Association Between Maintenance of Certification Examination Scores and Quality of Care for Medicare Beneficiaries

Eric S. Holmboe, MD; Yun Wang, PhD; Thomas P. Meehan, MD, MPH; Janet P. Tate, MPH; Shih-Yieh Ho, PhD, MPH; Katie S. Starkey, MHA; Rebecca S. Lipner, PhD

Background: The relationship between physicians' cognitive skill and the delivery of evidence-based processes of care is not well characterized.Therefore, we set out to determine associations between general internists' performance on the American Board of Internal Medicine maintenance of certification examination and the receipt of important processes of care by Medicare patients.

Methods: Physicians were grouped into quartiles based on their performance on the American Board of Internal Medicine examination. Hierarchical generalized linear models examined associations between examination scores and the receipt of processes of care by Medicare patients. The main outcome measures were the associations between diabetes care, using a composite measure of hemoglobin A_{1c}, and lipid testing and retinal screening, mammography, and lipid testing in patients with cardiovascular disease and the physician's performance on the American Board of Internal Medicine examination, adjusted for the number of Medicare patients with diabetes and cardiovascular disease in a physician's practice panel; frequency of visits; patient comorbidity, age, and ethnicity; and physician training history and type of practice.

Results: Physicians scoring in the top quartile were more likely to perform processes of care for diabetes (composite measure odds ratio [OR], 1.17; 95% confidence interval [CI], 1.07-1.27) and mammography screening (OR, 1.14; 95% CI, 1.08-1.21) than physicians in the lowest physician quartile, even after adjustment for multiple factors. There was no significant difference among the groups in lipid testing of patients with cardiovascular disease (OR, 1.00; 95% CI, 0.91-1.10).

Conclusion: Our findings suggest that physician cognitive skills, as measured by a maintenance of certification examination, are associated with higher rates of processes of care for Medicare patients.

Arch Intern Med. 2008;168(13):1396-1403

THE OVERALL QUALITY OF care for Medicare patients is suboptimal.[1,2] Research has not clearly identified which physician factors affect the quality of care for older patients. A systematic review found a negative relationship between physician experience and performance.[3] Several recent studies have found a positive association between performance on some process quality measures and whether a physician was board certified.[4-7] However, the attainment of board certification is a relatively blunt assessment tool, as the majority of US-trained physicians ultimately achieve board certification.[8] One outstanding question is whether actual cognitive skill, as measured by a reliable high-stakes examination, is associated with quality of care as assessed by validated performance measures. Public support for repeated cognitive testing has been strong, with 87% of patients who were surveyed saying that they believe a physician should take an examination of knowledge periodically.[9,10]

For editorial comment see page 1365

However, some in the physician community have questioned the need for a recertification examination, citing the lack of evidence that maintenance of certification (MOC) provides meaningful information about physician competence.[11] One Canadian study did find that family physicians' scores on an initial licensing examination had predictive validity for future performance on a number of quality measures up to 6 years after the examination.[12] However, to our knowledge, no study has yet examined the concurrent relationship between cognitive skills and quality of care.

Author Affiliations: American Board of Internal Medicine, Philadelphia, Pennsylvania (Drs Holmboe and Lipner and Ms Starkey), and Qualidigm, Middletown, Connecticut (Drs Wang, Meehan, and Ho and Ms Tate).

The primary objective of this study was to examine the association between cognitive skills, as measured by performance on a general internal medicine MOC examination, and performance on a set of Center for Medicare and Medicaid Services quality measures for patients with diabetes and mammography screening, as well as lipid testing for cardiovascular disease. Our hypothesis is that internists with higher cognitive skills will provide better quality of care, as assessed by claims-based processes of care measures: hemoglobin A_{1c} testing, lipid testing, and retinal eye examinations for patients with diabetes, mammography screening for women, and lipid testing for patients with cardiovascular disease.

METHODS

PHYSICIANS AND PATIENTS

The sampling method is outlined in the **Figure**. Potential subjects were internists who initially certified in internal medicine between 1990 and 1995, had a universal physician identification number, and did not hold a subspecialty certificate. We then excluded physicians who did not have between 30 and 800 Medicare beneficiaries in their practice. First, because the response variable at the patient level follows a binomial distribution, we selected a lower limit of 30 cases to ensure that the response variable approximated to a normal distribution. Second, we selected an upper limit of 800 to avoid a situation in which physicians may share 1 universal physician identification number across physicians or groups; 800 beneficiaries represented the 95th percentile for Medicare patient volume. Third, the Center for Medicare and Medicaid Services limits the number of patient records it will release, and we wanted to ensure that we only sampled actively practicing physicians. Finally, we excluded physicians who attempted their first initial certification examination before 1985 or their first MOC examination before 2000, did not take all 3 internal medicine modules on the MOC examination (60 questions per module), or canceled their MOC enrollment. From this cohort, we generated a random sample of 5000 physicians to assemble the patient cohorts for analysis.

