

The Relationship Between Licensing Examination Performance and the Outcomes of Care by International Medical School Graduates

John J. Norcini, PhD, John R. Boulet, PhD, Amy Opalek, MS, and W. Dale Dauphinee, MD

Abstract

Purpose

The Step 2 Clinical Knowledge (CK) examination of the United States Medical Licensing Examination sequence is a requirement for the certification of international medical graduates (IMGs) by the Educational Commission for Foreign Medical Graduates. An association between scores on the test and the quality of care later provided by those who take it is central to its use in certification and licensure. The purpose of this study was to determine whether there is a relationship between scores on Step 2 CK and patient outcomes for IMGs.

Method

This is a retrospective observational study of the 60,958 hospitalizations from 2003 to 2009 in Pennsylvania where the principal diagnosis was acute myocardial infarction or congestive heart failure and the attending physician (N = 2,525) was an IMG who had taken the Step 2 CK. The main measures were the three-digit scores on Step 2 CK and in-hospital mortality.

Results

After adjustment for severity of illness, physician characteristics, and hospital characteristics, performance on Step 2 CK had a statistically significant

inverse relationship with mortality. Each additional point on the examination was associated with a 0.2% (95% CI: 0.1%–0.4%) decrease in mortality. The size of the effect is noteworthy, with each standard deviation (roughly 20 points) equivalent to a 4% change in mortality risk.

Conclusions

These findings provide evidence for the validity of Step 2 CK scores. Given the magnitude of its relationship with patient outcomes, the results support the use of the examination as an effective screening strategy for licensure.

Approximately one-quarter of the practicing physicians in the United States are international medical graduates (IMGs).¹ To take the first step to entry into the U.S. postgraduate clinical education system, these physicians must be certified by the Educational Commission for Foreign Medical Graduates (ECFMG). The requirements for ECFMG certification have changed over time, but in 1992 the first IMG took the Step 2 Clinical Knowledge (CK) examination of the United States Medical Licensing Examination (USMLE) licensing sequence. Although there are some studies documenting the relationship between scores on this examination and other educational measures, there appear to be no data indicating whether they have a relationship with actual performance in practice.^{2–4} This study provides data on

the association between scores of IMGs on Step 2 CK and patient outcomes.

The number of IMGs seeking graduate training in the United States has varied over time as a function of a number of factors.¹ Given the fact that IMGs have attended medical school programs in more than 200 nations, it is not surprising that they are a heterogeneous group. However, one important distinction is between those IMGs who are U.S. citizens (USIMGs) and those who are not (non-USIMGs). USIMGs make up roughly one-fifth of all certified IMGs, are more motivated to take graduate training in the United States, and tend to perform less well than non-USIMGs on some examinations.^{5,6}

Compared with U.S. medical school graduates (USMGs), IMGs as a group tend to fill unmet needs in the U.S. health care system. They are more likely to practice in rural areas and urban areas of poverty, and more often choose primary care specialties where there are both shortages and maldistribution of physicians.⁵ IMGs have increased the diversity of the workforce and have contributed to U.S. health care in all areas of medical practice.^{1,5,7}

IMGs must complete a rigorous credentialing process before they are able to practice in the United States. First, they must be certified by the ECFMG, which requires primary source verification of educational credentials (diploma and transcript) and successful performance on the first two steps of the USMLE, which includes three examinations (Step 1, Step 2 CK, and Step 2 Clinical Skills [CS]). Upon ECFMG certification, IMGs are eligible to enter accredited residency training in the United States. To then get a state medical license, IMGs must complete at least one year of residency training (required years of training differ by state) and pass the USMLE Step 3. Upon completion of a graduate training program, they can apply for specialty board certification, which requires further examination.

Despite the rigor of this process, there have been questions about the competence of IMGs.^{8,9} However, current work has demonstrated that they are performing well on educational measures, and a majority have achieved specialty board certification.¹⁰ Moreover, a recent study has shown that the patient outcomes of IMGs are comparable to those of USMGs for congestive heart failure (CHF) and acute myocardial infarction (AMI).¹¹

Please see the end of this article for information about the authors.

