Review Article

The Quality of Care of International Medical Graduates: How Does It Compare to That of U.S. Medical Graduates?

Stephen S. Mick
Maureen E. Comfort
University of Michigan

This article reviews the literature on the quality of care provided by foreign-trained physicians (international medical graduates or IMGs) compared with that of U.S. medical graduates (USMGs). As concerns are raised about IMGs in the U.S. physician workforce, there are suggestions that IMGs do not deliver care equal in quality to that of USMGs. The review of process and outcome studies finds little support for this claim. However, lower IMG levels of performance on structural measures of quality like credentialing examinations exist and may indicate quality differences. Because no consistent evidence exists that there is a connection between IMG test scores and process or outcome measures of quality of care, whether test scores matter in clinical practice and its outcome is uncertain. Until research shows the contrary, one should be cautious in accepting IMG-USMG quality arguments to support policy to reduce the size of the IMG component of the physician workforce.

One of the most enduring policy issues surrounding international medical graduates (IMGs)—also known as foreign medical graduates (FMGs)—is whether they provide care of quality equal to that given by U.S. medical graduates (USMGs).[1] Historically, some observers have argued that people

Address correspondence and requests for reprints to Stephen S. Mick, Department of Health Management and Policy, School of Public Health, University of Michigan, 109 Observatory, Ann Arbor, MI 48109. The authors wish not only to thank Madelaine Pfahler and Mark Mullen for their

Medical Care Research and Review, Vol. 54 No. 4, (December 1997) 379-413

trained in non-U.S. medical schools, in a language other than English, would be unlikely to practice medicine as well as USMGs (Torrey and Taylor 1973; Weiss et al., "The Effect," 1974; Derbyshire 1975). The issue remains a focal point of debate today, especially in view of the dramatic increase since 1989 of the number of IMG residents in U.S. hospitals and the strong tendency of IMGs to remain in this country after completion of training.[2] Organizations such as the Institute of Medicine (1996), the Pew Health Professions Commission (1995), the Council on Graduate Medical Education (1995a), and a consortium of professional medical groups led by the Association of American Medical Colleges (1997) have all called for reductions in the number of IMGs in residency training.

The policy context of the IMG quality of care debate consists of (1) the consensus among organized medicine and certain federal agencies that the United States now has a surplus of physicians (Rivo, Jackson, and Clare 1993; Mullan, Politzer, and Davis 1995), (2) the increasing level of public expenditures (especially Medicare) to support graduate medical education (Congressional Budget Office 1995), and (3) the impact of managed care plans on reducing the number of specialists that the nation needs (Council on Graduate Medical Education 1995b). IMGs are thought to exacerbate all these problems by contributing to the physician surplus, adding to the amount spent on graduate medical education, and increasing specialization. In addition to these problems, the Institute of Medicine (1996) has called into question the quality of care provided by IMGs, thus adding one more reason to reduce the nation's reliance on them.[3]

The explanation for the lower level of IMG quality rests in the "circumventing the continuum" argument (Lockett 1975): IMGs do not participate in the sequential stages of education and training of U.S. physicians, including graduation from a 4-year undergraduate college or university, a 4-year medical school curriculum, a 3- to 5-year residency program and possibly 1 or 2 extra years in a fellowship position, all interspersed with rigorous entrance and screening examinations such as the Medical College Admission Test (MCAT), the U.S. Medical Licensure Examination (USMLE), and specialty board examinations.

---

research assistance in this work but also to acknowledge the review and comment of Robert M. Politizer and Stephen Tise of the Bureau of Health Professions, U.S. Department of Health and Human Services. In addition, the critiques of two anonymous reviewers were especially helpful. Support for this research is based, in part, on funding provided by the Bureau of Health Professions, but the opinions expressed in this article are those of the authors alone. This article, submitted to *Medical Care Research and Review* on April 1, 1996, was revised and accepted for publication on June 20, 1997.

NEW CONTRIBUTION

The present study is the first in over two decades to assess whether there is reliable research to support the hypothesis of poorer IMG quality. Several decades ago, investigators found little evidence in this vein (see Sutnick 1970). A few years later, Williams and Brook (1975), in their then state-of-the-art review of IMG quality, concluded the following:

> Available data on quality do not permit unequivocal conclusions about the level of care provided by FMGs. Inferences have been made, primarily on the basis of structural variables, that FMGs . . . are likely to provide lower-quality care than fully qualified USMGs. Until more hard data become available, policy decisions regarding the regulation of FMGs must be based largely on "expert opinion" regarding quality of care provided by FMGs, or on grounds other than whether higher or lower quality of care is provided by FMGs. (P. 570)

These authors challenged the relationship between traditional structural indicators of quality such as test scores and quality of care actually delivered. They appealed for more research on process and outcome measures of quality. Through a systematic review of the literature, we determine what progress has been made since the mid-1970s in addressing the question of IMG-USMG quality.

CLASSIFYING STUDIES OF QUALITY

The structure/process/outcome scheme of Donabedian (1980) provides a way to organize the disparate studies on IMG-USMG quality.[4] Structural measures of quality include those based on test scores, board certification rates, and the like. Process measures include those based on procedures and guidelines followed, or not followed, in the act of providing care. Outcome measures include those based on mortality rates, complication rates, rehospitalization rates, malpractice experience, among others. These aspects of quality are sequentially linked and, to a certain extent, dependent on their precursors—for example, quality of outcomes is partly dependent on structural and process quality of care. The exact relationship between structure, process, and resultant outcomes has not been clearly delineated and awaits the results of quality-of-care research. Nevertheless, the hypothesized linkage of structure or process measures with outcome measures is the usual basis for studies that evaluate quality.

Between 1975 and mid-1997, the IMG quality research literature consisted of 88 publications. These studies were identified in the following way. First, the senior author developed an extensive file of IMG-related research and

commentary that covered the period 1948-1997. All post-1975 studies that treated the IMG quality issue formed the first group of studies reviewed. Second, we performed computer searches using various terms to trace studies that had any bearing on the IMG quality issue. We then excluded the innumerable reports that simply contained statistics on IMG and USMG performance (i.e., pass rates and point distributions) on various examinations that medical students and graduates routinely took during and after their training.

These sources provided the bulk of the studies that were the basis of the present review. A close reading of them revealed bibliographic references to a few more relevant studies, and through this "snowball" technique, we arrived at the figure of 88 publications. The majority of these (46, or 52 percent) focused on structural measures of quality, with the remainder about evenly divided between process measures (22, or 25 percent) and outcome measures (20, or 23 percent). Of all these studies, we reported the results of 48 in this article with the excluded 40 being entirely in the structural realm. This was because the latter were based on previously published data and were not original contributions.

The first important characteristic common to this research was its weak methodological rigor. Most studies had multiple deficiencies, and only a few had but one major deficiency—for example, limited, small, or nonrandom samples; questionable or narrow measures of quality; weak research designs making apparent differences difficult to attribute to an IMG-USMG distinction; lack of control variables; lack of statistical significance tests. A meta-analysis was impossible because of the absence of a commonly agreed upon dependent quality variable and the lack of experimental research designs. The only feasible option was to report those studies that had the fewest methodological weaknesses.[5]

STRUCTURAL STUDIES

Structural measures have the advantage of being multiple and plentiful. Many measures are also reliable, to the extent that they rely heavily on standardized examinations that have been repeated in the same form over a number of years. These, however, tend to be paper-and-pencil tests of academic knowledge only and only recently include assessments of diagnostic or treatment planning skills (we treat these latter in the section on process studies). Clearly, structural measures evaluate some portion of what is important for a physician to know in order to provide quality care; however, what portion they do not assess is unclear. As such, it may be best to consider most such measures as screens, functioning to eliminate those persons whose lack of knowledge makes it improbable that they will provide quality care.

EXAMINATIONS AS STRUCTURAL MEASURES OF QUALITY

The most common way IMGs are contrasted to USMGs is by comparing examination scores, the classical structural measure. We did not intend to treat this complicated and vast literature (and did not include any of it in our count of research publications on IMG-USMG quality) for two reasons. First, it consisted mostly of test results with neither commentary nor comparison to USMGs. Second, our objective was to move beyond test results as indicators of quality, a need recognized by Williams and Brook (1975) 20 years ago. However, because there is still heavy reliance on test performance by both the medical and the health services research communities (see Institute of Medicine 1996), we were obliged to discuss the subject.

The Educational Commission for Foreign Medical Graduates (ECFMG) is the organization charged since 1956 to test the preparedness of IMGs to undertake graduate medical education in the United States (see ECFMG 1985). The high failure rate of IMGs on its examinations has buttressed the argument that they are less competent than USMGs.[6] For the period 1958 through 1969, the number of IMGs passing a particular administration of the examination varied from a low of 30.7 percent to a high of 51.0 percent (ECFMG 1969). Overall, of those IMGs who took the ECFMG examination during this period, 58.0 percent eventually passed (Dublin and Oesterling 1987). Yet, it is a fallacy to conclude that low pass rates mean that IMGs are less competent than USMGs. The examination can be seen as a screen to weed out incompetent IMGs.

To this there are two objections. First is the issue of *how well* IMGs do on this examination compared with USMGs. This point can, in turn, be subdivided into three problems: (1) the problem of repeat examination takers, (2) the problem of differences in IMG-USMG scores, and (3) the problem of differences in rigor of the examinations taken by IMGs compared with those taken by USMGs. The second objection stems from the belief that many IMGs have entered, and are entering, into various forms of U.S. medical practice *without* taking and passing the ECFMG examination, leading to an absence of any quality screen. We now discuss these points.