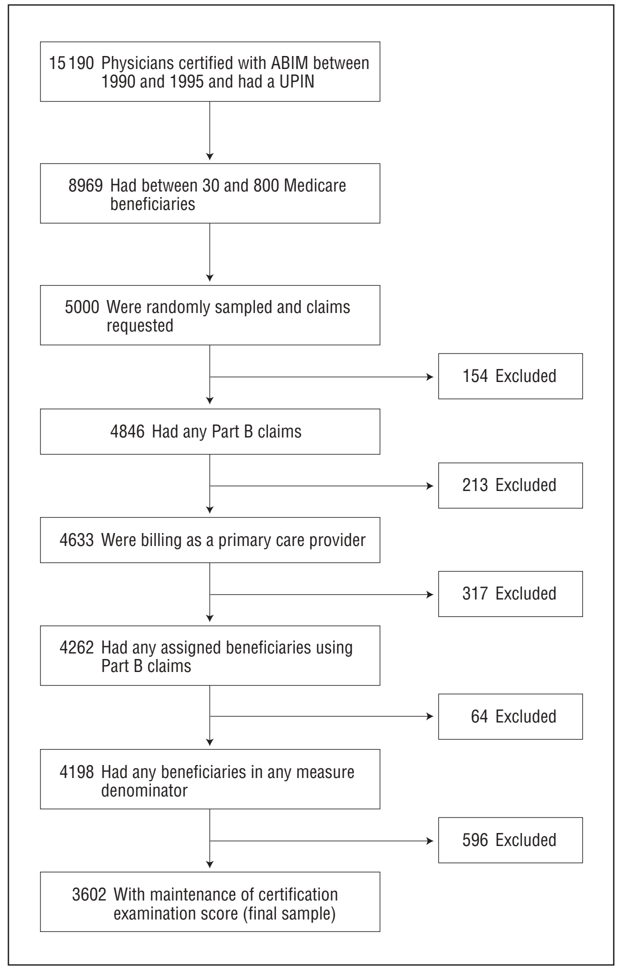

Figure. Assembly of physician sample. ABIM indicates American Board of Internal Medicine; PCP, primary care physician; UPIN, universal physician identification number.

Medicare claims data and enrollment files from 2002 and 2003 were used to identify patients seen by the physician cohort, to determine the performance measures for which they were eligible and whether they received them, and to ascertain relevant patient characteristics and comorbidities. All 2002 and 2003 Medicare claims were obtained for the beneficiaries to ensure that services were counted regardless of who provided them and to ensure proper assignment to the appropriate primary care physician. Assignment of beneficiaries to physicians was done using only Part B office visit claims for medical care. Part A outpatient claims were not used for physician assignment in order to keep the study population more homogeneous and because the unique physician-beneficiary linkage is less clear. Since physicians sometimes bill under more than 1 specialty, a provider was considered a primary care physician if the majority of his or her Part B claims were billed as internal medicine, general practice, family practice, or geriatrics. Beneficiaries were assigned to the physician whom they saw most often (or most recently in case of ties) in each calendar year.[13] Those beneficiaries who were assigned to the same physician, from the cohort described above, in both 2002 and 2003 were included in the study sample. Therefore, a patient had to have a minimum of 2 visits to the physician: 1 in 2002 and 1 in 2003. This requirement eliminated patients who switched physicians late in 2002 or who were new patients in 2003. Visits to specialists were not used for assignment. Eligible Medicare beneficiaries had 24 months of fee-for-service Part B coverage in 2002 and 2003 and were alive as of December 31, 2003. We then restricted the physician cohort only to those physicians who had actually taken the MOC examination between 2000 and 2005. This resulted in a final study sample of 3602 physicians and a total of 220 340 beneficiaries eligible for 1 or more of the performance measures. The study was approved by the New England Institutional Review Board.

MOC EXAMINATION

Recognizing the need for the periodic assessment of practicing physicians, the American Board of Internal Medicine (ABIM) instituted time-limited internal medicine certificates in 1990.[10-14] All internists must now complete the MOC program every 10 years to receive a new certificate. To renew a certificate, physicians must complete a set of formative self-assessment activities and pass a rigorous recertification examination. The ABIM MOC examination consists of 180 single-best-answer questions. The examination evaluates the extent of the candidate's knowledge and clinical judgment (ie, cognitive skills) in areas in which an internist should demonstrate competence, including treatment of both common and uncommon conditions that have important consequences for patient care (www.abim.org). The examination also includes questions about diagnosis, treatment, and prevention.