Correspondence should be addressed to Dr. Norcini, FAIMER, 3624 Market St., 4th Floor, Philadelphia, PA 19104; telephone: (215) 823-2170; e-mail: jnorcini@faimer.org.

Acad Med. 2014;89:1157–1162.

First published online May 21, 2014
doi: 10.1097/ACM.0000000000000310

Interestingly, the performance of non-USIMGs for these conditions was better than that of USMGs.

Among the requirements for ECFMG certification is successful performance on the USMLE Step 2 CK examination. According to the USMLE, the content of this examination provides the foundation for the “safe and competent practice of medicine.”¹² The examination measures the application of medical knowledge, skills, and clinical science required for the provision of patient care under supervision. It is a single-day, computer-based examination administered at sites throughout the world.

Because it is required for ECFMG certification and for licensure in the United States, Step 2 CK scores have been used as the basis for several studies.^{2-4,13}

The vast majority of experts judge the content to be clinically relevant, appropriate for the examination, and useful in clinical practice.³ Moreover, there is little redundancy with Step 2 CS (an examination based on an assessment of clinical skills using standardized patients).²

Although the studies referenced above speak to the utility of the Step 2 CK examination, the scores from neither this examination nor any of the others in the USMLE sequence have been compared with actual performance in practice. Such data are important in the context of the regulatory and administrative burdens in U.S. health care.¹⁴ Our goal with this study was to fill this gap by comparing Step 2 CK scores with in-hospital mortality for CHF and AMI. We chose these conditions because they are common, the outcomes are important, and they are often treated by doctors who are IMGs.

Method

Sources of data

We conducted this study using the inpatient records for Pennsylvania from January 1, 2003, to December 31, 2009, which were obtained from the Pennsylvania Health Care Cost Containment Council (PHC4). All hospitals in Pennsylvania (except Veterans Administration hospitals, skilled nursing facilities, and state psychiatric hospitals) are required to submit a patient record to PHC4 each

time there is a discharge. Uniform Billing data are sent directly to PHC4, while the patient’s demographics, comorbid conditions, and other important clinical information from the beginning of their stay (such as the history, laboratory data, symptoms, and pathophysiology) are sent to MediQual.¹⁵ MediQual applies a statistical model which yields the Atlas Admission Severity index reflecting the probability of death where 0 indicates no clinical instability (<0.001), 1 is minimal instability (0.002–0.011), 2 is moderate instability (0.012–0.057), 3 is severe instability (0.058–0.499), and 4 is maximal instability (0.50–1.0). These data were sent to PHC4 where they were combined with the Uniform Billing data. For this study, we combined patients with index values of 0 and 1.

In these records, the attending physician was identified for each hospitalization. PHC4 defines this as the physician who was primarily responsible for the patient’s medical care and who certified the necessity for services.

From these data we took the 685,774 hospitalizations where the principal diagnosis was AMI or CHF. We chose these diagnoses because they occur frequently and are often used to judge the quality of care.¹⁶ Using a common physician identifier, we matched these data with the 2010 American Medical Association (AMA) Masterfile, which contains information on all physicians who reside in the United States and have met credentialing requirements for recognition.

Hospitalizations for both conditions were excluded if the patient was under 18 years, discharge status was not available, or the patient transferred from another short-term facility. In addition, we excluded hospitalizations for AMI if the patient’s admission status was not available or if the patient transferred to another short-term facility.

Of the 685,774 hospitalizations, we could not match 4,579 (<1%) with a physician in the Masterfile. Further, we eliminated hospitalizations where there were diagnosis-related exclusions (80,440 or 12%), severity of illness information was missing (11,000 or 2%), and the attending physician graduated before 1959, when the ECFMG began its

work (3,735 or <1%). This left 585,990 hospitalizations.

We then matched these hospitalizations with information from the ECFMG. The ECFMG files contain data on all IMGs who have applied for its certification, which is required for admission to graduate training in the United States. The ECFMG files also contain the scores of candidates on the USMLE Step 2 CK examination, which was first administered in 1992. After we eliminated the 445,825 hospitalizations where USMGs were the attending physician, as well as those 75,086 hospitalizations where IMGs were the attending physician and they did not take USMLE Step 2 CK (they took a comparable examination before 1992), 60,958 hospitalizations attended by 2,525 doctors were left for analysis.

