

Maintenance of Certification and Its Association With the Clinical Knowledge of Family Physicians

Thomas R. O'Neill, PhD, and James C. Puffer, MD

Abstract

Purpose

To examine the relationship between maintenance of certification (MOC) and the clinical knowledge demonstrated by family physicians as they move further away from formal training.

Method

Performances of 10,801 examinees—2,440 seeking initial certification; 8,361 seeking MOC—on the summer 2009 American Board of Family Medicine (ABFM) certification examination were compared across 30 cohorts that represented recent residency program graduates and already-certified family physician diplomates with varying years of clinical

experience. Experience was defined as the time in years since the year of initial certification. This study employed a natural-groups, cross-sectional design; however, it was used to draw longitudinal inferences.

Results

Family physicians who maintained certification performed better than recent graduates. They increased their examination scores by almost 17 points each successive time that they took the exam, with scores reaching their highest point 28 to 31 years after initial certification. Multiple comparison analyses confirmed that

the trend was significant; however, subanalyses revealed that this trend remained significant only for U.S. medical graduates (USMGs) but not international medical graduates. Those family physicians that did not maintain their certification performed significantly worse than recent graduates.

Conclusions

The findings suggest that ABFM diplomates who are USMGs and maintain their certification perform better on the ABFM certification examination with additional years of experience until approximately 30 years after residency training.

Certification in a medical specialty is a voluntary process in which the certifying board establishes criteria that physicians must meet if they desire formal acknowledgement of their professionalism and expertise through the certification process. Although certification in medical specialties has existed since 1917 when the American Board of Ophthalmology was created, the expectation that board-certified physicians be required to periodically pass an examination that assesses their clinical decision-making skills and the adequacy and currency of their fund of knowledge related to their specialties began only in the mid-1970s.¹ Until 1985, only four of the medical specialty boards that were members of the American Board of Medical Specialties (ABMS)

issued time-limited certificates² that required mandatory recertification by examination at periodic intervals. The remainder issued lifetime certificates after their physician diplomates passed the initial certification examination. These diplomates were never examined again. The move to periodic recertification by these specialty boards was driven by the awareness that a physician's clinical knowledge and skills needed to be continuously assessed.^{1,2}

The perception that clinical knowledge declines as a physician moves further away from formal training is prevalent. This perception is reinforced by a significant body of research. Meskauskas and Webster,³ using a cross-sectional design, examined performance on the 1974 internal medicine certification examination to determine whether age and the interval of time since residency were related to the diplomate's performance and found that they were negatively correlated. In a similar cross-sectional study, Norcini and colleagues⁴ equated the results of the 1980 internal medicine recertification examination to those of the 1979 initial certification examination and found that age and the length of time since residency were both

negatively associated with test scores. Similar results have been demonstrated with surgeons⁵⁻⁷ and family physicians.⁸ If studies that consider more narrow domains of medical knowledge—such as those of blood product transfusion, emergency contraception, HIV, and hypertension—are considered, then the systematic review conducted by Choudhry and colleagues⁹ provides several other examples in which similar findings were reported.

It would seem intuitive that recent graduates from residency training programs, exposed to the most current knowledge and sophisticated training, would demonstrate a better grasp of contemporary medical knowledge within their specialties than would physicians who were many years removed from formal training. However, no studies have examined the influence that rigorous mandatory requirements of continuing medical education (CME), assessment of performance in practice, and periodic reexamination necessary for maintaining certification has had to verify this conjecture. Therefore, we wished to answer the question of whether recent residency graduates outperform seasoned family physicians

Dr. O'Neill is vice president of psychometric services, American Board of Family Medicine, Lexington, Kentucky.

Dr. Puffer is president and CEO, American Board of Family Medicine, Lexington, Kentucky.

Correspondence should be addressed to Dr. O'Neill, 1648 McGrathiana Pkwy., Suite 550, Lexington, KY 40511-1247; telephone: (859) 269-5626 ext. 1225; e-mail: toneill@theabfm.org.

Acad Med. 2013;88:780-787.

First published online April 24, 2013

doi: 10.1097/ACM.0b013e3182900de4

on an examination that assesses the medical knowledge and clinical decision-making skills necessary for certification in family medicine. We found that this question can be empirically examined because the physicians who are applying for initial certification by the American Board of Family Medicine (ABFM) and the physicians who are taking the examination to maintain their certification must take the same examination.