QUALITY-OF-CARE MEASURES

The quality-of-care measures that were used in this study were receipt of at least 2 hemoglobin A_{1c} measurements in the previous year and a lipid test and retinal eye examination within the past year for patients with diabetes; mammography for women aged 65 through 74 years within the previous year; and lipid testing within the past year for patients with cardiac disease. Receipt of each service was determined using claims-based *Current Procedural Terminology* and *International Classification of Diseases, Ninth Revision, Clinical Modification*, codes. For diabetes, we used the perfect score method by combining all 3 processes of care for each patient to create a composite measure, assigning a value of 1 if all 3 measures had been performed and 0 if 2 or fewer measures had been performed.[13]

MOC SCORE

The candidates were divided into 3 percentile groups (<25th, 25-75th, and >75th) based on their first attempt on their ABIM MOC examination. We used the interquartile approach because the relationship between physicians' examination score and their quality of performance is not linear. Second, the cutpoints of less than 25th, 25th through 75th, and greater than 75th percentiles are a standard method to measure scorelike variables[15,16] and are used widely in practice.[17,18] The equated passing score for the MOC examinations was approximately 350.

STATISTICAL ANALYSIS

Bivariate analysis was performed to assess associations between patient characteristics and MOC scores. Patient characteristics were divided into 3 domains: demographics, comorbidities, and access to care during the previous 12 months. Demographics included age, sex (female vs male), race (white vs other), Medicaid status, and comorbidities, including history of percutaneous transluminal coronary angioplasty, coronary artery bypass graft, heart attack, heart failure, renal failure, depression, peripheral vascular disease, hypertension, chronic obstructive pulmonary disease, cancer, liver disease, dementia, diabetes, smoking, and obesity status. We also calculated a Deyo score for each patient.[19] The access-to-care domain includes the total number of primary care physician visits; any cardiologist, endocrinologist, or nephrologist visits (yes/no); hospital emergency department visits (yes/no); or hospitalizations (yes/no). The Mantel-Haenszel χ^2 test was used to compare dichotomous and categorical variables, and the Kruskal-Wallis rank test was used to compare continuous variables.

We used the hierarchical generalized linear model to evaluate associations between physician MOC scores and physician performance on quality measures. Because patients seeing the same physician are not independent, it is necessary to use a multilevel model to properly account for the clustering of patients by physician. The log-odds of the quality measures were modeled as a function of the MOC score represented by 2 dummy variables, with the lowest-score category as the reference.[20-22] Our intent was not to develop a predictive model for each measure; rather, it was to focus on associations between the MOC score and quality measures with and without adjustment for patient and physician characteristics. Physician characteristics included sex (female vs male), years since graduation, country of medical school (United States vs other), type of practice, and condition-specific patient volume. Patient characteristics were included to differentiate the effect of patient behavior from physician behavior.

Accordingly, we sequentially constructed 3 models for each measure. The first model was fitted without adjustment for either patient or physician characteristics; the second model was fitted with adjustment for patient demographics, comorbidity, and access to care; and the third model was fitted with adjustment for patient and physician characteristics. All hierarchical generalized linear models were fitted with a random physician-specific effect to account for within-physician correlation of the measures. The hierarchical generalized linear models separated the within-physician variation from the between-physician variation. We calculated 95% confidence intervals for each estimate from the models. All statistical analyses were conducted using hierarchical linear modeling version 5.0 (Scientific Software International Inc, Lincolnwood, Illinois) and SAS version 9.02 (SAS Institute Inc, Cary, North Carolina).

RESULTS

SUBJECTS

Between 2002 and 2003, the study cohort of 3602 general internists who took a recertification examination cared for more than 220 000 Medicare patients with diabetes, patients with cardiovascular disease, and women who were eligible for mammography screening. **Table 1** displays the physician characteristics and bivariate results. Of note, approximately 37% of the cohort were women, 26% had trained in a foreign medical school, and nearly 23% failed the initial certification examination on their first attempt. Approximately 60% of these physicians worked in practices of 10 or fewer physicians, with nearly 20% working as solo practitioners.

BIVARIATE RESULTS

Table 1 shows that physicians who were not graduates of a US or Canadian medical school, who had graduated more than 20 years ago from medical school, and who were solo practitioners were more likely to score in the lowest quartile on the MOC examination. **Table 2** highlights important patient variables and receipt of processes of care. Of note, patients with diabetes who were male, African American, or Hispanic; who received fewer visits; or who did not see a specialist were less likely to receive "perfect care" (eg, all 3 diabetes process-of-care measures). The same relationships held true for mammography (sex not applicable) and lipid testing in patients with cardiovascular disease. The relationship with the Deyo comorbidity score varied by condition: patients with diabetes and a Deyo score of 3 or higher were more likely to receive all 3 processes of care, whereas patients with cardiovascular disease were modestly less likely to receive a lipid test.

Table 3 provides the bivariate analyses of physician characteristics and performance on the quality measures. Regarding the diabetes and mammography measures, physicians who were not graduates of a US or Canadian medical school, who had graduated more than 20 years ago from medical school, who were male, and who were solo practitioners did worse on the performance measures for diabetes and mammography. For example, those physicians in the lowest quartile had a compliance rate 6.2% lower than physicians in the top quartile for mammography screening and 5.0% lower for the diabetes composite measure.