Because all candidates for ECFMG certification sign a release that permits their deidentified data to be used for research purposes and the other data are publicly available for research purposes, no ethical approval was required for this study.

Data elements

The patients’ age, sex, race, principal diagnosis, admission severity, and discharge status (which indicated mortality) were available from the PHC4 data. These records also indicated the facility where the patient was treated.

The physicians’ self-reported specialization and specialty board certification were available from the AMA Masterfile via agreement with the American Board of Medical Specialties. Data on whether physicians were international graduates, their citizenship at entry to medical school, and their number of attempts and scores on USMLE Step 2 CK were available from the ECFMG. Only the three-digit score from the physician’s first attempt at USMLE Step 2 CK is used in this study (80% of the doctors passed on their first attempt). Most scores on this test range from 140 to 260, and their equivalence is maintained over years by statistical methods.

We calculated two additional variables. For each physician, we tallied the number of CHF and AMIs treated. For each facility, we used a list from the Pennsylvania Office of Rural Health to determine if its county location was urban or rural.¹⁷

Analyses

We calculated descriptive statistics for patients, physicians, and facilities using the hospitalizations as the basis for analysis. A multivariate model was used to assess the relationship between patient mortality and the scores of physicians on USMLE Step 2 CK. We adjusted the model for severity of illness on admission, whether the principal condition was AMI, facility volume and rural location, whether the doctor was a certified and self-identified family physician, whether the doctor was a certified and self-identified internist, whether the doctor was a certified and self-identified cardiologist, whether the doctor was a non-U.S. citizen at entry to medical school, and physician case volume. Because patients are clustered within physicians and physicians are clustered within facilities, generalized estimating equations were applied (GENMOD procedure, SAS version 9.1, SAS Institute).

Potential confounding and biasing variables

We undertook analyses to address the possibility that confounding variables might be influencing the results. Because physicians' clustering within hospitals

may be nonrandom, we correlated average USMLE scores with mortality for each facility. The resulting correlation of .01 ($P > .05$) was not statistically significant.

We were also concerned that there might be an effect of specialization, so we ran the multivariate analysis limiting the physicians to self-designated family doctors, internists, and cardiologists. The nature of the relationship between scores and mortality was similar to the results reported below.

Previous work has shown that the time since medical school graduation (which is highly correlated with age) has an association with physician performance.^{11,18} We tested the influence of this variable and found that it did not have a statistically significant association with mortality. This is likely due to the fact that the participants in the study were relatively close to graduation, having all attempted the USMLE Step 2 CK examination in 1992 or later.

Results

Characteristics of physicians, patients, and hospitals

Of the 2,525 physicians in the study, 2,152 were non-USIMGs and 373

were USIMGs. There were 243 (10%) board certified self-designated family medicine doctors, 1,169 (46%) board certified self-designated internists, and 134 (5%) board certified self-designated cardiologists. In addition, 248 physicians (10%) identified themselves as practitioners of these disciplines but were not board certified in them. These 1,794 physicians managed 50,855 (83%) of the hospitalizations analyzed. The remaining 731 physicians had 48 different practice specialties, and they managed 10,103 hospitalizations (17%).

There were 173 hospitals included in the study. These facilities had from 4 to 18,262 hospitalizations, with a median of 2,462.

Table 1 presents information on the number of hospitalizations stratified by the characteristics of the facilities and physicians, including data on patient volume and performance on the USMLE Step 2 CK examination. The non-USIMGs had slightly higher patient volume during the period of study than the USIMGs. The non-USIMGs also had higher test scores than the USIMGs. Both of these effects were statistically significant.