Background

With its inception in 1969 as the 20th medical specialty board approved by the ABMS, the ABFM became the first medical specialty board to issue time-limited certificates and to mandate recertification¹⁰; the first cohort of examinees was in 1970. The ABFM accomplished this by requiring diplomates to accumulate 300 CME credits, conduct a review of performance in practice via office record review, and successfully pass an examination at least every 7 years to maintain their certification. The ABFM replaced its traditional 7-year recertification paradigm with a new maintenance of certification (MOC) process in 2002—Maintenance of Certification for Family Physicians (MC-FP). This new process allowed diplomates who successfully met the requirements of each of its four components to increase the period of time permitted before they needed to be reexamined to 10 years rather than 7. The last year in which candidates had to test once every 7 years was 2009, although they could test in year 6 if they wanted more examination opportunities in case they should fail.

The majority of examinees that take the ABFM MC-FP examination pass. The overall pass rates for the period from 2007 to 2009 have ranged from 80% to 82%, and the pass rates during this period for those not taking the examination because of a prior failure have been 85% to 88%.¹¹ The examination is offered twice yearly.

Method

Participants

The participants were the 10,801 examinees (2,440 to obtain initial certification; 8,361 to maintain their

certification) that took the summer 2009 administration of the ABFM MC-FP examination. These examinees took the test not as part of an experiment but, rather, for the purpose of earning or maintaining their board certification. The 73 examinees who were initially certified between 1970 and 1974 were excluded from the analysis because completing a residency program was not a requirement at that time. The remaining examinees were placed into 30 mutually exclusive cohorts based on their year of initial certification. Examinees who had not yet earned their initial certification were placed into the 2009 cohort.

As a condition of participating in the ABFM certification program, all candidates and diplomates must agree to the terms of the ABFM privacy policy, which delineates the manner in which the ABFM may use the physician's personal information. We deidentified the data before analyzing them, and the ABFM's executive staff reviewed the research before external distribution to ensure that the ABFM's privacy policies had not been violated.

Instruments

The examination administered was the 2009 ABFM MC-FP examination,¹² consisting of 350 scored items; 260 are from a "common core" defined by the ABFM examination test plan specifications. A description of the content categories, the percentage of items assigned to each category, and the rationale for using these categories and percentages are well documented.¹³ The remaining 90 items are from two 45-item, topic-specific modules that are selected by the examinee from a menu of eight modules. These modules are ambulatory family medicine, child and adolescent care, geriatrics, women's health, maternity care, emergent/urgent care, hospital medicine, and sports medicine. We scaled all items by difficulty onto a common metric using the dichotomous Rasch¹⁴ model, a form of item response theory commonly used in high-stakes testing. Using those item-difficulty values, we also computed ability estimates for the examinees using the Rasch model. Candidate ability estimates were then converted to scaled scores that range from 200 to 800.

Design

This study employed a natural-groups, cross-sectional design; however, it was

used to draw longitudinal inferences. Longitudinal data were not available for this study because before 2008, the scoring scales were not equated across years. The most recent data for which this type of analysis is appropriate were the data from 2009 because the ABFM replaced its traditional seven-year recertification paradigm with a new MOC process in 2002. Our analysis would not have been possible using data from the 2010, 2011, and 2012 examinations because the majority of diplomates certifying or recertifying in 2003 or 2004 successfully met their MC-FP requirements, drastically reducing the number of diplomates taking the examination in 2010, 2011, and 2012.

A *cohort* was composed of all those who took the examination who shared the same year of initial certification. Those already-certified family physicians taking the summer 2009 examination were not randomly selected from the population of diplomates across all years of initial certification; rather, the examinees were those who—no matter when they were initially certified—opted to take the examination in such a way as to maximize the amount of time between tests without having a gap in their certification.

Statistical analyses

For each cohort, the number of examinees and the mean score was computed, as well as the standard deviation and standard error of that mean. Each cohort was assigned to one of three mutually exclusive groups representing (1) initial certifiers, (2) examinees with uninterrupted certification, and (3) examinees with gaps in their certification.