HOW WELL IMGs PERFORM

Repeat Examination Takers

Previously, an examinee could take the ECFMG examination a number of times until a passing mark was attained. For example, during the period 1969-1982, 205,542 IMGs took the examination at least once; of those who did not pass on the first try, 71,890 took the examination a second time. A third

attempt was made by 39,030 IMGs; a fourth attempt by 22,629; a fifth, by 13,713; a sixth, by 8,658 (Dublin and Oesterling 1987). However, the more times an IMG took the examination, the lower his or her probability of passing. The pass rate of first examination takers mentioned above was 40.0 percent; for 5th-time takers, 20.0 percent; for 10th-time takers, 13.0 percent; and 11.1 percent for 20th-time takers (Dublin and Oesterling 1987). The fact that some IMGs have taken the ECFMG examination more than once does not necessarily lead to the conclusion that they are incompetent, particularly given the sizable correlation between passing and whether English was the native tongue of the examinee (Mick and Mou 1991). However, repeat testing may reduce the screening effectiveness of the exam.

Differences in IMG-USMG Examination Scores

Lower scoring by IMGs than USMGs may be more pertinent. Because the questions in all versions of the ECFMG examination were derived from questions used on previous National Board examinations, the proportion of U.S. students correctly answering each question is known. One is able therefore to compute the rates at which U.S. students could have been *expected* to pass the ECFMG's examinations and compare them to the actual performance of IMGs (see ECFMG 1982). Such comparisons have always shown a lower *actual* pass rate among IMGs when compared with a higher expected pass rate for USMGs. For example, on the ECFMG examinations given between 1958 and 1984, from 30 percent to 50 percent of IMGs passed on the first try, whereas the expected rate of USMGs was always more than 90 percent. More to the point, the scores of IMGs were lower. In the July 1981 examination, for example, the IMG mean was 71.4 points, whereas the expected mean for USMGs was 82.2 points (ECFMG 1982).

If quality is a function of performance on the ECFMG examinations, it would be hoped that a substantial number of IMGs passing would well exceed the minimum pass level, set at 75 points. As Lowin (1975) points out, of IMGs passing the July 1974 examination, for example, 70 percent achieved no better than 4 points beyond the minimum passing grade. The conclusion that successful IMGs were not passing with scores as high as the expected scores of USMGs cannot be easily dismissed, and this is one area of the structural realm of quality where there may be substantive differences between IMGs and USMGs.

Differences in IMG-USMG Examination Rigor

The third problem is that the ECFMG's examinations have not been as rigorous as those required of USMGs. As an example, consider the Foreign

Medical Graduate Examination in the Medical Sciences (FMGEMS), implemented in July 1984.[7] FMGEMS was a 2-day affair with day 1 concentrating on the basic medical sciences and day 2 concentrating on the clinical sciences. Like the original ECFMG examination, an English language test was also given. The examination was strictly comparable in question-to-question difficulty of parts 1 and 2 of the National Boards; FMGEMS differed in that it contained about one half the number of questions that parts 1 and 2 routinely contained and each of these parts of the National Boards lasted 2 days, not 1 day. Because a longer examination is more reliable in that there is a reduced probability that guessing will lead to a passing score, critics noted that a "double standard" existed. In a study of the first seven administrations of FMGEMS, Mick and Mou (1991) found a full pass rate (i.e., passing day 1, day 2, and English) of 23.0 percent for the entire population of IMGs sitting for any part of any of the seven FMGEMS administrations.[8]

Testing requirements changed dramatically when the ECFMG—in concert with the National Board of Medical Examiners (NBME)—adopted a policy, implemented in September 1989, that all IMGs were required to take exactly the same set of examinations (parts 1 and 2) as those offered by the NBME to obtain ECFMG certification. The FMGEMS continued to be offered until its last administration in July 1993. Selected examination results are shown in Table 1.[9]

In the most recent change phased in since 1991, the NBME and the ECFMG have established the USMLE as the *only* examination henceforth given to both IMGs and USMGs (Swanson et al. 1992).[10] This three-part series of examinations has also replaced the Federation Licensing Examination (FLEX), which had been administered since 1968 and was accepted by all state medical boards. Thus, the USMLE now links certification by the ECFMG more closely to the medical licensing process in the United States. Those IMGs (and USMGs) successfully certified in steps 1 and 2 will be eligible to take step 3 and thus will have completed all the examination requirements for licensure in all 54 licensing jurisdictions of the United States (Sutnick, Shafron, and Wilson 1992).

Of the examinees sitting for the USMLE step 1 administration, the percentage of "first takers" who passed was 50.2, and the percentage of "first takers" of step 2 who passed was 52.3. The comparative percentages of those passing for steps 1 and 2 during the second administration of the USMLE were 48.0 and 44.7, respectively (ECFMG 1994). The data for the entire year of 1994 are interesting: the overall (repeaters and first-time takers) pass rate of NBME-registered medical graduates (i.e., USMGs and Canadian medical graduates) was 86 percent. The comparable figure for ECFMG-registered IMGs was 44 percent (National Board of Medical Examiners 1995). Although the pass rate

TABLE 1 Selected Comparative Examination Pass Rates Taken by International Medical Graduates (IMGs)[a] (in percentages)

	Examination Pass Rate[b]	*Individual Pass Rate*[b]
Standard ECFMG[c] examination—1958 to 1983		
March 1958		
First takers	51.0	51.0
Repeaters	N.A.	N.A.
September 1965		
First takers	41.9	38.9
Repeaters	N.A.	64.0
1969-1982 aggregated		
First takers	N.A.	40.0
Repeaters	N.A.	58.0
Visa Qualifying Examination (VQE)—1977 to 1983		
1983 Basic and clinical science combined		
First takers	32.0	N.A.
Repeaters	14.0	N.A.
Foreign Medical Graduate Examination in the Medical Sciences (FMGEMS)—1984 to 1988		
July 1984-July 1987 aggregated		
Day 1 and day 2 combined		
First takers	N.A.	N.A.
Repeaters	N.A.	23.0
December 1987		
Day 1		
First takers	39.0	N.A.
Repeaters	26.0	N.A.
Day 2		
First takers	47.0	N.A.
Repeaters	26.0	N.A.
National Board of Medical Examiners (NBME)—1989 to 1993		
September 1991		
Part 1		
First takers	41.0	N.A.
Repeaters	38.0	N.A.
Part 2		
First takers	52.0	N.A.
Repeaters	39.0	N.A.

TABLE 1 continued

	Examination Pass Rate[b]	*Individual Pass Rate*[b]
United States Medical Licensing Examination (USMLE)—1994 to present		
June 1995—September 1995 aggregated		
Step 1		
First takers	55.0	N.A.
Repeaters	29.0	N.A.
Step 2		
First takers	56.0	N.A.
Repeaters	40.0	N.A.
Step 3		
First takers	70.0	N.A.
Repeaters	31.0	N.A.

Source: American Medical Association (1982, 1984); Dublin and Oesterling (1987); Educational Commission for for Foreign Medical Graduates (1994); Margulies and Block (1969); Mick and Mou (1991); National Board of Medical Examiners, 1996.
Note: N.A. = not applicable.
a. Pass rates generally exclude performance on the English language portions of the ECFMG's examinations. English language pass rates are always high; for example, in the FMGEMS example in this table, the overall English pass rate rate was 56.4 percent.
b. Examination pass rates are based on the number of examinations taken at a given administration; individual pass rates are based on the pass rate per examinee, regardless of the number of times the person took the examination.
c. ECFMG = Educational Commission for Foreign Medical Graduates.

of USMGs and Canadians was nearly twice that of IMGs, the IMG figures were well within, if not superior to, the levels of performance of past IMG examinees (Table 1). Recent data on step 3 of the USMLE have shown that the overall pass rate for USMGs and Canadians was 94 percent and that for IMGs was 58 percent (National Board of Medical Examiners 1996).

Throughout the nearly 40-year period that the ECFMG has been conducting its screening and credentialing functions, its tests have become progressively more rigorous. The intriguing question is how it has been that over time IMGs have consistently passed within the same band of rates despite internal variation related to factors such as the country where the IMG's medical school was located[11] or whether the IMG was a foreign national or a U.S. citizen.[12] One would have expected an increasingly high failure rate over time as examination difficulty approached and finally equaled that given to USMGs, yet this is clearly not the case. The over time consistency of pass rates shown in Table 1 suggests that historical differences of rigor between examinations

taken by IMGs and USMGs may have been overstated. With the implementation of the USMLE, the issue has now become irrelevant for most newly arrived IMGs.

UNCERTIFIED IMGs PRACTICING MEDICINE

The second key argument about testing and certification of IMGs concerns the phenomenon of IMGs performing clinical work who have neither taken nor passed any of the ECFMG's various examinations. This point became public in the mid-1970s when Weiss and his research team (Weiss et al., "Foreign Medical Graduates," 1974; Kleinman, Brandt, and Weiss 1975) coined the term *medical underground* to underscore their contention that the United States was being inundated with IMGs in medical roles who had not passed the ECFMG examination.

In a study of 4,035 IMGs having taken the January 1973 examination in U.S.-located examination centers, 48 percent of the 3,935 respondents were working in the health field at the time of the examination, that is, were not ECFMG certified yet were working in the health sector (Weiss et al., "Foreign Medical Graduates," 1974). Telephone interviews of a sample of 850 revealed that 73 percent of the 513 who reported working in the health field were involved in direct patient care and 64 percent of these were employed in hospitals. Analyses of specific job duties revealed large numbers functioning independently and in unsupervised settings.

The existence of non-ECFMG-certified IMGs is probably a fact, although its dimensions and its qualification as a "problem" are matters of debate. Studies continue to show that uncertified IMGs, or certified IMGs unable to obtain a license or a clinical position, exist (Politzer, Yesalis, and Katzoff 1989). There is nothing irregular or unusual about a foreign-educated physician immigrating to the United States without ECFMG certification. In fact, Smith and Fowkes (1983), in a study of some 1,210 unlicensed IMGs in California (many of whom were ECFMG certified), found that two thirds of the IMGs studied came to the United States after restrictive immigration legislation of 1977 and were refugees from countries in political upheaval. Others were the spouse of someone else who had migrated, the spouse of a U.S. citizen, or in one of a number of other situations that were (and are) completely legal and ordinary. In general, the two important issues raised include how large the "reserve pool" of unemployed but certified IMGs is and whether they are working in clinical settings and performing clinical acts that only a fully licensed physician should be allowed to undertake. Both questions are difficult to answer definitively, and no current research exists on the topic. However, both the ECFMG and the Federation of State Medical Boards are cur-

rently concerned that a certain number of IMGs may be filling residency slots without having passed any certification examination because of loop holes in the "H" visa category, which does not require possession of an ECFMG certificate to be awarded. This is an area in need of further research.