Table 1

Number of Hospitalizations in Pennsylvania, Stratified by the Characteristics of the Facilities and the Physicians, 2003–2009^a

Characteristic	Physicians		All hospitalizations	P value
	Non-USIMGs (N = 2,152)	USIMGs (N = 373)		
Facility, location of institution, no. (%)				
Urban	38,243 (73)	6,197 (72)	44,440 (73)	
Rural	14,120 (27)	2,398 (28)	16,518 (27)	>.05 ^b
Total no.	52,363	8,595	60,958	
Physicians, board certified and self-designated specialty, no. (%)				
Board certified and self-identified family medicine	2,353 (64)	1,331 (36)	3,684 (6)	<.001 ^c
Board certified and self-identified internal medicine	29,579 (87)	4,469 (13)	34,048 (56)	<.001 ^c
Board certified and self-identified cardiologist	5,712 (93)	429 (7)	6,141 (10)	<.001 ^c
Other	14,719 (86)	2,366 (14)	17,085 (28)	<.001 ^c
Total no.	52,363	8,595	60,958	
Physician volume and Step 2 CK score, mean (SD)				
Volume	24 (44)	23 (36)	24 (43)	<.001 ^d
Step 2 CK	197 (32)	190 (30)	196 (31)	<.001 ^d
Total no.	221	213	220	

Abbreviations: USIMG indicates international medical graduates who are U.S. citizens; Step 2 CK, United States Medical Licensing Examination Step 2 Clinical Knowledge examination.

^aFrom a study of the relationship between scores on the Step 2 CK examination and patient outcomes for international medical graduates.

^bP value calculated using chi-square.

^cP value calculated using chi-square and comparing each specialty group with all other physicians.

^dP value calculated using two-sided Student *t* test.

As shown in Table 2, there were 60,958 hospitalizations from 2003 to 2009 with a principal diagnosis of CHF or AMI and where the attending physician was an IMG who took the Step 2 CK examination. Patient mortality was 4.4% overall, 2.9% for CHF, and 7.9% for AMI. Differences between the groups of physicians were

statistically significant, with patients of non-USIMGs having lower mortality.

Multivariate analyses

Table 3 presents the results of the multivariate analysis with USIMGs as the reference group. It includes parameter estimates, confidence intervals, and the

adjusted odds ratios for all variables included in the analysis. In this section, to aid in the interpretation of the results, we report the results as changes in relative risk.

Adjusting for characteristics of the patients, physicians, and facilities, performance on the USMLE Step 2 CK examination had a statistically significant inverse relationship with mortality. Each additional point on the USMLE examination was associated with a 0.2% (95% CI: 0.1%–0.4%) decrease in mortality.

Facility volume and location did not have statistically significant associations with the log odds of mortality, but several physician characteristics did. Being a certified and self-designated family doctor was associated with a decrease of 37% (CI: 16% to 59%) in mortality. Likewise, being a certified and self-designated internist was associated with a 27% (CI: 15% to 39%) decrease in mortality. Individual physician case volumes were inversely related to the log odds of mortality: Each additional AMI or CHF hospitalization was associated with a 0.1% (CI: 0.1% to 0.2%) decrease in patient mortality. Finally, patients of non-USIMGs had 20% lower mortality (CI: 1% to 38%) than patients of USIMGs.

Discussion

In this study, we looked for a relationship between scores on the Step 2 CK examination and in-hospital mortality for patients with CHF or AMI. Conditioning on a number of patient, physician, and facility characteristics, we found that better examination performance was associated with a decrease in patients' relative risk for mortality. The size of the effect was noteworthy, with each standard deviation (roughly 20 points) equivalent to a 4% change in relative risk.

Controlling for all other factors, having a doctor who is a board-certified and self-designated family physician or internist is associated with lower relative risk of patient mortality. Likewise, having an attending physician with more experience with these conditions also produces lower relative risk of mortality. Finally, patients of non-USIMGs had a 20% lower relative risk of mortality than the patients of USIMGs. These results are consistent with previous work.^{11,19–21}

It is important to recognize that the findings reported here likely understate

Table 2
Characteristics of the Patients by Attending Physician, Derived From All Hospitalizations in Pennsylvania, 2003–2009^a