Across the 30 different cohorts, multiple comparison procedures could be conducted on the mean scaled scores for all 435 possible cohort pairs; however, that would drastically increase the experiment-wise alpha error without contributing to the evaluation of the question we were attempting to answer. Instead, we decided to exclude the cohort of examinees from years prior to the requirement of residency training and then separately consider the cohorts that had uninterrupted certification and those that had gaps in their certification. We further decided to group the cohorts by how many recertification cycles

(zero to five) they had achieved. As a final constraint, the mean score of any group would be compared only with that of the group immediately preceding it—for instance, initial certification would be compared with one recertification, or one recertification would be compared with two recertifications. This resulted in nine planned comparisons. To maintain an experiment-wise alpha level of .05, a Bonferroni correction was computed that set the level for significance at .005. This procedure was repeated to break

out results by U.S. medical graduates (USMGs) and international medical graduates (IMGs) to see whether a trend was attributable largely to the influence of one group or the other.

Results

There were radically different numbers of examinees in the “year of initial certification” cohorts (see Table 1). The vast majority of the MC-FP examinees were taking the examination “on

cycle,” thereby avoiding a gap in their certification. The low-volume cohorts could be described as testing “off cycle,” having experienced a gap in their certification. The 2003 cohort had only four examinees and therefore has near-zero inferential stability. It is included in Table 1 and Figure 1 for the sake of completeness, but it was otherwise excluded from further consideration.

To understand the grouping rationale, it is helpful to begin by looking at the

Table 1
Summary Data From the 2009 American Board of Family Medicine Examination by Year of Examinee’s Initial Certification*

Cohort year [†]	Group [‡]	Mean score (SD) [SE]			Total N	USMGs’ N	IMGs’ N
		Total	USMGs [§]	IMGs [§]			
1975	UC	495 (120) [20]	499 (117) [20]	465 (160) [80]	37	33	4
1976	UC	510 (107) [10]	— (–) [–]	— (–) [–]	105	103	1
1977	UC	515 (110) [8]	522 (104) [8]	378 (83) [24]	193	179	12
1978	UC	524 (121) [7]	534 (118) [7]	407 (105) [22]	315	285	23
1979	GAP	420 (120) [18]	441 (118) [20]	355 (105) [32]	47	36	11
1980	GAP	428 (118) [16]	445 (121) [18]	354 (67) [21]	54	44	10
1981	UC	538 (114) [7]	548 (109) [7]	405 (74) [17]	250	230	18
1982	UC	555 (119) [7]	562 (115) [7]	427 (108) [28]	264	247	15
1983	UC	551 (117) [7]	562 (108) [7]	433 (146) [31]	263	240	22
1984	UC	561 (121) [6]	571 (115) [6]	433 (124) [24]	350	323	26
1985	GAP	454 (109) [17]	474 (112) [20]	386 (61) [20]	40	31	9
1986	GAP	445 (94) [15]	445 (93) [18]	434 (124) [24]	37	27	9
1987	GAP	448 (103) [11]	465 (104) [13]	398 (90) [18]	91	64	24
1988	UC	529 (103) [4]	534 (101) [5]	453 (91) [15]	534	492	35
1989	UC	528 (112) [5]	537 (109) [6]	435 (103) [16]	436	394	41
1990	UC	530 (106) [4]	537 (103) [4]	449 (113) [18]	578	533	42
1991	GAP	422 (118) [30]	419 (126) [35]	440 (70) [50]	15	13	2
1992	GAP	371 (81) [16]	379 (88) [21]	354 (69) [23]	27	18	9
1993	GAP	414 (82) [13]	417 (86) [16]	389 (61) [20]	37	27	9
1994	GAP	393 (77) [7]	401 (84) [9]	375 (54) [7]	135	79	54
1995	UC	496 (112) [4]	515 (108) [4]	416 (89) [7]	808	636	161
1996	UC	506 (108) [3]	518 (104) [4]	429 (89) [7]	1,016	830	159
1997	GAP	462 (145) [32]	482 (153) [38]	355 (44) [22]	21	16	4
1998	GAP	388 (103) [28]	399 (107) [32]	347 (90) [52]	14	11	3
1999	GAP	422 (81) [16]	414 (80) [17]	410 (71) [50]	26	22	2
2000	GAP	384 (68) [8]	379 (67) [9]	347 (90) [52]	80	56	23
2001	GAP	397 (64) [5]	399 (67) [6]	391 (57) [9]	163	117	45
2002	UC	509 (100) [2]	520 (98) [2]	457 (90) [4]	2,348	1,914	417
2003	UC	445 (81) [41]	— (–) [–]	— (–) [–]	4	3	1
2009	IC	476 (86) [2]	493 (87) [2]	448 (76) [3]	2,440	1,530	901

* Excluding the 73 examinees who were initially certified between 1970 and 1974, a total of 10,728 examinees (2,440 seeking initial certification, 8,288 seeking maintenance of certification) took the 2009 administration of the test.