Table 1. Physician Characteristics and Maintenance of Certification (MOC) Score Percentiles

Physician Characteristics	MOC Score Percentile <25th (<435[a])	50th (435-580[a])	>75th (>580[a])	Total	P Value
No. of physicians	862	1853	887	**3602**	
US or Canadian medical school, %	57.2	75.9	87.6	74.3	<.001
Male, %	63.8	63.7	61.7	63.2	.72
Years since medical school graduation as of 2003, %					<.001
11-20	75.4	82.9	89.0	82.6	
21-44	18.0	9.1	4.5	10.1	
Missing data	6.6	8.0	6.5	7.3	
No. of Medicare beneficiaries cared for, %					.01
<50	33.9	32.1	36.4	33.6	
50-150	39.1	35.3	33.7	35.8	
>150	27.0	32.6	29.9	30.6	
Practice type, %					<.001
Solo	32.0	18.6	10.7	19.9	
Group, 2-10 physicians	37.5	42.3	36.8	39.8	
Group, ≥11 physicians	15.3	22.1	25.7	21.3	
Academic hospital–based	3.8	7.9	16.3	9.0	
Missing data/other	11.4	9.1	10.5	10.0	
Residency program director's overall clinical competence rating, %					<.001
4	4.6	3.3	0.5	2.9	
5	25.3	14.1	5.6	14.7	
6	32.6	29.1	16.2	26.8	
7	26.3	30.5	32.0	29.9	
8	7.5	17.2	32.9	18.7	
9	1.4	4.2	11.8	5.4	
Missing data	2.2	1.7	0.9	1.6	
No. of tries to achieve initial certification in internal medicine, %					<.001
1	54.5	77.9	97.3	77.1	
2	23.3	15.2	2.3	13.9	
3	12.6	4.9	0.1	5.6	
4	6.4	1.5	0.2	2.3	
5-8	3.1	0.6	0.1	1.1	
Initial internal medicine certification examination score, %					<.001
<300	10.4	3.1	0.2	4.1	
300-399	30.2	14.8	1.6	15.2	
400-499	50.0	45.4	17.2	39.6	
500-599	9.2	34.5	59.8	34.7	
≥600	0.2	2.2	21.2	6.4	

[a] Indicates equated score on MOC examination.

Table 3 also provides an additional analysis that breaks down the MOC score into 5 levels. Physicians who scored below an equated score of 300 (with a passing score of approximately 350) had a compliance rate that was 8.6% lower for mammography screening and 9.3% for the diabetes composite measure compared with physicians with an equated score greater than 600.

RESULTS OF HIERARCHICAL MODELS

Table 4 shows the results of the hierarchical models. After adjustment for patient and physician characteristics, there was a significant association for performance on the diabetes and mammography quality-of-care process measures and examination scores. By rough estimation, patients with diabetes were 17% more likely to receive all 3 diabetes processes of care and women were 14% more likely to undergo mammography if they were cared for by physicians who scored in the top quartile compared with the lowest quartile group. The 1 measure that did not show a significant association with MOC scores was lipid testing for patients with peripheral vascular and cardiovascular disease. However, more than 59% of all patients with this condition saw a cardiologist subspecialist, compared with only 6% of the patients with diabetes who saw an endocrinologist. Therefore, performance on this measure could have been confounded by cardiologist performance.

COMMENT

To our knowledge, this is one of the first studies to examine the relationship between physician cognitive skills,