Characteristic	No. (%) of patients whose attending physicians were:		All hospitalizations	P value
	Non-USIMGs	USIMGs		
Age				
18–49	3,908 (7)	491 (6)	4,399 (7)	
50–64	10,102 (19)	1,442 (17)	11,544 (19)	
65–74	10,077 (19)	1,705 (20)	11,782 (19)	
75–84	15,517 (30)	2,760 (32)	18,277 (30)	
≥85	12,757 (24)	2,197 (26)	14,954 (25)	
Total	52,363 (86)	8,595 (14)	60,958	<.001 ^b
Sex				
Female	26,853 (51)	4,281 (50)	38,152 (33)	
Male	25,509 (49)	4,314 (50)	76,599 (67)	<.05 ^b
Total	52,362 ^c	8,595	114,752	
Race				
White	41,983 (80)	7,385 (86)	49,368 (81)	
Black	7,878 (15)	801 (9)	8,679 (14)	
Asian or Pacific Islander	180 (<1)	17 (<1)	197 (<1)	
Native American or Eskimo	22 (<1)	1 (<1)	23 (<1)	
Other/mixed	1,142 (2)	123 (1)	1,263 (2)	
Unknown	1,149 (2)	267 (3)	1,418 (2)	<.001 ^b
Total	52,354 ^c	8,594 ^c	60,948 ^c	
Condition				
AMI	15,420 (29)	2,477 (29)	17,897 (29)	
CHF	36,943 (71)	6,118 (71)	43,061 (71)	
Total	52,363	8,595	60,958	
Admission severity				
None or minimal	11,496 (22)	1,697 (20)	13,193 (22)	
Moderate	28,249 (54)	4,680 (54)	32,929 (54)	
Severe	12,076 (23)	2,097 (24)	14,173 (23)	
Maximal	542 (1)	121 (1)	663 (1)	<.001 ^b
Total	52,363	8,595	60,958	
In-hospital deaths^d				
AMI	1,183 (7.7)	234 (9.5)	1,417 (7.9)	<.001 ^b
CHF	1,030 (2.8)	239 (3.9)	1,269 (2.9)	<.001 ^b
Total	2,213	473	2,686	

Abbreviations: USIMG indicates international medical graduates who are U.S. citizens; Step 2 CK, United States Medical Licensing Examination Step 2 Clinical Knowledge examination; AMI, acute myocardial infarction; CHF, congestive heart failure.

^aFrom a study of the relationship between scores on the Step 2 CK examination and patient outcomes for international medical graduates.

^bP value calculated using chi-square.

^cResponses were not given by all participating respondents for these categories.

^dHospitalizations for that condition.

Table 3

Parameter Estimates, 95% CI, and Adjusted Odds Ratios for In-Hospital Mortality, for all Hospitalizations in Pennsylvania, 2003–2009^a

Parameter	Parameter estimate	95% CI	Adjusted odds ratio
Condition			
AMI	0.460	0.350, 0.570	1.584
Admissions severity			
Moderate	1.621	1.340, 1.903	5.058
Severe	3.303	3.025, 3.580	27.194
Maximal	5.388	5.076, 5.700	218.765
Attending physician			
Step 2 CK score	-0.002	-0.004, -0.001	0.998
Board certified and self-identified family medicine	-0.371	-0.585, -0.157	0.690
Board certified and self-identified internal medicine	-0.270	-0.389, -0.152	0.763
Board certified and self-identified cardiologist	-0.119	-0.318, 0.080	0.888
Non-USIMG	-0.198	-0.377, -0.012	0.820
Case volume	-0.001	-0.002, -0.001	0.999
Facility			
Case volume	-0.000	-0.000, 0.000	1.000
Rural location	0.061	-0.053, 0.174	1.063

Abbreviations: AMI indicates acute myocardial infarction; Step 2 CK, United States Medical Licensing Examination Step 2 Clinical Knowledge examination; USIMG, international medical graduates who are U.S. citizens.

^aFrom a study of the relationship between scores on the Step 2 CK examination and patient outcomes for international medical graduates. To aid in the interpretation of the results, these results are reported as changes in relative risk in the text.

the true relationship between Step 2 CK scores and patient outcomes. IMGs who did not ultimately pass the Step 2 CK examination or did not pass the other examinations in the USMLE sequence were not eligible to train and practice in the United States, so they are not included in this study. Likewise, those who passed, but could not acquire a residency position in the United States, are not included. Therefore, the relationships described in this study are subject to restriction of range, which attenuates the magnitude of the reported relationships. Moreover, there are a number of subcomponents in the Step 2 CK examination, and some of them are more relevant to CHF and AMI than others. The score–outcomes relationship might be greater had we been able to focus only on the relevant sections of the examination.