† A cohort was composed of all those who took the examination who shared the same year of initial certification.

‡ The mutually exclusive groups were composed of examinees with uninterrupted certification (UC) or with a gap in certification (GAP), or who were seeking initial certification (IC). At the time of this examination, diplomats were required to take the exam once every seven years but were prohibited from taking it earlier than their sixth year into their seven-year cycle (after which cycle they were required to test again).

§ USMG indicates United States Medical Graduate; IMG, international medical graduate.

number of examinees (the bars in Figure 1) in each cohort. The initial certification cohort, 2009, has 2,440 examinees. This volume was generally expected, given the number of residents completing their training. Going backward, the next cohort with a large number of candidates is 2002. This is because those physicians, who initially certified in 2002, needed to pass the examination in 2009 to remain certified. Interestingly, almost no one who was certified in 2003 chose to take the examination in their sixth year, which was 2009. Going backward from 2002, the next cohorts with a large number of candidates were 1995 and 1996, which were six and seven years before 2002. Following the same logic, the rest of the “on cycle” cohorts were placed into groups that represent their number of recertifications. In short, the cohorts from 1975, 1976, 1977, 1978, 1981, 1982, 1983, 1984, 1988, 1989, 1990, 1995, 1996, and 2002 represent, almost without exception, diplomates who had successfully maintained continuous certification since their initial certification. Conversely, the cohorts in this data set from 1979, 1980, 1985, 1986, 1987, 1991, 1992, 1993, 1994, 1997, 1998, 1999, 2000, and 2001 represent diplomates who have at some point experienced a gap in their certification history.

Uninterrupted certification

When the performance of the initial certification group (represented in Figure 1 with a diamond) was compared with the performance of the groups who had more than six years of postresidency experience and had maintained their certification without interruption (represented in Figure 1 with a square), the diplomates with postresidency experience performed better. Among the uninterrupted certification groups (see Table 2), statistically significant gains were observed from zero to one, two to three, and three to four recertifications. The gain from one to two was not significant, and the change from four to five was a significant drop. This suggests that a trend exists that is unlikely to have occurred by chance.

Gap in certification

When the performance of the initial certification group was compared with the performance of the cohorts who had more than six years of postresidency experience but had gaps in their

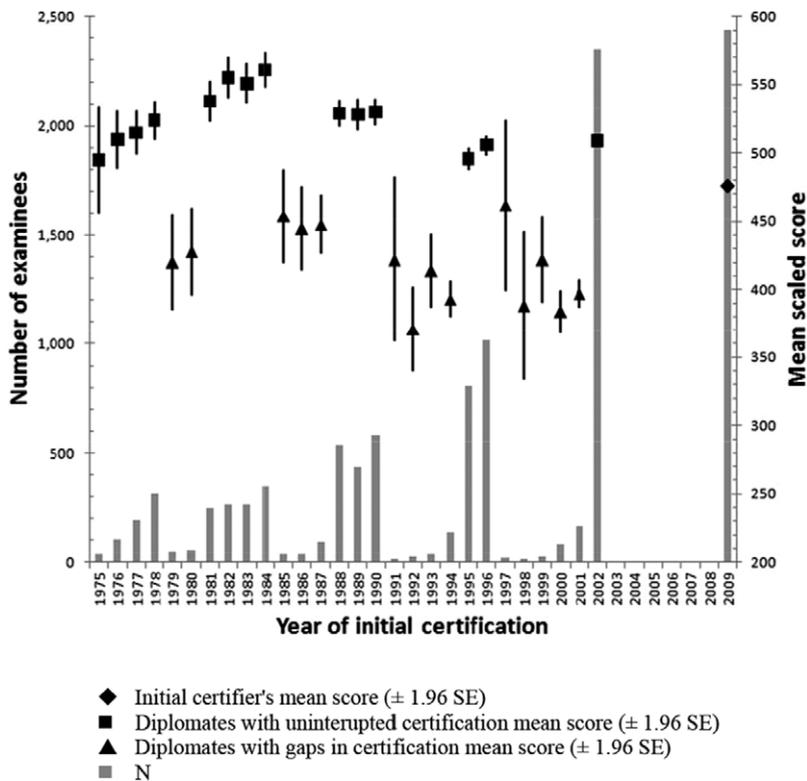


Figure 1 All groups' mean scores (95% confidence interval) and examinee volume by year of initial certification. Performances of 10,728 examinees—2,440 seeking initial certification; 8,288 seeking maintenance of certification—on the summer 2009 American Board of Family Medicine certification examination were compared across 30 cohorts that represented recent residency program graduates and already-certified family physician diplomates with varying years of clinical experience.