STRUCTURAL MEASURES OTHER THAN EXAMINATIONS

Licensure is another structural measure of quality that relies, in part, on criteria other than test results. However, the problem with licensure as a measure of physician quality is that its attainment has varied by state, has been different for USMGs and IMGs within a state, and has been as much a function of political pressure and physician workforce competition (Stevens and Vermeulen 1972). Goldblatt et al. (1975), using a 1971 follow-up database of all 1963 U.S. interns and residents, reported an association between licensure status and visa status: 43 percent of IMGs on exchange-visitor visas were licensed; the figures for permanent residents and naturalized U.S. citizens were 76 percent and 89 percent, respectively. Therefore, the overall licensure rate of IMGs of 66 percent was a grossly misleading statistic to compare with the 93 percent licensure rate of USMGs. The appropriate comparison would be the 89 percent figure for naturalized citizens, yielding only a 4 percent difference.[13]

Historic fluctuations in state licensure requirements for licensure are well documented (Butter and Sweet 1977), and as the 1990s began, licensing procedures continued to be characterized by much variation. Most states had clear differences of endorsement for USMGs and IMGs. Next, there was the well-known requirement that IMGs pass a different series of examinations than USMGs to obtain licensure: only USMGs and Canadian medical graduates could take the NBME parts 1, 2, and 3. IMGs took the FLEX.[14] Finally, most states required that IMGs complete more years of graduate medical education (GME) than USMGs; only 18 states required an equal number of years of GME for both IMGs and USMGs. Unlike most other states, California, Florida, New York, Ohio, Texas, and Virginia had reasonably similar procedures for both IMGs and USMGs (U.S. General Accounting Office 1990). Overall, the General Accounting Office found that states had no uniform endorsement standards or requirements to determine quality and that disagreement existed across state licensing boards as to the amount and kind of documentation that were required to make such a determination.

As late as 1993 the situation was as follows: 48 boards required that all IMGs have passed the ECFMG examination or hold ECFMG certificates before sitting for the FLEX. Nine state licensing boards permitted IMGs to take the FLEX before they had graduate training, but even these states required (as did

all others) at least 1 year of training before a license was granted. Thirteen states required 2 years, and 27 states required 3 years of training in an accredited GME program before allowing a license to be conferred. This apparent simplicity is actually belied by the existence of state-by-state differences in the licensing process (American Medical Association 1993). California required the passing of the first part of FLEX before approval to undertake GME. An oral examination in general medicine was also required. Oregon required GME of not less than 3 years in not more than two specialties, in not more than two U.S. or Canadian hospitals. IMGs in North Carolina had to take a special examination prepared and given by the state itself. New Hampshire demanded proof that the IMG was committed to practice medicine in New Hampshire. Rhode Island placed a limit of five attempts at the FMGEMS. Because of this variation, the use of licensure as a measure of quality is problematic.

Specialty Board Certification

A "final stage" in the testing careers of many physicians is the preparation for, and taking of, a specialty board certification examination, the last structural examination measure discussed here. There are 23 specialty board societies in the United States, and each has its own procedures for certification. There are few comparative IMG-USMG studies across all specialties, but there are some. Mick and Rubino (1992) studied all ECFMG-certified IMGs between 1969 and 1982 and compared them to all USMGs who began their residency training during the same period. They found that the comparative proportions of board-certified medical graduates were 46.9 percent for IMGs and 74.0 percent for USMGs. However, data on the proportion of IMGs in a given specialty who are board certified are very difficult to obtain. The problem resides in not knowing the denominator, that is, the number of IMGs (or USMGs) practicing in the relevant specialties who are and are not board certified. The American Medical Association Physician Masterfile, for example, determines a physician's specialty through a self-report, and it is not always clear that people reporting themselves as pediatric allergists actually underwent a postdoctoral fellowship designed to qualify the physician to be "board eligible," that is, qualified to take the board examination.

The difficulty of correctly estimating board certification rates is underscored in studies by Levit, Sabshin, and Mueller (1974) and Levit and Holden (1978). They examined samples of USMGs (no IMGs), followed up their residency and postdoctoral training careers, and then looked at the results of their specialty board examinations. The fact that it took an average of 12 years following receipt of a medical degree before a USMG had entered the certifi-

cation process or was board certified underscores the methodological problem of knowing the certification rates of a given cohort. In short, special studies that track IMGs (and USMGs) are required to determine what is a valid denominator for each of the 23 specialty boards.

Internal Medicine

Certain specialties, such as internal medicine, do have comparable IMG-USMG data. The examination of the American Board of Internal Medicine (ABIM) is the most commonly taken certification examination. Benson, Meskauskas, and Grosso (1981) documented the now familiar IMG-USMG pattern: comparing all examinees between 1975 and 1980 having taken the ABIM certifying examination revealed pass rates of 79 percent to 82 percent for USMGs, 15 percent to 38 percent for U.S.-citizen IMGs, and 27 percent to 45 percent for all other IMGs having taken an examination for the first time. All study participants had completed the same general postdoctoral requirements. Work by others (Norcini et al. 1986) confirmed these differences, although Norcini, Shea, and Benson (1991), studying 37,000 test takers over a recent 6-year period, showed that the difference between IMGs and USMGs narrowed over time. Studies on board certification performance in internal medicine, then, not unlike other examination results, showed a consistently lower score for IMGs than for USMGs. Since non-board-certified physicians, be they IMGs or USMGs, can and do practice medicine, the score differences may be important.

Other Studies

A study that coincidentally examined IMG-USMG differences questioned whether the knowledge of practicing internists had deteriorated over time or was out-of-date (Ramsey et al. 1991). A small, apparently nonrandom sample of 289 practicing internists (who had been board certified from 5 to 15 years previously) from New York, New Jersey, and Pennsylvania agreed to take an 82-item multiple-choice examination based on questions from the 1988 ABIM certifying examination. There was a statistically significant negative relationship ($r = -0.30$) between test scores and number of years elapsed since initial certification. The average scores of the IMGs versus the USMGs on the original certifying examination were 467.1 and 523.8, respectively, a statistically significant difference ($p < 0.0001$); and the correct answers on the study examination were 46.7 percent and 55.9 percent, for IMGs and USMGs, respectively ($p < 0.0001$). A multiple regression analysis predicting study test scores revealed that graduation from a U.S. medical school produced a statistically

significant standardized regression coefficient of 2.95 ($p = 0.008$), controlling for practice setting, subspecialty practice, time since certification, ABIM certification score, and university-based residency program. However, three other variables had larger coefficients. The uncertain nature of the sample severely limits the generality of the results.

Knowledge of genetics concepts and facts was studied by Hofman et al. (1993), who mailed questionnaires to 1,795 primary care physicians and psychiatrists. USMGs had a mean knowledge score of 74.2 percent correct, whereas IMGs achieved 66.9 percent correct ($p < 0.001$). The researchers also found that IMGs had a statistically significantly lower total genetics knowledge score than did USMGs, controlling for year of medical school graduation, high versus low exposure to genetics in practice, likelihood that a practitioner would offer a genetics test before it had become standard practice, exposure to drug company information, medical school courses in genetics, and postgraduate exposure to genetics. All of these variables were statistically significant predictors, with the added variation explained by having been an IMG being 2.8 percent (25.2 percent of all variation was explained by the model). By contrast, year of graduation from medical school explained 11.3 percent of the total variation.

These two studies support the general finding from the research employing structural measures discussed above that differences between USMG and IMG levels of medical knowledge exist. The relevant question becomes whether these differences are meaningful, given their magnitude relative to other factors, and whether they are valid predictors of differences in quality of care provided clinically.

PROCESS MEASURES OF QUALITY

Process measures of quality improve on structural measures in that they test not only knowledge but also the application of knowledge and skills in clinical practice; they are therefore one step closer, causally and conceptually speaking, to quality of outcomes. Their limitations are a lack of universally accepted measures, expense and difficulty of implementation, and the resultant lack of widespread use, and, where in use, lack of large-sample studies comparing IMG and USMG performance.

PRE-1975: THE DACSO-HALBERSTAM STUDIES

A series of pre-1975 studies done by Halberstam and Dacso (1966a, 1966b); Dacso, Antler, and Rusk (1968); and Halberstam, Rusk, and Taylor (1970) are

mentioned only because they remain virtually the only of their type and therefore deserve note. The researchers used supervisors' ratings to compare IMGs and USMGs according to their knowledge of basic medical science and clinical medicine. Although the IMG ratings were generally lower than those of USMGs, they were regarded as "satisfactory." These studies remain virtually the only of their type, and although they supported an IMG-USMG quality differential, they suffered from a lack of "blind" evaluation by raters, self-selected participating hospitals, and small nonrandom samples.

THE RHEE STUDIES

In the mid- to late 1970s, methodologically improved studies emerged that assessed process aspects of quality between IMGs and USMGs. Rhee and his colleagues (Rhee 1976, 1977a, 1977b; Rhee et al. 1981) published a series of studies based on an initial sample of 506 physicians having made 3,316 hospital discharges in 16 diagnostic categories from 22 nonfederal hospitals in the state of Hawaii during 1968.

The key dependent variable measuring physician quality was the Physician Performance Index (PPI), a measure that examined performance "according to the level of physician's compliance to the medical norms in the provision of patient care in offices and hospitals" (Rhee 1976).[15] Controlling for selected variables, Rhee (1976) found no difference between graduates of U.S. medical schools with an emphasis on teaching and specialization, U.S. medical schools with an emphasis on practice, foreign medical schools in medically more advanced countries, foreign medical schools in medically less advanced countries, and unknown medical schools. The medical school background of each medical graduate was explored in more detail in another article (Rhee 1977b), and the results were similar, except when specialists (IMGs and USMGs) practiced outside their areas of specialization. The latter practiced below-average medicine as measured by the PPI, a finding reinforced by yet another publication (Rhee et al., 1981), which revealed that general practitioners working outside the general practice domain practiced a quality of medicine below that of specialists.