Table 2. Performance Measure Rates by Beneficiary Characteristics

Beneficiary Characteristics	Diabetes Measures: No. of Patients With Diabetes	% of Patients Receiving Service: Eye Examination	Lipid test	HbA_{1c} test	All 3 Services	*P* Value	Mammography: No. of Female Patients	With Mammogram, %	*P* Value	HEDIS CAD Measure: No. of Patients With CAD	With Lipid Test, %	*P* Value
Demographics												
Age, y						<.001						<.001
<65	11 582	46.2	70.1	82.0	34.0		15 327	46.6	<.001	7204	70.0	
≥65	40725	61.8	77.8	86.8	49.4		118 404	58.0		42 322	77.8	
Sex						<.001						<.001
Male	24 360	55.3	76.2	85.0	43.8					27 934	77.7	
Female	27 947	61.0	76.0	86.5	47.9					21 592	75.4	
Race						<.001						<.001
White	41 167	59.2	77.0	86.2	47.2		115 628	57.9	<.001	43626	77.3	
Black	8230	54.5	71.4	84.6	40.4		13 137	50.8		4199	70.5	
Hispanic	962	55.0	73.9	82.7	42.3		1330	42.8		517	72.5	
Asian	894	60.0	79.9	83.3	47.5		1798	40.1		563	79.9	
Other	1054	56.6	73.4	81.0	43.5		1838	48.7		621	73.6	
Access to care during 2002 and 2003												
No. of PCP visits in 2003												<.001
1-2	10 776	51.7	69.4	79.7	37.7	<.001	46 385	57.1	<.001	13293	72.8	
3-4	19 200	58.9	78.4	87.8	47.8		46 219	58.5		17 005	79.3	
≥5	22 331	61.0	77.3	86.9	48.4		41 127	54.2		19 228	77.1	
Any visit to cardiologist									<.001			<.001
Yes	14 953	61.6	80.6	86.3	50.1	<.001	24193	54.0		29 211	79.9	
No	37 354	57.0	74.2	85.5	44.3		109 538	57.3		20 315	72.1	
Any visit to endocrinologist									<.001			NA
Yes	2489	64.6	75.5	85.4	48.9	.003	2485	49.7		1975	75.7	
No	49 818	58.0	76.1	85.8	45.8		131 246	56.8		47 551	76.7	
Any visit to nephrologist						<.001						<.001
Yes	4028	71.6	81.8	89.8	61.0		4571	57.6		2165	83.1	
No	48 279	57.2	75.6	85.4	44.7		129 160	56.7		47 361	76.4	
Any visit to ED						<.001			<.001			<.001
Yes	19 132	57.3	73.2	84.6	43.2		40 188	51.6		19 530	73.2	
No	33 175	58.9	77.7	86.4	47.6		93 543	58.9		29 996	78.9	
Any hospitalization						<.001						<.001
Yes	20 572	57.2	72.6	84.3	42.9		37 784	48.7	<.001	24 603	73.3	
No	31 735	59.1	78.3	86.7	48.0		95 947	59.8		24 923	80.0	
Medical history												
Deyo score[19]						<.001						<.001
<3	28 883	52.8	76.9	85.2	42.1		116 966	58.1	<.001	37 264	77.5	
≥3	23 424	65.2	75.0	86.5	50.8		16 765	46.8		12 262	74.1	

Abbreviations: CAD, coronary artery disease; ED, emergency department; HbA_{1c}, hemoglobin A_{1c}; HEDIS, Health Plan Employer Data and Information Set; NA, not applicable; PCP, primary care physician.

as measured by the ABIM MOC examination, and basic quality-of-care performance measures. We found a positive correlation between examination scores and performance on all but 1 quality-of-care performance measure. Regarding lipid testing among patients with cardiac disease, nearly 60% of the patients were also cared for by a cardiologist, thus mitigating the influence of the general internist. Our results are consistent with those of a previous study that examined quality of care in fee-for-service Medicare patients, but our study extends understanding of the role of physician factors in several important ways.[7] First, we specifically explored the relationship of levels of physician knowledge and quality of care, finding a positive association. We used a direct measure of physician competence in cognitive skill, instead of only demographic characteristics, which have been used in previous studies.[3] Second, we looked at performance on key quality-of-care measures in close time proximity to performance on a knowledge examination. Third, we used hierarchical models to specifically examine the physician as the unit of analysis and to adjust for multiple patient and physician demographic factors previously shown to affect quality.[7,13]

Our results, like those of Tamblyn et al,[12] who investigated a licensure examination in Canada, raise an important question as to why we found a relationship between physician competency in knowledge and performance on basic quality-of-care processes that many investigators believe should be handled by nonphysicians.[23-26] One potential explanation is that many physician offices currently lack effective systems to ensure that these processes of care occur without direct physician involvement.[27,28] Therefore, the burden to order these studies still falls to the physician, and physicians with better cognitive skills may be more effective in remembering to do the "right thing" or in creating systems. Alternatively, physicians with higher levels of cognitive skills competency may either gravitate toward or create better systems of care delivery. Further-

