (and conversely, underadjustment in lower-intensity regions or hospitals). Although our study did not show a significantly higher rate of survival among beneficiaries who moved to regions with higher-intensity practices, this result should not be interpreted as implying that greater diagnostic intensity offers no benefits. Rather, it underscores the need for research to determine the specific clinical settings in which greater diagnostic intensity does — or does not — confer a benefit.

Our findings nonetheless have implications for health care reform. Comparative-effectiveness studies could be biased by the well-documented differences in diagnostic intensity across hospitals.[26] Under public reporting programs, patients may be subject to harm to the extent that their own choices or their physicians' referrals are based on biased risk-adjusted quality measures. Capitation systems and bundled payments for episodes of care could also be distorted. The differences are not likely to be trivial, on the basis of our analyses. By the end of the study, beneficiaries who had moved to quintile 5 regions (those with the highest intensity of practice) had risk scores that were, on average, 19% higher than those of beneficiaries who had moved to quintile 1 regions (those with the lowest intensity of practice). Since patients had similar baseline health status, these differences are plausible estimates of the differences in diagnostic intensity across these regions. Under a public reporting or payment program that relied on the unmodified HCC scores, capitated reimbursement rates would be as much as 19% higher in the high-intensity regions solely because of bias related to diagnostic practice, particularly since the CMS has been relying to a greater extent on HCC scores in adjusting payments to Medicare Advantage plans.[27] In addition, our results suggest that when HCC scores are overestimated by 19%, risk-adjusted rates of death would appear to be 15% lower.

We recognize that biases related to diagnostic intensity are not the only challenge confronting risk adjustment. A major concern about both payment reforms and performance-measurement initiatives is their potential for adversely affecting behavior. For example, if providers are more highly compensated for treating patients with more diagnoses, they could conceivably be inclined to perform more intensive screening and diagnostic testing, with clear effects on costs and uncertain effects on health outcomes. Alternatively, risk-adjustment models could fail to account for the difficulty of caring for truly high-risk patients or those whose care is made more difficult owing to challenges such as language barriers, poor health literacy, or lack of social support, encouraging some providers to avoid or stop providing care for such patients. Such concerns only underscore the importance of continued efforts to advance the development of unbiased methods of risk adjustment as health care reform proceeds.

These challenges could become more manageable as comprehensive electronic health records are implemented. To help improve risk adjustment, such systems would need to incorporate both nonclinical factors that may predict a patient's lack of adherence to clinical advice (e.g.,

homelessness or poverty) and clinical data that are less subject to bias that is due to differences in diagnostic practices. Examples of such data include stage and grade in the case of patients with cancer and ejection fraction in the case of those with congestive heart failure. It is also possible that measures of health risks reported by patients (e.g., smoking and exercise patterns) and functional status (physical, social, and role function) could be incorporated in risk-adjustment models to improve their performance.

The newly passed health care reform legislation includes substantial increases in funding for comparative-effectiveness research programs and establishes major initiatives that will move Medicare and Medicaid toward bundled payment systems. Our findings underscore the need for additional efforts to advance risk-adjustment methods as reform proceeds.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003;349:2117-27.
2. Ong MK, Mangione CM, Romano PS, et al. Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure. Circ Cardiovasc Qual Outcomes 2009;2:548-57.
3. Medicare payment basics: hospital acute inpatient services payment system. Washington, DC: Medicare Payment Advisory Commission, 2008.
4. Skilled nursing facilities prospective payment system: overview. Baltimore: Centers for Medicare and Medicaid Services, 2009. (Accessed May 7, 2010, at http://www.cms.hhs.gov/SNFPPS/01_overview.asp.)
5. Fisher ES. Paying for performance — risks and recommendations. N Engl J Med 2006;355:1845-7.
6. Silverman E, Skinner J. Medicare upcoding and hospital ownership. J Health Econ 2004;23:369-89.
7. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. N Engl J Med 1994;331:1493-9.
8. Schnitt SJ, Connolly JL, Tavassoli FA, et al. Interobserver reproducibility in the diagnosis of ductal proliferative breast lesions using standardized criteria. Am J Surg Pathol 1992;16:1133-43.
9. Black WC, Welch HG. Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. N Engl J Med 1993;328:1237-43.
10. Urban BA, Fishman EK, Kuhlman JE, Kawashima A, Hennessey JG, Siegelman SS. Detection of focal hepatic lesions with spiral CT: comparison of 4- and 8-mm interscan spacing. AJR Am J Roentgenol 1993;160:783-5.
11. Stafford RS, Misra B. Variation in routine electrocardiogram use in academic primary care practice. Arch Intern Med 2001;161:2351-5.
12. Sirovich B, Gallagher PM, Wennberg DE, Fisher ES. Discretionary decision making by primary care physicians and the cost of U.S. health care. Health Aff (Millwood) 2008;27:813-23.
13. 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;138:288-98.
14. *Idem*. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med 2003;138:273-87.
15. Barnato AE, Herndon MB, Anthony DL, et al. Are regional variations in end-of-life care intensity explained by patient preferences? A study of the US Medicare population. Med Care 2007;45:386-93.
16. Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA. Chronic conditions and risk of in-hospital death. Health Serv Res 1994;29:435-60.
17. Pope GC Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financ Rev 2004;25:119-41.
18. Ezzat S, Sarti DA, Cain DR, Braunstein GD. Thyroid incidentalomas: prevalence by palpation and ultrasonography. Arch Intern Med 1994;154:1838-40.
19. Lederle FA, Walker JM, Reinke DB. Selective screening for abdominal aortic aneurysms with physical examination and ultrasound. Arch Intern Med 1988;148:1753-6.
20. Lu-Yao G, Albertsen PC, Stanford JL, Stukel TA, Walker-Corkery ES, Barry MJ. Natural experiment examining impact of aggressive screening and treatment on prostate cancer mortality in two fixed cohorts from Seattle area and Connecticut. BMJ 2002;325:740.
21. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312:1604-8.
22. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? JAMA 2000;283:2975-8.
23. Skinner JS, Staiger DO, Fisher ES. Is technological change in medicine always worth it? The case of acute myocardial infarction. Health Aff (Millwood) 2006;25(2):w34-w47.
24. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ. Variations in the longitudinal efficiency of academic medical centers. Health Aff (Millwood) 2004;Suppl Web Exclusives:VAR19-VAR32.
25. Barnato AE, Chang CC, Farrell MH, Lave JR, Roberts MS, Angus DC. Is survival better at hospitals with higher "end-of-life" treatment intensity? Med Care 2010;48:125-32.
26. Wennberg JE, Fisher ES, Goodman DC, Skinner J. Tracking the care of patients with severe chronic illness: the Dartmouth atlas of health care 2008. Hanover, NH: The Dartmouth Atlas Project, 2008.
27. 2006 Risk adjustment data basic training for Medicare Advantage organizations: participant guide. Baltimore: Centers for Medicare and Medicaid Services, 2006.