Rhee hypothesized that forces subsequent to medical education such as amount of experience, features of residency training, and practice setting would affect performance on process measures (Rhee 1977a). Using the same data set described above, Rhee demonstrated the "present work environment" had more influence on quality of care as measured by the PPI than physicians' formal medical training, and that this influence was more pronounced on those physicians with less training and less pronounced on those with more training.

Rhee et al. (1986), again using the PPI, studied a random sample of 14,203 patient episodes by 1,156 physicians in ambulatory care settings drawn from discharge lists in a midwestern state. There was no support for the hypothesis of IMG-USMG differences: IMGs provided equal care to the USMGs, and sometimes the IMGs provided even marginally better care than USMGs.

THE SAYWELL-STUDNICKI STUDIES

Saywell and Studnicki (1976), in a cross-sectional study, used two medical audit procedures to evaluate the performance of care provided by attending physicians and house-staff physicians in eight diagnostic categories at 22 hospitals in Maryland and western Pennsylvania. A total of 6,980 medical records were abstracted from eight diagnostic categories for 1,321 physicians, of whom 985 were USMGs and 331 IMGs. A possible self-selection bias could have been present because participation of hospitals was voluntary: the 22 responding hospitals were among 42 asked to participate, although no differences in bed size, admissions, occupancy rate, personnel per bed, medical school affiliation, and urban-rural distinction existed between the two groups.

Saywell et al. (1979) used two types of inpatient hospital audits: the Payne Process Audit and the JCAH [Joint Commission for the Accrediation of Hospitals] Performance Evaluation Program (PEP) Audit.[16] The results indicated that although there was evidence of a strong interaction between hospital and type of physician for many of the diagnoses, there was no significant overall difference in performance between USMG and IMG attending physicians or residents (Saywell et al. 1980, 1983). The largest and most consistent differences in physician performance were associated with hospital characteristics, not physician characteristics. Saywell et al. ("An Examination," 1981) and Saywell et al. ("A Comparision," 1981) found similar results when examining inappropriate utilization of hospital resources.

In short, hospital characteristics appeared to be the stronger set of variables in explaining differences in their quality measures. However, these studies, based on one cross-sectional study group, were unable to untangle the causal issue of whether "bad" hospitals produced "bad physicians" or whether "bad physicians" (including IMGs) were selected into hospitals and made the hospitals "bad" (Baskin 1980). Unfortunately, this question remains unanswered.

THE CLINICAL SKILLS ASSESSMENT (CSA)

The ECFMG has announced that a clinically oriented examination, simulating care delivery, called the Clinical Skills Assessment (CSA), will be re-

quired of IMGs no sooner than mid-1996 (ECFMG 1994). The examination consists of (1) clinical encounters with standardized patients to assess history taking, physical examination and communication skills, (2) laser video-disk scenarios to assess identification and interpretation of diagnostic procedures, (3) written clinical vignettes to assess diagnosis and management skills, and (4) spoken English evaluation.

Conn (1986) demonstrated the feasibility of administering the CSA, and Conn and Cody (1989) compared 635 IMGs and a reference sample of 123 USMGs and concluded that the clinical skills of 28 percent of the IMGs who took the CSA were found to be inadequate when compared with those of graduates of U.S. schools. In 1993, Sutnick et al. (1993) reported that the CSA possessed high reliability and improved the predictability of residents' performance in the hospital as against then current ECFMG examinations. Still, there are no large-sample studies that compared IMG and USMG performance on the CSA, and its utility as a measure of quality is limited, at least at present.

OTHER STUDIES

Other studies using standardized patients and kindred approaches have yielded some information regarding IMG-USMG quality of care. Schnabl, Hassard, and Kopelow (1991) found that reports by standardized patients of physicians' interpersonal skills were about equally positive regardless of USMG or IMG status. Another approach, the In-Training Examination (ITE), a test intended to determine the knowledge base of second-year residents in general medicine, relates to ambulatory care and employs case vignettes requiring clinical decision making. Results of the ITE reported by Garibaldi et al. (1994) showed that during the period 1988-1993, the average scores of IMGs were always below those of USMGs for each level of residency training (first, second, and third years). For example, the average score of first-year internal medicine USMG and IMG residents in 1988 was 62.0 and 59.8, respectively. Comparable scores in 1993 for USMG and IMG residents were 57.0 and 55.9, respectively. Nevertheless, the 1995 ITE results showed that the IMG average score was higher than that of USMGs for each of the three levels of training (e.g., for first-year USMG and IMG residents, 60.0 and 61.9, respectively) (Waxman, Garibaldi, and Subhiyah 1996).

Another study used an "objective structured clinical examination," composed of nine physical diagnosis and test interpretation stations using paid volunteers with known diagnoses and physical findings (Dupras and Li 1995). Of the 51 second-year internal medicine residents at the Mayo Clinic, 10 were IMGs. The average scores of IMGs were no different than those of the USMGs (56 versus 57, respectively).

Finally, there is the question of language. Part and Markert (1993) found that the interview language skill of 46 IMGs in a first-year internal medicine residency program did not correlate with the mean ratings of residency performances. The lack of significance may have been due to the small sample, because other studies (cf. George, Young, and Metz 1989) have suggested that language skills of IMGs are predictive of performance. There is research and commentary suggesting that language ability plays a role in quality of medical services (Woloshin et al., 1995), but, interestingly, there is no consensus literature on the IMG-USMG language issue.

This group of "other process studies" presents contradictory IMG-USMG findings, with the major problem being the existence of but a small number of studies. Clearly, more work needs to be done in this area.

OUTCOME STUDIES

GENERAL

Outcome studies have almost never been done with the explicit purpose of comparing IMGs and USMGs. Instead, when present, the IMG-USMG distinction has been used as one of a number of control variables. Notwithstanding this lack of focus, the few studies that have been done are revealing.

Burns and Wholey (1991) studied the correlation of patient, hospital, and physician characteristics on length of stay (LOS) and mortality rates for 27 diagnosis-related groups (DRGs), using an unnamed western state's entire population of 1988 discharges. The DRGs examined included 11 medical and 5 surgical conditions. The IMG variable in this study was the percentage of treating physicians in each hospital who were IMGs. Results suggested that both hospital and physician characteristics were important predictors of both outcome measures, but the authors concluded that "graduation from a foreign medical school, board certification, and general/family practice specialty exert little influence on length of stay" (p. 259). With regard to mortality, the authors reported that "board certification and graduation from a foreign medical school (IMG) exert little influence on the odds of patient mortality" (p. 265).

Complications from 1,302 patients having undergone carotid endarterectomies in 1981 in three large geographic areas were studied by Brook et al. (1990). Of these patients, 11.3 percent had postoperative stroke or heart attack or died within 30 days of the operation. Using logistic regressions with multiple control variables, including clinical variables and adjustment for patient age, race, income, and gender as well as provider volume and other provider- and hospital-based controls, the researchers found that if the sur-

geon was an IMG (but not one from Canada or Western Europe), the average complication or death rate rose to 19.6 percent ($p < 0.05$). This was the only significant covariate other than illness severity and comorbidity although the authors cautioned that a relatively small number of variables were examined and that "the data are 10 years old, and the literature does not consistently indicate that foreign medical graduates provide worse care" (p. 752).

Tussing and Wojtowycz (1993), using a retrospective review of 65,784 obstetrical deliveries by 1,740 different physicians in New York State (excluding New York City) in 1986, examined the cesarean section rate in this large sample. A cross-sectional study, it employed various databases to develop measures of physician characteristics, fetal conditions (dystocia and fetal distress), maternal history of cesarean section, maternal characteristics, and delivery organization characteristics. The overall cesarean rate for the sample was 27.8 percent (compared with 24.1 percent nationally), and the rate for IMGs was 29.4 percent, statistically significant at the 0.01 level.[17] Other variables with significantly higher rates were medical graduates who had graduated by 1977 (28.1 percent) and deliveries by board-certified obstetrician-gynecologists (28.6 percent).

MEDICAL MALPRACTICE

Another outcome measure of quality is medical malpractice. Notwithstanding problems with malpractice as a measure of quality, there are at least some comparative IMG-USMG studies available.[18] Two Michigan-based studies suggest no IMG-USMG differences. Richards (1980) examined the 906 malpractice claims between 1974 and 1979 involving Michigan physicians and found that 17 percent of IMGs had been sued compared to the fact that 26 percent of all Michigan physicians were IMGs. Next, Zannoth (1981) examined 443 suits that had been litigated between 1980 and 1981, but that had been initiated from as early as 1970. For each year of the initiation of the suit, the percentage of IMGs involved was computed, and the conclusion was that although the percentage of IMGs involved had risen over the period, it did not equal the percentage of IMGs in practice. In neither study were any variables of relevance controlled.

An early study on comparative malpractice was prepared by the American Medical Association Office of the General Counsel (American Medical Association 1972). In a "recent 12-month period," 107 court decisions involving 125 physicians in medical liability cases were studied. Of these 125 physicians, 108 were born in the United States, 3 in Canada, and 14 elsewhere, but only 8 (6.4 percent) of the defendants were graduates of medical schools outside the United States and Canada. Furthermore, of the 125 defendants, 58 experienced

an "unsuccessful" outcome of the litigation, and of these 4 (6.9 percent) were IMGs. Because the proportion of IMGs in the United States hovered between 15 percent to 20 percent in the early 1970s, it can be seen that IMGs were not being sued at anything close to their representation in the population.

Rovit's study (1978) of neurological surgeons compared 211 IMGs, 44 of whom were graduates of Canadian medical schools, and 1,210 USMGs (1,421 total). In a nonrandom sample study of 2,200 neurosurgeons with a poor response rate, Rovit showed no preponderance of suits against the IMGs compared with USMGs.