This observational study has some limitations. We included data on patients, physicians, and facilities in an effort to limit the effect of potential biases and confounders. In addition, we conducted supplementary analyses intended to

rule out effects such as specialization. Nonetheless, there might be other variables, not available in our data set, which might clarify these findings. For instance, PHC4 provides guidance in terms of the identification of the attending physician, but there may have been variability across facilities.

The study was done with only two inpatient conditions in one state, and the results may or may not generalize to other conditions and locations. Further, it is conceivable that those with better scores on the USMLE Step 2 CK examination attended better residency training programs, which in turn led to better patient outcomes. However, the relationship between the USMLE Step 2 CK and patient mortality exists even after controlling for specialization and board certification, which should be sensitive to these same issues. Further research is needed to address these limitations.

Future research should also address the association between the USMLE Step 2 CS examination and the outcomes of

care.¹⁴ This examination provides an assessment of clinical skills based on the use of standardized patients, and thus it might be expected to have an even larger association with the outcomes of care. Further, there is little redundancy between the USMLE Step 2 CK and CS examinations.² Unfortunately, we could not include Step 2 CS performance in our study because it was first introduced several years after the USMLE Step 2 CK and there are not yet enough patient data for good analysis.

Our findings provide evidence for the validity of Step 2 CK scores. Moreover, they are consistent with the growing literature suggesting that national, high-stakes examinations have a positive relationship with patient outcomes.^{11,19–23} It is challenging to gather this type of validity evidence because it is not possible to randomize to treatment, and the patient outcomes follow the test at a significant time interval. Nonetheless, it is critical to collect data that speak to this relationship to justify the utility of the examinations.¹⁴ This study indicates that, given the magnitude of their association with patient outcomes, these exams are an effective screening test for physician licensure.

Funding/Support: This work was funded by the Foundation for Advancement of International Medical Education and Research and the Educational Commission for Foreign Medical Graduates.

Other disclosures: None reported.

Ethical approval: Reported as not applicable. The American Medical Association Masterfile and the Pennsylvania Health Care Cost Containment Council data are deidentified and publicly available datasets. Candidates for Educational Commission for Foreign Medical Graduates certification sign a release allowing the use of their data for research purposes.

Disclaimers: This work does not necessarily reflect the opinions of the Foundation for Advancement of International Medical Education and Research and the Educational Commission for Foreign Medical Graduates.

Dr. Norcini is president and chief executive officer, Foundation for Advancement of International Medical Education and Research, Philadelphia, Pennsylvania.

Dr. Boulet is associate vice president for research and data resources, Foundation for Advancement of International Medical Education and Research, Philadelphia, Pennsylvania.

Ms. Opalek is information scientist, Foundation for Advancement of International Medical Education and Research, Philadelphia, Pennsylvania.

Dr. Dauphinee is senior scholar, Foundation for Advancement of International Medical Education and Research, Philadelphia, Pennsylvania.