Table 3. Performance Measure Rates by Physician Characteristics

Physician Characteristics	Diabetes Measures: No. of Patients With Diabetes	Diabetes Measures: No. of Patients Receiving Service, %: Annual Eye Examination	Annual Lipid Test	Twice Yearly HbA_{1c} Test	All 3 Services	Mammography: No. of Female Patients	Mammography: With Annual Mammogram, %	HEDIS CAD Measure: No. of Patients With CAD	HEDIS CAD Measure: With Annual Lipid Test, %
US or Canadian medical school									
Yes	36 874	59.6[b]	75.8	68.7[b]	39.6[b]	101 440	59.2[b]	35 201	77.0[a]
No	15 433	55.3	76.7	59.8	33.9	32 291	48.8	14 325	75.8
Sex									
Male	39 439	57.6[b]	75.6[b]	65.0[b]	36.8[b]	88 079	54.2[b]	38 933	76.4[a]
Female	12 868	60.7	77.4	69.4	41.4	45 652	61.5	10 593	77.7
Years since medical school graduation									
11-20	42 323	58.8	76.3	67.4	38.7	110 444	57.4	40 190	77.0[b]
21-44	6224	54.9	75.3	57.0	32.2	12 759	48.2	5678	74.9
Measure specific patient volume, %									
Lowest 25%	1225	58.0[b]	74.7	62.2[b]	36.1	4140	49.5[b]	1361	72.1[b]
Middle 50%	20 959	59.7	75.8	65.1	38.6	55 761	54.3	18 270	76.0
Highest 25%	30 123	57.4	76.3	66.9	37.5	73 830	58.9	29 895	77.3
Practice type									
Solo	13 136	55.8[b]	75.9[b]	61.3[b]	34.4[b]	30 650	51.6[b]	12 885	76.0[b]
Group, 2-10 physicians	21 508	59.2	76.3	67.6	39.2	57 724	57.5	20 698	77.2
Group, ≥11 physicians	10 937	60.4	77.6	71.3	41.7	30 048	60.1	10 450	77.5
Academic hospital–based	2541	58.0	70.2	60.8	33.3	5031	57.6	1754	71.6
Military or government	219	51.6	62.6	50.2	28.3	431	50.1	189	63.0
Missing data/other	3966	57.0	75.0	63.5	35.5	9847	56.9	3550	77.0
Examination scores (first attempt)									
Internal medicine certification									
<300	2439	56.7[b]	76.1	60.2[b]	34.3[b]	5604	50.6[b]	2167	76.7
300-399	7881	57.6	76.5	62.6	36.4	19824	54.1	7482	77.4
400-499	20 728	58.2	76.3	65.5	37.3	52 903	56.5	19 883	76.7
500-599	18 054	58.4	75.6	68.5	39.1	47 249	58.0	16 930	76.4
≥600	3205	62.3	75.8	68.9	41.4	8151	60.4	3064	76.2
Maintenance of certification									
<300	1218	55.8[b]	74.7[a]	57.8[b]	30.5[b]	2714	50.9[b]	1157	76.1[b]
300-399	6761	56.3	76.6	61.7	34.3	16 685	53.2	6541	75.2
400-499	16 227	58.0	76.9	66.1	38.1	41 027	55.2	15 297	77.9
500-599	18 840	58.6	75.3	67.3	38.6	49 510	58.1	17 919	76.5
≥600	9261	60.2	75.9	68.0	39.8	23 795	59.5	8612	76.0
Lowest 25%									
<435	12 390	56.4[b]	76.9	62.8[b]	35.0[b]	30 970	53.1[b]	11 834	76.4
Middle 50%									
435-580	27 480	58.5	75.7	66.5	38.3	70 698	57.0	26 154	77.0
Highest 25%									
>580	12 437	59.9	76.1	68.3	40.0	32 063	59.3	11 538	76.3

Abbreviations: CAD, coronary artery disease; HbA_{1c}, hemoglobin A_{1c}; HEDIS, Health Plan Employer Data and Information Set.
[a] $P < .01$.
[b] $P < .001$.

more, high performance on the examination may be a marker for more effective physician behaviors in other competencies that promote better care. Previous work supports this hypothesis. Ramsey et al[29] found that physician peers rated colleagues who scored higher on the certification examination more favorably. Shea et al[30] found that examinees who scored higher on the examination received higher ratings in other competencies from their program directors.

Also, research in cognitive psychology has also deepened our understanding of the link between "knowing" and "action."[31] The first step in the clinical reasoning process is to successfully integrate information obtained from the patient with the physician's existing medical knowledge to create a representation of the patient's problem. Physician competence in cognitive skills is an essential element for accurate diagnosis and clinical decision making. Next, physicians should evaluate their initial problem representation to ensure that they have not missed important information from the patients or to identify a gap in their knowledge that necessitates using information resources to acquire new knowledge. However, research suggests that physicians too often skip this important evaluation step and move directly to the last step: action.[32] Our understanding of this clinical reasoning process demonstrates the strong link between cognitive skills, the cognitive skills of integration and synthesis, and actual patient care actions. The ABIM examination is a highly reliable and valid evaluation of clinical decision-making skills and currently represents the only summative evaluation method in the MOC program. Our findings and those of other studies suggest that cognitive skill is an important foundational competency, one that is essential to facilitate other care activities.

This study has several limitations. First, we assessed only the technical aspects of quality. Second, the quality

Table 4. Hierarchical Generalized Linear Models to Evaluate the Association Between Physician Test Score and Performance

Model[a]	Odds Ratio (95% Confidence Interval) Diabetes Mellitus (Perfect Score)	Mammography	Coronary Artery Disease
Unadjusted			
<25th percentile	1 [Reference]	1 [Reference]	1 [Reference]
25th-75th percentile	1.15 (1.07-1.23)	1.19 (1.12-1.25)	1.04 (0.97-1.13)
>75th percentile	1.24 (1.15-1.35)	1.33 (1.25-1.41)	0.99 (0.91-1.08)
Adjusted for patient characteristics[b]			
<25th percentile	1 [Reference]	1 [Reference]	1 [Reference]
25th-75th percentile	1.12 (1.05-1.20)	1.13 (1.08-1.19)	1.01 (0.94-1.10)
>75th percentile	1.24 (1.15-1.34)	1.27 (1.19-1.34)	0.98 (0.89-1.07)
Adjusted for patient and physician characteristics[c]			
<25th percentile	1 [Reference]	1 [Reference]	1 [Reference]
25th-75th percentile	1.08 (1.01-1.16)	1.06 (1.01-1.12)	1.02 (0.94-1.11)
>75th percentile	1.17 (1.08-1.27)	1.14 (1.08-1.21)	1.00 (0.91-1.10)

[a] Mean (SD) between-physician performances on measure were 0.27 (0.52) for diabetes mellitus, 0.34 (0.58) for coronary artery disease, and 0.17 (0.42) for mammography, respectively.