certification (represented in Figure 1 with a triangle), it is clear that the family physicians with postresidency experience performed worse. It is important to note that the number of examinees with certification gaps accounts for less than 10% of all recertifiers, whereas the uninterrupted certification group accounts for more than 90%. Among the groups with gaps in their certification (see Table 2), statistically significant decreases were observed from zero to one recertifications. The decreases from one to two and from three to four were not significant. There was a significant gain from two to three, but it was not large enough to offset the decrease from zero to one. Among the gapped groups, only two comparisons were unlikely to have occurred by chance, and those were in opposite directions.

Comparing USMGs with IMGs

Figure 2 presents key findings for USMGs, and Figure 3 presents such findings for IMGs. In instances where a cohort had less than two examinees in either the USMG or IMG group, the data were excluded

because a standard deviation could not be computed, and the score of an individual examinee might be able to be surmised.

Multiple comparison procedures for the USMG groups whose members had uninterrupted certification demonstrated a very similar pattern of increases to that of the total group; however, only the increases from zero to one recertifications and three to four recertifications were statistically significant (see Table 3). The decrease from four to five recertifications was similar to the trend in the entire group; both trends were statistically significant. For IMGs, it was unusual for any of the cohorts to exceed the mean score of the initial certifiers. Multiple comparison procedures for the IMG group showed only one statistically significant change, a decrease, and it did not suggest a general trend. For the groups with gaps in their certification, multiple comparison procedures broken out by type of medical education were not carried out because a discernible trend had not emerged earlier from the combined sample.

Table 2

Multiple Comparison Procedures for Mean Scores by Cohort Group, for the Entire Sample of Examinees Seeking Recertification in the 2009 American Board of Family Medicine Examination

Group	Smaller number of recertifications			Larger number of recertifications			Mean increase	t	df	P value
	Number of recertifications	Mean (SD)	N	Number of recertifications	Mean (SD)	N				
Uninterrupted certification	0	480.7 (86.7)	2,440	1	513.1 (102.7)	3,364	32.4	13.0*	5,670.5	.000 [†]
	1	513.1 (102.7)	3,364	2	514.8 (110.9)	1,386	1.7	0.5	2,412.3	.619
	2	514.8 (110.9)	1,386	3	533.2 (108.4)	970	18.4	4.0	2,354.0	.000 [†]
	3	533.2 (108.4)	970	4	551.0 (121.0)	1,442	17.7	3.8	2,223.7	.000 [†]
	4	551.0 (121.0)	1,442	5	515.8 (110.3)	335	-35.2	-5.2	537	.000 [†]
Gap in certification	0	480.7 (86.7)	2,440	1	403.7 (79.2)	304	-77.0	-15.8*	399.0	.000 [†]
	1	403.7 (79.2)	304	2	400.3 (82.6)	214	-3.4	-0.5	516.0	.637
	2	400.3 (82.6)	214	3	453.4 (102.1)	168	53.1	5.5*	317.0	.000 [†]
	3	453.4 (102.1)	168	4	428.9 (118.5)	101	-24.5	-1.8	267.0	.074

* Indicates that equal variances were not assumed in the t test for equality of means; however, in all cases, the equal variance formula would have led to the same significance conclusion as did the unequal variance formula.

† Indicates a statistically significant difference. (The significance level was adjusted to .005 to maintain an experiment-wise alpha level of .05 across all nine t tests.)

The mean score of the initial certifiers was lower for IMGs ($\bar{x} = 446$) than for USMGs ($\bar{x} = 493$). Because there were over 900 IMGs in the 2009 cohort, their lower mean score noticeably lowered the mean score for the total group ($\bar{x} = 476$) in Figure 1. As the result of having a higher baseline, the score increase over time for the USMG diplomates without gaps in their certification (Figure 2) seems less dramatic than when both USMGs and IMGs are considered together (Figure 1). Nevertheless, the trend was still discernible.