The U.S. General Accounting Office (1987) analyzed data from a random sample of 31,395 malpractice claims closed in 1984 by 25 insurers.[19] Of the 71,930 physicians in the study, 16,780 were IMGs (23.3 percent). In 1981, apparently the latest data available at that time, 110,542 (23 percent) of all physicians in the U.S. workforce were IMGs, and the General Accounting Office concluded that IMGs were not more likely to be involved in malpractice claims than were USMGs.

Sloan et al. (1989) studied the universe of all physicians practicing in Florida at least 3 years between 1975 and 1980, during which there were 5,934 paid claims against physicians. "No payment," "low payment," and "high-very high" payment were used as proxies for (decreasing) quality, and their use of multinomial logit analysis makes this the most ambitious effort to control for confounding factors that might affect an IMG-USMG comparison. The key result for our purpose was that physicians with degrees from less economically developed countries had about the same claims experience as other groups of physicians, including those with U.S. and Canadian degrees.[20]

Finally, Schwartz and Mendelson (1989) examined demographic data on some 920 physicians who, between 1983 and 1987, lost their medical malpractice insurance and applied to a "surplus-lines" company that would insure anyone. Of the sample, 21.1 percent of surplus-lines physicians were IMGs, whereas the age- and specialty-adjusted IMG population proportion was 22.6 percent, a nonsignificant difference. In one subspecialty—plastic surgery—there was a significant overrepresentation of IMGs: 42 percent in the surplus-lines sample versus 18 percent IMGs in the general population in this sample ($p < 0.05$).

SUMMARY

Table 2 summarizes eight of the methodologically strongest studies on comparative IMG-USMG quality. These studies were included in the table because they met most of the following methodological criteria: sample size was large or random, or an entire study population was used; control or

TABLE 2 Selected Published Studies of IMG-USMG[a] Quality of Care

Author(s)	*Date*	*Sample*	*Type of Study*	*Measure of Quality*	*Statistical Procedure*	*General Findings*	*Comment*
Structure							
Hoffman et al.	1993	1,140 nongeneticist physicians responding to a genetics test	Cross-sectional study of test scores to determine genetics knowledge	Structure; percentage questions answered correctly	Bivariate comparisons; OLS[b] multiple regression controls	Statistically significant difference between IMGs and USMGs: 66.9 percent versus 74.2 percent, respectively, answering correctly; regression shows significant IMG effect adding 2.8 percent explained variation	Most important factor in explaining test scores was year of medical school graduation (11.3 percent of variation). Nonresponse rate of study = 35 percent, limiting generality; most frequent problem was in understanding probabilities
Process							
Rhee (1977b)	1976	454 physicians; 2,517 patient discharges in 15 diagnostic categories in Hawaii	Retrospective review of medical records; cross-sectional study with control variables	Process; norm process compliance measure by Payne and Lyons (1972): Physician Performance Index (PPI)	Multiple Classification Analysis (MCA)	Medical school of graduation not related to PPI scores after controlling for other structural variables	Related studies using the same sample, PPI scores, and statistical technique show greater importance of amount of experience, type of hospital, and other organizational variables
Process							
Saywell, Studnicki, et al.	1979 through 1983	Self-selected sample 22 Maryland and Pennsylvania hospitals: 8 discharge groups of 6,980 patients of 1,321 attending physicians	Retrospective review of medical records; cross-sectional study with few control variables	Process; PPI scale of Payne et al., and JCAH PEP[c] compliance scales	Two-way analysis of variance (ANOVA)	IMG-USMG distinction not statistically significant; hospital effect significant, and interaction effect between medical graduate type and hospital also significant	Reports in the series study residents and attendings and conclude that IMG-USMG distinction is "not very useful" because hospital and patient-level variables emerged as significant. Quality of hospital was a salient factor

(continued)

TABLE 2 continued

Author(s)	*Date*	*Sample*	*Type of Study*	*Measure of Quality*	*Statistical Procedure*	*General Findings*	*Comment*
Outcome							
Tussing and Wojtowycz	1993	65,784 births in 1986 in New York State by 1,740 different physicians, from New York State Live Birth File. Entire population of live births	Data abstracted from Live Birth File and matched with hospital discharge data. Cross-sectional study with control variables	Outcome; cesarean section rates	Probit multiple regression	IMGs had significantly higher cesarean section rates than USMGs after controlling for a variety of other physician traits, patient traits, medical indications, and other variables	Study design does not permit testing of causal factors leading to higher IMG rates; lower levels of skill, language problems, and defensive medicine are several proposed hypotheses
Outcome							
Schwartz and Mendelson	1989	920 physicians applying for "surplus lines" malpractice insurance (for physicians with bad claims experience or other problems); time period 1983-1987	Retrospective/ cross-sectional analysis; sample adjusted to represent distribution of characteristics of U.S. physician workforce	Outcome; percentage of sample IMG surplus line applicants	Adjustment of sample to reflect population; bivariate distribution	Percentage IMGs in adjusted sample 21.2 percent versus 22.6 percent IMGs in the U.S. physician workforce; no statistically significant difference	Plastic surgery overrepresented in surplus lines applications; IMGs more likely to be in plastic surgery, thus more IMGs in this specialty seeking surplus lines; but general findings controlling for specialty is no IMG-USMG difference
Outcome							
Sloan et al.	1989	Florida physicians involved in malpractice claims between 1975 and 1988	Retrospective/ cross-sectional study with controls	Outcome; dollar level of payments	Multinominal logit analysis; controls	IMGs had no different claims payment experience than did non-IMGs in medical specialties, obstetrics-anesthesiology, and surgical specialties, except *fewer* high-payment settlements compared to USMGs	Generally, study found no connection of claims settlement experience with medical school of graduation or prestige of medical school

Outcome							
Burns and Wholey	1991	Unnamed western state's (probably Arizona) population of discharges for 11 medical and 5 surgical diagnosis-related groups (DRGs) for all of 1988	Retrospective review of medical records to determine length of stay and mortality rates; cross-sectional study with controls	Outcome; adjusted mortality rates and length of stay	Multiple regression and analysis of covariance (ANCOVA)	Medical school of graduation (IMG-USMG distinction) exert little influence on length of stay or mortality rates	Study focused on volume-outcome effects, controlling for large number of other variables, including IMG-USMG, but that turned out to be generally insignificant expecially in relation to other controls
Outcome							
Brook et al.	1990	1,302 patients having caritod endarterectomy in 3 "large" geographic areas in 1981	Retrospective review of medical records to determine postoperative outcomes	Outcome; postoperative stroke, heart attack, 30-day postdischarge mortality	Multiple regression (logistic)	11.3 percent of patients had complications or died; IMG status other than Canadian or Western European had significantly higher complication and mortality outcome	Illness severity and comorbidity also had significant impact on complications and mortality. No other physician or hospital controls had significant impacts

a. IMG-USMG = international medical graduate–U.S. medical graduate.

b. OLS = ordinary least squares.

c. JCAH PEP = Joint Commission for the Accreditation of Hospitals Performance Evaluation Program.

quasi-control groups were present; control variables were used; multiple regression techniques were used; statistical tests were employed; some attempt at standardization of quality variables was made. A general summary of structure-, process-, and outcome-oriented studies follows.

Structural Measures

The review of examination results is the one area where an argument may exist that supports lower IMG quality. ECFMG examination scores showed a pattern of lower IMG test performance. A caveat, however, is that it is unclear what the testing reveals. The tendency of IMGs to achieve roughly equal pass rates over the years on progressively longer, and presumably more difficult, ECFMG certification examinations suggests that something else is occurring than just a quality screen function. One hypothesis is that IMGs have done and do only that which they need to do to get certified, that is, to pass the examination. Passing is sufficient; the score is secondary.

One major disappointment was that we found no methodologically rigorous studies that linked examination score performance with process or outcome measures of quality. In short, little progress seems to have been made since the mid-1970s' challenge issued by Williams and Brook (1975). With the implementation of the CSA, it may be possible to correlate examination scores with simulated process measures.[21] On the other hand, whether the CSA simulation will predict actual performance during the course of residency training and beyond is as yet an unanswered question.

The existence of a "pool" of uncertified (by the ECFMG) IMGs in the United States may pose a hazard for patient care if it can be reliably demonstrated that these persons are, in fact, delivering medical services. However, the possibility that this phenomenon may exist is unrelated to the quality issue of IMGs who are ECFMG certified and who are in residency training programs; who are licensed to practice medicine; who do, in fact, practice medicine in every sort of practice setting in which USMGs are also found. The "medical underground" issue requires specific study and policy recommendations aimed to eliminate the practice in much the same way as there are laws and processes intended to eliminate medical quackery.

Studies of licensure should be undertaken. Apart from some FLEX data based on self-selected samples, this arena of structural measures has yielded little information on IMGs versus USMGs. There is a promising research effort possible as the results of the USMLE begin to be disseminated, and this is probably where future work should concentrate.

There is a clear need for more work on the implications of quality for those IMGs (and USMGs) who are not board certified. Although some specialty areas like internal medicine have clear data on board examination performance, others do not appear to have much published information on the subject. Furthermore, linking this structural measure with process and outcome measures and distinguishing between IMGs and USMGs is still a research frontier, echoing the conclusion of Williams and Brook (1975). It is still uncertain whether board certification matters despite the belief in many health care organizations that it is a valid proxy for quality.

In sum, our review reveals virtually no progress during the past 20 years in producing useful information on the meaning of IMG-USMG structural measures of quality such as certification examination scores, licensure, and board certification. Work that examines other structural measures is scarce. A paper-and-pencil test indicates that genetic knowledge appears to be somewhat lower among IMGs than USMGs. In short, while there are examination differences and a few other indicators of structural quality differences, there is no indication of what it means with respect to process and outcome measures. We admit that there may be a connection between structural measures of quality and process (or outcome) measures; Donabedian (1980) and others have long argued that an important and substantive connection exists. However, at least in the realm of IMG-USMG quality of care studies, the connection has yet to be demonstrated.