[b] Patient characteristics include demographics (age, race, sex [except for mammography measure], dual-eligible status, and the interaction of dual eligibility and age); comorbidity (dementia, acute myocardial infarction, cancer, chronic obstructive pulmonary disease, congestive heart failure, hypertension, peripheral vascular disease, renal failure, depression, smoking, obesity, liver disease, diabetes [for lipid and mammography measures], percutaneous transluminal coronary angioplasty, and coronary artery bypass grafting); and access to care during 2002 and 2003 (number of visits to primary care physicians, any visit to a specialist [cardiologist, endocrinologist, or nephrologist], or any emergency department visit or hospitalization).

[c] Physician characteristics include sex, US medical school graduate (yes/no), condition-specific patient volume, and practice types (solo and large group or academic vs small group).

measures used are quite modest in stringency; eg, only 2 hemoglobin A_{1c} determinations were required per year, and the measures for lipid testing and mammography were only required once in a 1-year period. Yet, despite this leniency in the measures, we still found a significant relationship between knowledge and performance for diabetes care and mammography screening. Third, we limited our evaluation to physicians with a minimum number of Medicare beneficiaries, and this restriction did not allow an analysis of the association between very low practice volume and knowledge. In a previous study, we found that low-volume physicians performed more poorly on basic diabetic processes of care, and we excluded low-volume providers from this study.[13]

Fourth, our study examined only the competency of knowledge and did not directly investigate competency in quality improvement and systems-based practice. Fifth, our study looked only at process of care measures, not at outcome measures. A recent study questioned the link between process and outcome measures and mortality rates in the hospital setting.[33] However, other studies found that attribution of performance at the physician level is higher with process measures than with outcome measures.[34,35] Sixth, the physician cohort was limited to those who first certified between 1990 and 1995. Therefore, we cannot examine the relationship of knowledge and performance with older physicians. Given the recent meta-analysis showing a decline in knowledge and performance over time, the relationship between knowledge and quality of care should be an area for future investigation.[3] Yet, even among a relatively young cohort of physicians, we still found a relationship between cognitive skills and quality-of-care performance. Finally, we excluded physicians who did not take an MOC examination. We fitted additional models with a dummy variable for physicians who did not take the test and found that there was no difference in performance between physician groups with and without MOC scores. However, the distribution of scores from the physicians' initial certification examination was similar to that of the analyzed cohort and thus likely explains the lack of an association.

Our study findings help to build the evidence link between cognitive skills and quality of care. The MOC examination is part of the medical profession's self-regulatory activities to demonstrate to the public physicians' commitment to maintaining competence in cognitive skills. If the relationship between the examination performance and clinical performance is confirmed in other studies, then using an assessment of cognitive skills through a reliable examination as a quality measure may be reasonable. Clearly, further work is needed to explore the relationship between cognitive skills and competency in other skills, such as accuracy in diagnosis and treatment decisions, and between various competencies and other domains of quality.

Accepted for Publication: December 10, 2007.
Correspondence: Eric S. Holmboe, MD, American Board of Internal Medicine, 510 Walnut St, Ste 1700, Philadelphia, PA 19106 (eholmboe@abim.org).
Author Contributions: Dr Holmboe had full access to all 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:* Holmboe, Meehan, Tate, Ho, and Lipner. *Acquisition of data:* Holmboe, Meehan, Tate, Ho, Starkey, and Lipner. *Analysis and interpretation of data:* Holmboe, Wang, Meehan, Tate, and Lipner. *Drafting of the manuscript:* Holmboe and Starkey. *Critical revision of the manuscript for important intellectual content:* Holmboe, Wang, Meehan, Tate, Ho, and Lipner. *Statistical analysis:* Wang, Tate, Ho, and Lipner. *Admin-*

istrative, technical, and material support: Meehan, Tate, Ho, and Starkey. *Study supervision:* Holmboe, Meehan, and Lipner.

Financial Disclosures: Drs Holmboe and Lipner and Ms Starkey are all employed by ABIM. Drs Wang, Meehan, and Ho and Ms Tate are all employed by Qualidigm. Qualidigm received payment from ABIM Foundation to conduct the analyses.

Funding/Support: This study was supported by a grant from ABIM.