References

- 1 Boulet JR, Cooper RA, Seeling SS, Norcini JJ, McKinley DW. U.S. citizens who obtain their medical degrees abroad: An overview, 1992–2006. *Health Aff (Millwood)*. 2009;28:226–233.
- 2 Harik P, Clauser BE, Grabovsky I, Margolis MJ, Dillon GF, Boulet JR. Relationships among subcomponents of the USMLE Step 2 Clinical Skills Examination, the Step 1, and the Step 2 Clinical Knowledge Examinations. *Acad Med*. 2006;81(10 suppl):S21–S24.
- 3 Cuddy MM, Dillon GF, Clauser BE, et al. Assessing the validity of the USMLE Step 2 clinical knowledge examination through an evaluation of its clinical relevance. *Acad Med*. 2004;79(10 suppl):S43–S45.
- 4 Cuddy MM, Swanson DB, Dillon GF, Holtman MC, Clauser BE. A multilevel analysis of the relationships between selected examinee characteristics and United States Medical Licensing Examination Step 2 Clinical Knowledge performance: Revisiting old findings and asking new questions. *Acad Med*. 2006;81(10 suppl):S103–S107.
- 5 Boulet JR, Norcini JJ, Whelan GP, Hallock JA, Seeling SS. The international medical graduate pipeline: Recent trends in certification and residency training. *Health Aff (Millwood)*. 2006;25:469–477.
- 6 Norcini J, Anderson MB, McKinley DW. The medical education of United States citizens who train abroad. *Surgery*. 2006;140:338–346.
- 7 Norcini JJ, van Zanten M, Boulet JR. The contribution of international medical graduates to diversity in the U.S. physician workforce: Graduate medical education. *J Health Care Poor Underserved*. 2008;19:493–499.
- 8 Mick SS, Comfort ME. The quality of care of international medical graduates: How does it compare to that of U.S. medical graduates? *Med Care Res Rev*. 1997;54:379–413.
- 9 Winward ML, Ripkey DR, Case SM, Morrison CA. Performance of foreign medical graduates on the Clinical Science component of the United States Medical Licensing Examination: Initial and ultimate pass rates. In: *Proceedings of the Eighth International Ottawa Conference on Medical Education and Assessment*. Philadelphia, Pa: National Board of Medical Examiners; 1998:67–74.
- 10 Garibaldi RA, Subhiyah R, Moore ME, Waxman H. The in-training examination in internal medicine: An analysis of resident performance over time. *Ann Intern Med*. 2002;137:505–510.
- 11 Norcini JJ, Boulet JR, Dauphinee WD, Opalek A, Krantz ID, Anderson ST. Evaluating the quality of care provided by graduates of international medical schools. *Health Aff (Millwood)*. 2010;29:1461–1468.
- 12 United States Medical Licensing Examination. *Bulletin*. <http://www.usmle.org/bulletin/overview/#threesteps>. Accessed March 9, 2014.
- 13 Kleshinski J, Khuder SA, Shapiro JJ, Gold JP. Impact of preadmission variables on USMLE Step 1 and Step 2 performance. *Adv Health Sci Educ Theory Pract*. 2009;14:69–78.
- 14 Lehman EP, Guercio JR. The Step 2 Clinical Skills exam—a poor value proposition. *N Engl J Med*. 2013;368:889–891.
- 15 Cardinal Health. *MediQual™ Atlas Outcomes*. <http://www.mediqua.com/products/atlasoutcomes.asp>. Accessed March 9, 2014.
- 16 Agency for Healthcare Research and Quality. *Guide to Inpatient Quality Indicators: Quality of Care in Hospitals—Volume, Mortality, and Utilization, Version 3.1*. Rockville, Md: Agency for Healthcare Research and Quality (AHRQ); 2007.
- 17 Pennsylvania Office of Rural Health. *Definitions of rural*. <http://www.porh.psu.edu/>. Accessed March 9, 2014.
- 18 Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: The relationship between clinical experience and quality of health care. *Ann Intern Med*. 2005;142:260–273.
- 19 Norcini JJ, Lipner RS, Kimball HR. Certifying examination performance and patient outcomes following acute myocardial infarction. *Med Educ*. 2002;36:853–859.
- 20 Tamblyn R, Abramhamowicz M, Brailovsky C, et al. Positive association between licensing examination scores and selected aspects of resource use and quality of care in primary care. *JAMA*. 1998;280:989–998.
- 21 Tamblyn R, Abrahamowicz M, Dauphinee WD, et al. Association between licensure examination scores and practice in primary care. *JAMA*. 2002;288:3019–3026.
- 22 Norcini JJ, Kimball HR, Lipner RS. Certification and specialization: Do they matter in the outcome of acute myocardial infarction? *Acad Med*. 2000;75:1193–1198.
- 23 Wenghofer E, Klass D, Abrahamowicz M, et al. Physician scores on national qualifying examinations predict quality of care in future practice. *Med Educ*. 2009;43:1166–1173.