Process Measures

There has been little progress in the study of process measures since the "classical" studies of the research teams of Rhee and of Saywell and Studnicki. Where more recent studies exist, they have used the IMG-USMG distinction as a control variable. This means that no theoretical or substantive argument has been used to explain why any differences should exist, which begs the question in which one is interested. That is, studies in both the process (and outcome) domain implicitly use the IMG-USMG distinction as a structural proxy in and of itself for quality differences. Although there is nothing inherently wrong with this approach, there is no gain in it for elucidating why IMG-USMG quality might differ if it does differ. Where an effort has been made in this area, such as in the Rhee and Saywell-Studnicki research, no differences have been found between IMGs and USMGs when statistical controls were used. With the widespread use of clinical pathways and guidelines, there should be abundant opportunity to examine potential IMG-USMG differences. Furthermore, Peer Review Organization (PRO) data should also provide a rich source of comparative work.

Outcome Measures

Outcome research suffers from the same problem just discussed: a lack of the IMG-USMG comparison being at the core of the research. Although Tussing and Wojtowycz (1993) found a slightly higher, statistically significant, IMG cesarean section rate in a sample of New York State physicians, Burns and Wholey (1991) found little influence of being an IMG on hospital LOS and hospital mortality rates. However, Brook et al. (1990) found a nearly twofold increase in complications or mortality rate after carotid endarterectomy for IMGs who were neither Western European nor Canadian trained. Thus, two studies oriented toward procedures (cesarean section and carotid endarterectomy) suggest an IMG-USMG difference, and perhaps a useful research focus would be procedure-oriented studies. Nonetheless, the malpractice literature revealed no propensity of IMGs to be sued more often than USMGs.

In short, there is contradictory evidence that IMGs have different outcomes in medical practices than do USMGs, and the lack of an extensive literature hinders making too much of the lower IMG quality levels that have been found. It is important to add that few of these studies, as well as the process studies, examine any of the structural measures discussed earlier.

An exception is board certification. Tussing and Wojtowycz (1993) found that board-certified physicians in their sample had a significantly higher-than-average cesarean section rate (28.6 percent). Burns and Wholey (1991) found that board certification had little influence on LOS or mortality. Sloan et al. (1989) found that board-certified physicians did not generally have a different malpractice claims experience than non-board-certified physicians. These three studies are instructive, however, because they underscore our contention that the IMG-USMG distinction is not at the center of the research effort. Had it been, the IMG-USMG variables and the board certification variables would have been entered into the analysis as interaction terms, which they were not. Although this is not the only way one could examine the IMG quality issue more closely, it is surely an important way.

CONCLUSION

The IMG-USMG quality issue is not yet resolved scientifically. If the question is whether various structural measures (e.g., tests) show a consistent pattern of lower IMG performance, the answer is "yes." However, if one asks if we know whether IMG quality of care is lower than that of USMGs, the answer is "we don't know." It cannot be said that the reviewed literature proves no IMG-USMG quality differences or that as research improves and accumulates, distinctions might not be found. Still, the reviewed research

yields great ambiguity, and this, we believe, suggests two general conclusions.

First, our inability to provide a clear answer to the latter question may stem from the difficulty that researchers of quality of care have had and continue to have in measuring quality and in linking structure with processes and outcomes of care. Our objective was not to draw the enterprise of quality-of-care assessment into question, but it would be foolish not to consider that the failure to produce clearer evidence about IMG-USMG quality inheres in the complexities of measuring the phenomenon in the first place. One of these complexities is the possibility that the notion of quality may be linked to the national and historical context of a given medical care system. This suggests that quality itself is contingent and fluid, influenced as much by social and cultural forces as by sheer technical and procedural considerations (Stevens 1995). What could be more difficult than attempting to standardize perceptions of quality as represented, on the one hand, by USMGs and, on the other, by the dozens of nationalities and the hundreds of medical schools represented by America's IMGs?

Second, we conclude that great caution must be taken in asserting that lower IMG quality is a reason for limiting the arrival, training, and practice of IMGs in the United States. There may be other cogent reasons for limiting the IMG presence in the U.S. health care system, but evoking a quality differential does not seem justified at the present state of our knowledge. Therefore, we urge objectivity in the debate about IMGs and call for more research specifically on the quality issue. In a time of profound upheaval in the organization and delivery of care by physicians in this country, it is imperative that *opinion* about quality be reduced to a minimum.

NOTES

1. Canadian medical graduates are included among U.S. medical graduates. This is because of the practice of the American Medical Association and derivative organizations such as the Liaison Committee on Medical Education (LCME), which consider Canadian medical schools equivalent to U.S. medical schools. Canadian medical graduates have therefore entered U.S. residency positions in the same manner as U.S. medical graduates (USMGs), bypassing the special procedures for all other international medical graduates (IMGs).
2. The number of IMG residents in 1990 and in 1994 was 14,914 and 23,499, respectively, a 57.6 percent increase; the respective figures for USMG residents were 67,988, 74,333, and 9.3 percent (Graduate Medical Education 1994). The total number of IMGs in the U.S. physician workforce was estimated to be 139,086 in 1992, 23.0 percent of all physicians. Follow-up studies of IMG residents reveal that two thirds to 90 percent of any given sample will have remained in the United States (see Mick and Pfahler 1995, chap. 2, for a review).

3. Offsetting these reasons for reducing IMGs is the often-cited "gap filling" hypothesis: IMGs tend to locate in places, practice in specialties and employment settings, and serve populations that would be underserved if only USMGs were in the U.S. health care system. Recent evidence that supports this general line of thought has been reported by Mick and Lee (1996). This explains, in part, the ambiguity and contradictions that have been present in the IMG policy debate.
4. We have emphasized the notion of quality rather than that of competence because the latter term means having requisite abilities or skills that may lead to high scores on measures of structure, process, and outcome. The existence of competence is therefore to be inferred from a person's position on measures of quality.
5. We found no pattern among studies excluded on methodological grounds to support a conclusion of lower IMG quality. Interested readers may contact the author for a complete list of the articles reviewed for this report.
6. The Educational Commission for Foreign Medical Graduates (ECFMG) itself has never taken a position on whether IMGs are more or less competent than USMGs, thus adhering to a clear policy of being neutral in the debate.
7. The U.S. Department of Health and Human Services considered the Foreign Medical Graduate Examination in the Medical Sciences (FMGEMS) as equivalent to the NBME examination parts 1 and 2; thus, under the 1976 provisions of PL 94-484, foreign national IMGs were able—upon certification by the ECFMG (passing FMGEMS was the major but not the only step for certification)—to obtain a visa to enter the United States.
8. These figures are confounded somewhat by the fact that IMGs could take the parts of FMGEMS serially, and that at the time of the study, a large proportion (48.6 percent) of IMG examinees had passed at least one portion of FMGEMS, many of whom (11,078 persons) had not attempted the other portion. This produces the so-called right-censoring problem, in which the magnitude of the phenomenon of interest, here pass rates, is artificially lower than what it would be if more time had passed before performing the study. Thus, if the 11,078 IMGs had been removed from the original denominator of that used to compute the full pass rate, the rate would have risen to 29.8 percent.
9. Because the data are not published, it is extremely difficult to determine the pass rates on a single type of ECFMG examination. In the early 1990s, it was possible for an IMG to pass any of the following combination of examinations to secure this part of ECFMG certification: passing FMGEMS day 1 and day 2, passing NBME part 1 and part 2, passing FMGEMS day 1 and NBME part 2, passing NBME part 1 and FMGEMS day 2 (Bergen 1990). Such flexibility was allowed to assist IMGs in coping with the numerous changes in examination structure that had transpired since 1984. However, the number of IMGs electing these various routes is not something that is published, and thus precision in pass rate data is not as high as it might otherwise be and as it otherwise was during the long period when the ECFMG gave its standard 1-day examination beginning in 1958.
10. A passing performance on any of the ECFMG's previous examinations continues to be valid for ECFMG certification.

11. As an example, initial pass rates between 1969 and 1982 on the old 1-day standard ECFMG examination varied from 17.2 percent for Japanese IMGs to 97.1 percent for Australian IMGs, with countries like the Federal Republic of Germany (52.2 percent) or Taiwan (33.2 percent) in between (Dublin and Oesterling 1987).
12. For instance, overall pass rates for the first seven FMGEMS examinations for FNFMGs for whom English was their native language was 43.3 percent; for USFMGs for whom their native language was English, 32.6 percent (Mick and Mou 1991).
13. Interestingly, the study also showed that foreigners educated in U.S. medical schools experienced similar variation in licensure rates: exchange visitors, 64 percent; permanent residents, 89 percent; naturalized citizens, 91 percent. Because the training was in the same schools as U.S. citizens and because the admission requirements were presumably the same for both Americans and foreigners, it would be hard to argue that much else, other than individual state idiosyncrasies, was the principal factor affecting licensure rates.
14. The Federation Licensing Examination (FLEX) did allow some limited comparisons of IMG-USMG performance, but because most USMGs took the National Board examinations, the characteristics of USMGs who took the FLEX were not known. Between June 1968 and December 1972, some 27,138 candidates sat for the FLEX, and 75 percent of these were IMGs. The overall USMG failure rate was 14.8 percent, and that for IMGs was 50.4 percent (Merchant 1973).
15. The Physician Performance Index (PPI), developed by Payne and Lyons (1972)—based on an evaluation of medical norms within diagnoses by panels of physicians—encompassed history taking, physical examination, laboratory work, and therapy prescribed. A PPI score was computed for each episode, based on a retrospective review of patient records, and all PPI scores were aggregated and standardized. Each physician received a PPI score as the mean of all the cases he or she cared for.
16. The Payne system consists of three parts, one of which was the PPI used in the Rhee studies. The Performance Evaluation Program (PEP) Audit consisted of categories for justification of the diagnosis, outcome, and quality indicators.
17. Using the same data set, Tussing and Wojtowycz (1997) found that IMGs were significantly more likely to use electronic fetal monitoring (EFM) than the overall population, but so did more recently graduated physicians, board-certified physicians, physicians without professorial appointments, and physicians in closed-panel HMOs. Finally, whether high use of EFM is a marker for poor quality is uncertain, although it might be an indication of "defensive medicine," that is, a technique taken to avoid malpractice exposure.
18. Problems associated with this measure were discussed by Imperato (1986): (1) the lack of multivariate techniques to control for potentially confounding variables such as physician's specialty and the severity of the condition being treated; (2) the willingness of the population to sue a physician especially in terms of the demography of the population, for example, a recent immigrant population may be less

willing to sue a physician; (3) the need to differentiate between the number of suits brought against physicians versus the number that were successful.