REFERENCES

1. Jencks SF, Cuerdon T, Burwen DR, et al. Quality of medical care delivered to Medicare beneficiaries: a profile at state and national levels. *JAMA*. 2000;284(13):1670-1676.
2. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. *JAMA*. 2003;289(3):305-312.
3. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Ann Intern Med*. 2005;142(4):260-273.
4. Norcini JJ, Lipner RS, Kimball HR. Certifying examination performance and patient outcomes following acute myocardial infarction. *Med Educ*. 2002;36(9):853-859.
5. Norcini JJ, Kimball HR, Lipner RS. Certification and specialization: do they matter in the outcome of acute myocardial infarction? *Acad Med*. 2000;75(12):1193-1198.
6. Chen J, Rathore SS, Wang Y, Radford MJ, Krumholz HM. Physician board certification and the care and outcomes of elderly patients with acute myocardial infarction. *J Gen Intern Med*. 2006;21(3):238-244.
7. Pham HH, Schrag D, Hargraves JL, Bach PB. Delivery of preventive services to older adults by primary care physicians. *JAMA*. 2005;294(4):473-481.
8. Residency program statistics: pass rates and trends. American Board of Internal Medicine Web site. http://www.abim.org/resources/pass_rates_residency.shtm. Accessed May 19, 2007.
9. Brennan TA, Horwitz RI, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification status in the quality movement. *JAMA*. 2004;292(9):1038-1043.
10. Publications. American Board of Internal Medicine Web site. http://www.abim.org/moc/policies/aspx. Accessed June 13, 2007.
11. Grosch EN. Does specialty board certification influence clinical outcomes? *J Eval Clin Pract*. 2006;12(5):473-481.
12. Tamblyn R, Abrahamowicz M, Dauphinee WD, et al. Association between licensure examination scores and practice in primary care. *JAMA*. 2002;288(23):3019-3026.
13. Holmboe ES, Wang Y, Tate JP, Meehan TP. The effects of patient volume on the quality of diabetic care for Medicare beneficiaries. *Med Care*. 2006;44(12):1073-1077.
14. Wasserman SI, Kimball HR, Duffy FD; Task Force on Recertification. Recertification in internal medicine: a program of continuous professional development. *Ann Intern Med*. 2000;133(3):202-208.
15. Moore DS, McCabe GP. *Introduction to the Practice of Statistics*. 3rd ed. New York, NY: WH Freeman & Co; 1998.
16. Altman DG, Bland JM. Statistics notes: quartiles, quintiles, centiles, and other quantiles. *BMJ*. 1994;309(6960):996.
17. Wang OJ, Wang Y, Lichtman JH, Bradley EH, Normand SL, Krumholz HM. "America's Best Hospitals" in the treatment of acute myocardial infarction. *Arch Intern Med*. 2007;167(13):1345-1351.
18. Blake GJ, Pradhan AD, Manson JE, et al. Hemoglobin A_{1c} level and future cardiovascular events among women. *Arch Intern Med*. 2004;164(7):757-761.
19. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45(6):613-619.
20. Normand SLT, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: issues and applications. *J Am Stat Assoc*. 1997;92(439):803-814.
21. Shahian DM, Normand SL, Torchiana DF, et al. Cardiac surgery report cards: comprehensive review and statistical critique. *Ann Thorac Surg*. 2001;72(6):2155-2168.
22. Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical aspects of institutional performance. *J R Stat Soc [Ser A]*. 1996;159:385-444.
23. Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. *Ann Intern Med*. 1997;127(12):1097-1102.
24. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. *Health Aff (Milwood)*. 2001;20(6):64-78.
25. Wagner EH. The role of patient care teams in chronic disease management. *BMJ*. 2000;320(7234):569-572.
26. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. *JAMA*. 2002;288(14):1775-1779.
27. Holmboe E, Kim N, Cohen S, et al. Primary care physicians, office-based practice and the meaning of quality improvement: a qualitative study. *Am J Med*. 2005;118(8):917-922.
28. Audet AMJ, Doty MM, Shamasdin J, Schoenbaum SC. Measure, learn, and improve: physicians' involvement in quality improvement. *Health Aff (Milwood)*. 2005;24(3):843-853.
29. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA*. 1993;269(13):1655-1660.
30. Shea JA, Norcini JJ, Kimball HR. Relationships *of* ratings of clinical competence and ABIM scores to certification status. *Acad Med*. 1993;68(10)(suppl):S22-S24.
31. Gruppen LD, Frohna AZ. Clinical reasoning. In: Norman GR, van der Vleuten CPM, Newble DI, eds. *International Handbook of Research in Medical Education*. Dordrecht, the Netherlands: Kluwer Academic Publishers; 2002:205-230.
32. Eva KW. The aging physician: changes in cognitive processes and their impact on medical practice. *Acad Med*. 2002;77(10)(suppl):S1-S6.
33. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA*. 2006;296(22):2694-2702.
34. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med*. 2002;136(2):111-121.
35. Lipner RS, Weng W, Arnold GK, Duffy FD, Lynn LA, Holmboe ESA. Three-part model for measuring diabetes care in physician practice. *Acad Med*. 2007;82(10)(suppl):S48-S52.