Discussion

What our findings suggest

In general, the results demonstrate that for the individuals we studied, diplomates without gaps in their certification outperformed both initial certifiers and those who had gaps in their certification. The data also suggest that when there are no gaps in certification, a tendency for examination scores to improve with additional recertifications exists; however, this is a longitudinal assertion based on results from cross-sectional data.

Diplomates without certification gaps represented more than 90% of the recertification candidates. Examining the trend across these cohorts, a picture emerged that suggested that family physicians who maintain their certification without interruption continue to develop

their skills and enhance their knowledge after residency. More specifically, the results show that the family physicians studied typically increased their scores by almost 17 points every 7 years, with their scores reaching their highest point 28 to 31 years after initial certification. At this point, most of the physicians studied were 53 to 55 years old. However, on their next certification exam, 6 or 7 years later, their scores dropped on average about 35 points. At this point, the physicians were typically 59 to 61 years of age. This score dip may reflect these physicians redirecting some of their professional development energy into transitioning into retirement; however, even after the dip, this group still outperformed the initial certifiers. Less than 10% of the recertification candidates had gaps in their certification. For these diplomates (see Figure 1), a discernible trend beyond the initial drop is not apparent. They typically did not perform as well as the cohort attempting their initial certification, despite their additional experience.

In the 1990s, a greater number of IMGs were applying to take the ABFM certification examination. Although many passed, these applicants, as a group, do not traditionally perform as well on the test as do their U.S.-trained counterparts.^{11,15} Because of the cross-sectional design of this study, we analyzed the results separately for USMGs and IMGs to see whether combining these

groups either amplified or obscured a trend. The performance of USMGs in each of the cohorts appears to account for the significant trends demonstrated.

The evidence that length of time after initial certification is associated with lower recertification examination scores has been largely limited to the specialties of internal medicine and surgery. Our results suggest that in family medicine—a specialty that has always mandated CME, assessment of performance in practice, and periodic assessment of cognitive expertise by examination in order to maintain certification—practitioners’ cognitive expertise and clinical decision-making skills improve with clinical experience up to a certain point. The upward trend in our data peaking from 1980 to 1984 and the slight decline for the 1975 to 1978 cohorts may represent a general path of physician ability over the course of their career. These results might suggest that family physicians (1) begin with solid training in residency, (2) gain more experience through practice and professional development postresidency, and then, near the end of their careers, (3) channel some of the energy that previously was allocated for professional development into considerations related to transitioning into retirement.

This type of progression would be consistent with current research related to

the development of expertise, which shows that people normally need a minimum of 10 years (or 10,000 hours) of specifically focused, intense training before they can be at the top of their profession.¹⁶⁻¹⁸ The results of this study suggest that a high degree of proficiency can be reached on completion of residency, but it does not typically represent the same level of performance as that demonstrated by family medicine clinicians who for 25 to 30 years have continued to expand their knowledge base, hone their existing skills, and acquire new ones. The latter group's higher level of performance may partly result from its members' consistent and habitual process of continuing the professional development required of every board-certified family medicine physician for over 40 years.

This is congruent with Rhodes and colleagues¹⁹ work, which demonstrated that performance on the American Board of Surgery examination was directly correlated with the quantity and quality of CME. However, other factors may also explain our findings. Eva²⁰ has demonstrated that recognizing patterns or using already-acquired knowledge is largely unaffected by aging. Day and colleagues²¹ suggestion that physicians further away from initial certification may have trouble updating their knowledge base rather than experiencing a deterioration of their original knowledge base would also support this line of thinking.

In either case, when the argument is made that older diplomates demonstrate a decline in clinical knowledge, we have shown that this was not true with the family physicians we studied who were conscientious about maintaining their certification. Although physicians who were approximately 35 years postresidency had scores lower than those of their 28-year-postresidency counterparts, this decrease was typically less than the gains made in the prior 14 years. Even with the decrease, diplomates who had continuously maintained their certification typically outperformed initial certifiers. The issue of a decline in clinical knowledge and decision-making skills tended to be more of a problem with diplomates who had gaps in their certification, as evidenced by the data in this study.