19. A "closed claim" is one for which (1) a claim for damages is not made, (2) the plaintiff drops the claim, (3) the insurer and plaintiff agree to a financial settlement, (4) a court renders a verdict, or (5) a settlement is reached through arbitration.
20. Of nine groups of comparisons across three specialty groupings, only one major difference in claims experience by country of graduation was revealed, and in the opposite direction than expected: in the surgical specialty group, those from developed, non-English-speaking countries were significantly *more likely* to experience "low payments" than others, including graduates of the United States. The Harvard Medical Practice Study (Brennan et al. 1991) is another study of malpractice that controlled for some confounding factors, in investigating adverse events and claims in a sample of New York State hospitals. It reported only differences by clinical specialty, however, and did not differentiate between IMGs and USMGs.
21. Stillman et al. (1992) have shown a weak but statistically significant correlation between scores on the FMGEMS day 2 (clinical science) examination and the notes on the Clinical Skills Assessment (CSA) standardized patient encounter test ($r = 0.14$, $p < 0.05$). Further practice with, and refinement of, the CSA may lead to stronger correlation with structural measures. However, such a weak association may indicate that structural examination measures may, in the long run, be largely irrelevant indicators of quality.

REFERENCES

American Medical Association, Office of the General Counsel. 1972. Profile of Medical Defendants. *Journal of the American Medical Association* 219 (12): 1687-1688.

American Medical Association. 1982. *U.S. Medical Licensure Statistics 1980-1981 and Licensure Requirements 1982*. Chicago: American Medical Association.

———. 1984. *U.S. Medical Licensure Statistics 1982-1983 and Licensure Requirements 1984*. Chicago: American Medical Association.

———. Department of Physician Licensure and Career Resources. 1993. *U.S. Medical Licensure Statistics and Current Licensure Requirements*. Chicago: American Medical Association.

Association of American Medical Colleges. 1997. Consensus Statement on Physician Workforce. Internet Address: http://www.aamc.org. (March 3).

Baskin, D. 1980. FMG-USMG Physician Performance. *American Journal of Public Health* 70 (7): 743-744.

Benson, J. A., J. A. Meskauskas, and L. J. Grosso. 1981. Performance of U.S. Citizen-Foreign Medical Graduates on Certifying Examinations in Internal Medicine. *American Journal of Medicine* 71 (2): 270-273.

Bergen, S. S. 1990. ECFMG Certification. *Federation Bulletin* 77 (9): 260-268.

Brennan, T. A., L. L. Leape, N. M. Laird, L. Hebert, A. R. Localio, A. G. Lawthers, J. P. Newhouse, P. C. Weiler, and H. H. Hiatt. 1991. Incidence of Adverse Events in

Hospitalized Patients. Results of the Harvard Medical Practice Study I. *New England Journal of Medicine* 324 (6): 370-376.

Brook, R. H., R. E. Park, M. R. Chassin, J. Kosecoff, J. Keesey, and D. H. Soloman. 1990. Carotid Endarterectomy for Elderly Patients: Predicting Complications. *Annals of Internal Medicine* 113 (19): 747-753.

Burns, L. R., and D. R. Wholey. 1991. The Effects of Patient, Hospital, and Physician Characteristics on Length of Stay and Mortality. *Medical Care* 29 (3): 251-271.

Butter, I., and R. G. Sweet. 1977. Licensure of Foreign Medical Graduates: An Historical Perspective. *Milbank Memorial Fund Quarterly* 55 (2): 315-346.

Congressional Budget Office. 1995. *Medicare and Graduate Medical Education.* Washington, DC: U.S. Government Printing Office.

Conn, H. L., Jr. 1986. Assessing the Clinical Skills of Foreign Medical Graduates. *Journal of Medical Education* 61 (11): 863-871.

Conn, H. L., Jr., and R. P. Cody. 1989. Results of the Second Clinical Skills Assessment Examination of the ECFMG. *Academic Medicine* 64 (8): 448-453.

Council on Graduate Medical Education. 1995a. *Sixth Report Managed Health Care: Implications for the Physician Workforce and Medical Education.* Washington, DC: U.S. Department of Health and Human Services, Health Resources and Services Administration.

Council on Graduate Medical Education. 1995b. *7th Report: COGME 1995 Physician Workforce Funding, Recommendations for Department of Health and Human Services' Programs.* Washington, DC: U.S. Department of Health and Human Services, Health Resources and Services Administration.

Dacso, M. M., L. Antler, and H. A. Rusk. 1968. The Foreign Medical Resident Training in the United States. *Annals of Internal Medicine* 68 (5): 1105-1113.

Derbyshire, R. C. 1975. Warm Bodies in White Coats. *Journal of the American Medical Association* 232 (10): 1034-1035.

Donabedian, A. 1980. *Explorations in Quality Assessment and Monitoring, Vol. I: The Definition of Quality and Approaches to Its Assessment.* Ann Arbor, MI: Health Administration Press.

Dublin, T., and S. F. Oesterling. 1987. *A Study of Candidates for ECFMG Certification, 1969-1982.* Philadelphia: Educational Commission for Foreign Medical Graduates.

Dupras, D. M., and J.T.C. Li. 1995. Use of an Objective Structured Clinical Examination to Determine Clinical Competence. *Academic Medicine* 70 (11): 1029-1034.

Educational Commission for Foreign Medical Graduates. 1969. *Annual Report, 1969.* Philadelphia: Educational Commission for Foreign Medical Graduates.

———. 1982. *Annual Report, 1981.* Philadelphia: Educational Commission for Foreign Medical Graduates.

———. 1985. *Annual Report, 1984.* Philadelphia: Educational Commission for Foreign Medical Graduates.

———. 1994. USMLE Examinee Performance. *Outlook* 14-15:8.

Garibaldi, R. A., M. C. Trontell, H. Waxman, J. H. Holbrook, D. T. Kanya, S. Khoshbin, J. Thompson, M. Casey, R. G. Subhiyah, and F. Davidoff. 1994. The In-Training Examination in Internal Medicine. *Annals of Internal Medicine* 121 (2): 117-123.

George, J. M., D. Young, and E. N. Metz. 1989. Evaluating Selected Internship Candidates and Their Subsequent Performances. *Academic Medicine* 64 (8): 480-482.

Goldblatt, A., L. W. Goodman, S. S. Mick, and R. A. Stevens. 1975. Licensure, Competence, and Manpower Distribution: A Follow-Up Study of Foreign Medical Graduates. *New England Journal of Medicine* 292 (3): 137-141.

Graduate Medical Education. 1994. *Journal of the American Medical Association* 272 (9): 725-733.

Halberstam, J. L., and M. M. Dacso. 1966a. A Cross-Cultural Evaluation of Graduate Training in Physical Medicine and Rehabilitation by Foreign and United States Residents. *Archives of Physical Medicine and Rehabilitation* 47 (9): 621-632.

———. 1966b. Foreign and United States Residents in University-Affiliated Teaching Hospitals: An Investigation of United States Graduate Medical Education. *Bulletin of the New York Academy of Medicine* 42 (3): 182-208.

Halberstam, J. L., H. A. Rusk, and E. J. Taylor. 1970. Foreign Surgical Residents in University-Affiliated Hospitals: A Unique Case in Graduate Medical Education. *Annals of Surgery* 171 (4): 485-500.

Hofman, K. J., E. S. Tambor, G. A. Chase, G. Geller, R. R. Faden, and N. A. Holtzman. 1993. Physicians' Knowledge of Genetics and Genetic Tests. *Academic Medicine* 68 (8): 625-632.

Imperato, P. J. 1986. Medical Malpractice and Foreign Medical Graduates. *New York State Journal of Medicine* 86 (10): 507-508.

Institute of Medicine. 1996. *The Nation's Physician Workforce: Options for Balancing Supply and Requirements.* Washington, DC: National Academy Press.

Kleinman, J. C., U. C. Brandt, and R. J. Weiss. 1975. Postgraduate Training and Work Experience of Non-ECFMG Certified Physicians in the U.S. *Medical Care* XIII (3): 205-208.

Levit, E. J., and W. D. Holden. 1978. Specialty Board Certification Rates: A Longitudinal Tracking Study of U.S. Medical School Graduates. *Journal of the American Medical Association* 239 (5): 407-412.

Levit, E. J., M. Sabshin, and C. B. Mueller. 1974. Trends in Graduate Medical Education and Specialty Certification. *New England Journal of Medicine* 290 (10): 545-549.

Lockett, B. A. 1975. FMGs: Circumventing the Continuum. *Journal of Medical Education* 50 (1): 103-104.

Lowin, A. 1975. *FMGs: An Evaluation of Policy Related Research.* National Science Foundation, NSF-C214. Minneapolis, MN: InterStudy.

Margulies, H., and L. S. Bloch. 1969. *Foreign Medical Graduates in the United States.* Cambridge, MA: Harvard University Press.

Merchant, F. T. 1973. FLEX—Five Years of Unprecedented Growth (1968-1972). *Federation Bulletin* 60 (9): 293-304.

Mick, S. S., and S.-Y. D. Lee. 1996. *An Analysis of the Comparative Distribution of Active Post-Resident IMGs and USMGs in the United States in 1996.* Rockville, MD: U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions.

Mick, S. S., and T. W. Mou. 1991. The Foreign Medical Graduate Examination in the Medical Sciences (FMGEMS): An Analysis of Pass Rates of the July 1984 Through July 1987 Examinations. *Medical Care* 29 (3): 229-242.

Mick, S. S., and M. N. Pfahler. 1995. *Review and Synthesis of the Literature on Foreign Medical Graduates/International Medical Graduates, 1980-1994.* P.O. HRSA 94-961(P). Rockville, MD: U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions.

Mick, S. S., and F. Rubino. 1992. *The 1987 Career Characteristics of Foreign and U.S. Medical Graduates Who Entered the U.S. Medical System Between 1969 and 1982.* Ann Arbor: University of Michigan, School of Public Health, Department of Health Services Management and Policy.

Mullan, F., R. M. Politzer, and C. H. Davis. 1995. Medical Migration and the Physician Workforce: International Medical Graduates and American Medicine. *Journal of the American Medical Association* 273 (189): 1521-1527.

National Board of Medical Examiners. 1995. *National Board Examiner* 42 (1): 1.

———. 1996. *National Board Examiner* 43 (1): 5.

Norcini, J. J., J. A. Shea, and J. A. Benson. 1991. Changes in the Medical Knowledge of Candidates for Certification. *Annals of Internal Medicine* 114 (1): 33-35.

Norcini, J. J., J. A. Shea, G. D. Webster, and J. A. Benson. 1986. Predictors of the Performance of Foreign Medical Graduates on the 1982 Certifying Examination in Internal Medicine. *Journal of the American Medical Association* 256 (24): 3367-3370.

Part, H. M., and R. J. Markert. 1993. Predicting the First-Year Performances of International Medical Graduates in an Internal Medicine Residency. *Academic Medicine* 68 (11): 856-858.

Payne, B. C., and T. F. Lyons. 1972. *The Episode of Illness: Method of Evaluating and Improving Personal Medical Care Quality.* Ann Arbor: University of Michigan, School of Medicine.

Pew Health Professions Commission. 1995. *Critical Challenges: Revitalizing the Health Professions for the Twenty-First Century.* San Francisco: University of California Center for the Health Professions.

Politzer, R. M., C. E. Yesalis, and J. M. Katzoff. 1989. The Hidden Future Supply of Foreign Medical Graduates. *Medical Care* 27 (11): 1046-1057.

Ramsey, P. G., J. D. Carline, T. S. Inui, E. B. Larson, J. P. LoGerfo, J. J. Norcini, and M. D. Wenrich. 1991. Changes Over Time in Knowledge Base of Practicing Internists. *Journal of the American Medical Association* 266 (8): 1103-1107.

Rhee, S. 1976. Factors Determining the Quality of Physician Performance in Patient Care. *Medical Care* 14 (9): 733-750.

———. 1977a. Relative Importance of Physicians' Personal and Situational Characteristics for the Quality of Patient Care. *Journal of Health and Social Behavior* 18 (March): 10-15.

———. 1977b. U.S. Medical Graduates versus Foreign Medical Graduates: Are There Performance Differences in Practice? *Medical Care* 15 (7): 568-577.

Rhee, S., R. D. Luke, T. F. Lyons, and B. C. Payne. 1981. Domain of Practice and the Quality of Physician Performance. *Medical Care* 19 (1): 14-23.

Rhee, S., T. F. Lyons, B. C. Payne, and S. E. Moskowitz. 1986. USMGs Versus FMGs: Are There Performance Differences in the Ambulatory Care Setting? *Medical Care* 24 (3): 248-258.

Richards, J. 1980. Malpractice Claims Statistics 1974 to Date: General Surgeons Draw Most Cases; Great Majority Settle for Under $10,000. *Michigan Medicine* 79 (2): 25-28.

Rivo, M. L., D. M. Jackson, and F. L. Clare. 1993. Comparing Physician Workforce Reform Recommendations. *Journal of the American Medical Association* 270 (9): 1083-1084.

Rovit, R. L. 1978. Malpractice, Neurosurgery—1977: An Analysis of the AANS Questionnaire. *Clinical Neurosurgery* 25:673-684.

Saywell, R. M., J. A. Bean, R. L. Ludke, R. W. Redman, and G. J. McHugh. 1981. An Examination of the Relationships between Physicians' Clinical and Hospital-Utilization Performance. *Health Services Research* 16 (3): 285-303.

Saywell, R. M., J. A. Bean, J. Studnicki, and R. L. Ludke. 1981. A Comparison of Inappropriate Hospital Utilization: USMG-FMG Attending Physicians. *Philippine American Medical Bulletin* 20 (9): 4-10.

Saywell, R. M., G. J. McHugh, J. A. Bean, R. L. Ludke, and R. W. Redman. 1983. Practicing Physician and House Staff Performance of Patient Histories and Physical Examinations. *Quality Review Bulletin* 9 (2): 48-62.

Saywell, R. M., and J. Studnicki. 1976. *The USMG-FMG Quality of Care Study. Technical Report to DHEW*. Washington, DC: National Technical Information Service.

Saywell, R. M., J. Studnicki, J. A. Bean, and R. L. Ludke. 1979. A Performance Comparison: USMG-FMG Attending Physicians. *American Journal of Public Health* 69 (1): 57-62.

———. 1980. A Performance Comparison: USMG-FMG House Staff Physicians. *American Journal of Public Health* 70 (1): 23-28.

Schnabl, G. K., T. H. Hassard, and M. L. Kopelow. 1991. The Assessment of Interpersonal Skills Using Standardized Patients. *Academic Medicine* 66 (9): S34-S36.

Schwartz, W. B., and D. N. Mendelson. 1989. Physicians Who Have Lost Their Malpractice Insurance: Their Demographic Characteristics and the Surplus-Lines Companies That Insure Them. *Journal of the American Medical Association* 262 (10): 1335-1341.

Sloan, F. A., P. M. Mergenhagen, B. Burfield, R. A. Bovbjerg, and M. Hassan. 1989. Medical Malpractice Experience of Physicians: Predictable or Haphazard? *Journal of the American Medical Association* 262 (23): 3291-3297.

Smith, M. W., and V. K. Fowkes. 1983. Unlicensed Foreign Medical Graduates in California: Social and Demographic Characteristics and Progress Toward Licensure. *Medical Care* 21 (12): 1168-1186.

Stevens, R. A. 1995. International Medical Education and the Concept of Quality: Historical Reflections. *Academic Medicine* 70 (7, Suppl.): S11-S18.

Stevens, R., and J. Vermeulen. 1972. *Foreign Trained Physicians and American Medicine*. Publication no. (NIH) 73-324. Washington, DC: Department of Health, Education, and Welfare.

Stillman, P. L., M. B. Regan, H. L. Haley, J. J. Norcini, M. Friedman, and A. I. Sutnick. 1992. The Use of a Patient Note to Evaluate Clinical Skills of First-Year Residents

Who Are Graduates of Foreign Medical Schools. *Academic Medicine* 67 (10, Suppl.): S57-S59.

Sutnick, A. I. 1970. First Symposium on the Problems of Medical Graduates. *Journal of the American Medical Association* 213 (13): 2241-2246.

Sutnick, A. I., M. L. Shafron, and M. P. Wilson. 1992. Impact of the New United States Licensing Examination on the Certification Process of the Educational Commission for Foreign Medical Graduates. *International Journal of Dermatology* 31 (11): 798-799.

Sutnick, A. I., P. L. Stillman, J. J. Norcini, M. Friedman, M. B. Regan, R. G. Williams, E. K. Kachur, M. A. Haggerty, and M. P. Wilson. 1993. ECFMG Assessment of Clinical Competence of Graduates of Foreign Medical Schools. *Journal of the American Medical Association* 270 (9): 1041-1045.

Swanson, D. B., S. M. Case, D. E. Melnick, and R. L. Volle. 1992. Impact of the USMLE Step 1 on Teaching and Learning of Basic Biomedical Sciences. United States Medical Licensing Examination. *Academic Medicine* 67 (9): 553-556.

Torrey, E. F., and R. L. Taylor. 1973. Cheap Labor from Poor Nations. *American Journal of Psychiatry* 130 (4): 428-433.

Tussing, A. D., and M. A. Wojtowycz. 1993. The Effect of Physician Characteristics on Clinical Behavior: Cesarean Section in New York State. *Social Science and Medicine* 37 (10): 1251-1260.

———. 1997. Malpractice, Defensive Medicine, and Obstetric Behavior. *Medical Care* 35 (2): 172-191.

U.S. General Accounting Office. 1987. *Medical Malpractice: Characteristics of Claims Closed in 1984*. GAO/HRD-87-55. Washington, DC: U.S. Government Printing Office.

———. 1990. *Medical Licensing by Endorsement: Requirements Differ for Graduates of Foreign and U.S. Medical Schools*. GAO/HRD-90-120. Washington, DC: U.S. Government Printing Office.

Waxman, H. S., R. A. Garibaldi, and R. G. Subhiyah. 1996. Performance of U.S. and International Medical Graduates on the 1995 Internal Medicine In-Training Examination. *Annals of Internal Medicine* 125:158.

Weiss, R. J., J. C. Kleinman, U. C. Brandt, J. L. Feldman, and A. C. McGuinness. 1974. Foreign Medical Graduates and the Medical Underground. *New England Journal of Medicine* 290 (25): 1408-1413.

Weiss, R. J., J. C. Kleinman, U. C. Brandt, and D. S. Felsenthal. 1974. The Effect of Importing Physicians—Return to a Pre-Flexnarian Standard. *New England Journal of Medicine* 290 (26): 1453-1458.

Williams, K. N., and R. H. Brook. 1975. Foreign Medical Graduates and Their Impact on the Quality of Medical Care in the United States. *Milbank Memorial Fund Quarterly* 53 (4): 549-581.

Woloshin, S., N. A. Bickell, L. M. Schwartz, F. Gany, and H. G. Welch. 1995. Language Barriers in Medicine in the United States. *Journal of the American Medical Association* 273 (9): 724-728.

Zannoth, D. 1981. Here Are Latest Data on MSMS Malpractice Claims Study. *Michigan Medicine* 80 (32): 624-627.