Our results would seem to complement the work by Turchin et al.²² which demonstrated that compliance with

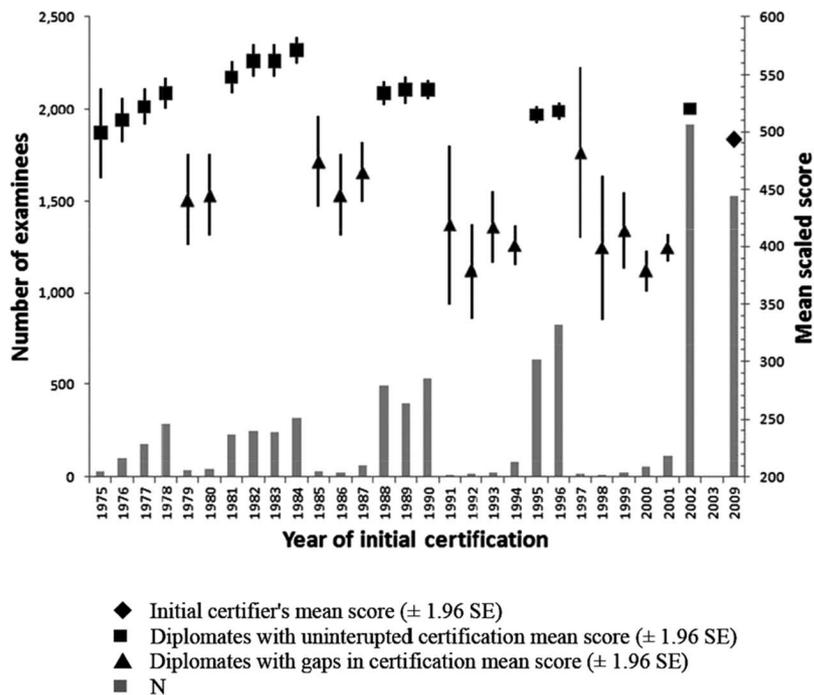


Figure 2 U.S. medical graduates' mean scores (95% confidence interval) and examinee volume by year of initial certification. Performances of 8,530 examinees—1,530 seeking initial certification; 7,000 seeking maintenance of certification—on the summer 2009 American Board of Family Medicine certification examination were compared across 30 cohorts that represented recent residency program graduates and already-certified family physician diplomates with varying years of clinical experience.

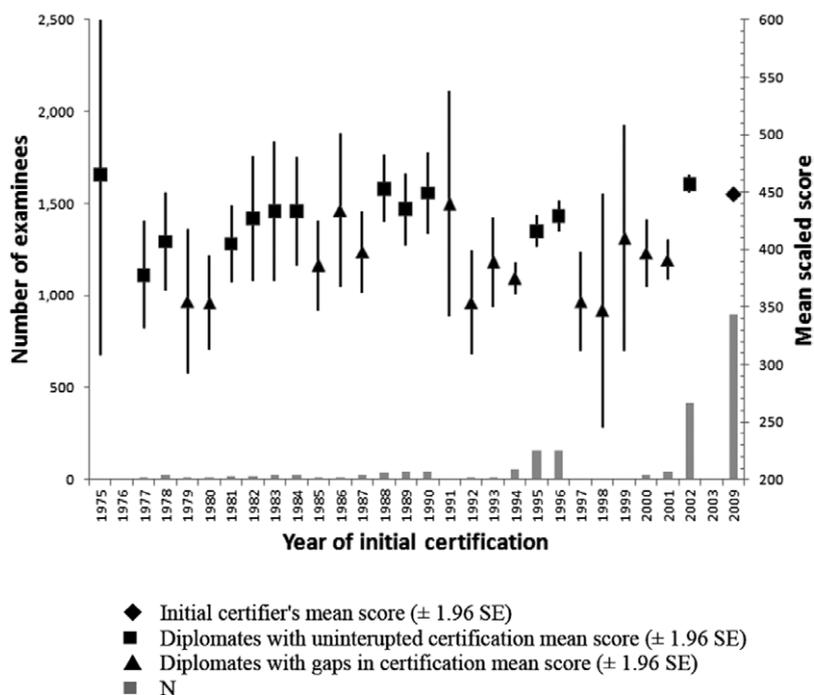


Figure 3 International graduates' mean scores (95% confidence interval) and examinee volume by year of initial certification. Performances of 2,092 examinees—901 seeking initial certification; 1,191 seeking maintenance of certification—on the summer 2009 American Board of Family Medicine certification examination were compared across 30 cohorts that represented recent residency program graduates and already-certified family physician diplomates with varying years of clinical experience.

Table 3

Multiple Comparison Procedures for Mean Scores by Type of Medical Education for Examinees With Uninterrupted Certification Who Sought Recertification in the 2009 American Board of Family Medicine Examination

Group	Smaller number of recertifications			Larger number of recertifications			Mean increase	t	df	P value
	Number of recertifications	Mean (SD)	N	Number of recertifications	Mean (SD)	N				
U.S. medical graduates	0	497.3 (87.9)	1,530	1	524.5 (100.1)	2,744	27.1	9.2*	3,516.6	.000†
	1	524.5 (100.1)	2,744	2	529.6 (106.7)	1,169	5.1	1.4*	2,082.6	.162
	2	529.6 (106.7)	1,169	3	540.2 (106.1)	886	10.6	2.2	2,053.0	.025
	3	540.2 (106.1)	886	4	560.3 (115.9)	1,325	20.1	4.2*	2,006.5	.000†
	4	560.3 (115.9)	1,325	5	519.9 (106.1)	315	-40.4	-6.0*	507	.000†
International medical graduates	0	452.0 (76.5)	901	1	453.6 (90.5)	576	1.6	0.38*	1,077.4	.732
	1	453.6 (90.5)	576	2	427.9 (94.7)	203	-25.7	-3.4	777.0	.001†
	2	427.9 (94.7)	203	3	448.1 (97.0)	76	20.2	1.6	277.0	.117
	3	448.1 (97.0)	76	4	425.4 (117.2)	104	-22.6	-1.4	178.0	.171
	4	425.4 (117.2)	104	5	402.6 (104.6)	17	-22.8	-5.2	119.0	.452

* Indicates that equal variances were not assumed in the t test for equality of means; however, in all cases, the equal variance formula would have led to the same significance conclusion as did the unequal variance formula.

† Indicates a statistically significant difference. (The significance level was adjusted to .005 to maintain an experiment-wise alpha level of .05 across all nine t tests.)

well-established practice guidelines in the management of hypertension in diabetic patients is better in physicians just out of training as well as those who have recently taken a recertification examination. This would imply that that regular assessment by examination is likely to compel physicians to remain current with recent practice guidelines and, therefore, more likely to comply with them.

Limitations

In this study, a natural-groups, cross-sectional design was used to draw longitudinal conclusions. To logically make these conclusions, one must assume that no systematic bias in the attrition process within cohorts exists. The attrition rate for ABFM diplomates under the age of 55 was rather small, so the potential impact of a systematic bias in attrition should also be small.¹⁵ One must also assume that each cohort was generally of equal ability in its initial certification year. Although no strong evidence exists to suggest that these requirements were violated, it is important to note that these requirements have been assumed, not demonstrated. The ABFM began the process of maintaining a single examination scale across years starting in 2007, which will make future longitudinal comparisons possible by using a stable frame of reference.

Conclusion

Our results strongly suggest that conscientious participation in the rigorous and structured processes required to maintain certification results in continued improvement in clinical knowledge over time. This knowledge is one of the critically important factors necessary to deliver high-quality care, as it is necessary to inform sound clinical judgment, make evidence-based decisions, deal with uncertainty, and manage the complexity and volume of patients typically seen by practicing physicians.^{23,24}

Acknowledgments: The editorial assistance of Ms. Nichole Lainhart is gratefully acknowledged.

Funding/Support: None.

Other disclosures: The study was conducted on a preexisting, deidentified data set.

Ethical approval: Not applicable.

References

- 1 Buyske J. For the protection of the public and the good of the specialty: Maintenance of certification. *Arch Surg.* 2009;144:101–103.
- 2 Ramsey PG, Carline JD, Inui TS, et al. Changes over time in the knowledge base of practicing internists. *JAMA.* 1991;266:1103–1107.
- 3 Meskauskas JA, Webster GD. The American Board of Internal Medicine recertification examination: Process and results. *Ann Intern Med.* 1975;82:577–581.
- 4 Norcini JJ, Lipner RS, Benson JA Jr, Webster GD. An analysis of the knowledge base of practicing internists as measured by the 1980

- recertification examination. *Ann Intern Med.* 1985;102:385–389.
- 5 Cruft GE, Humphreys JW Jr, Hermann RE, Meskauskas JA. Recertification in surgery, 1980. *Arch Surg.* 1981;116:1093–1096.
- 6 Rhodes RS, Biester TW. Certification and maintenance of certification in surgery. *Surg Clin North Am.* 2007;87:825–836, vi.
- 7 Lipner R, Song H, Biester T, Rhodes R. Factors that influence general internists' and surgeons' performance on maintenance of certification exams. *Acad Med.* 2011;86:53–58.
- 8 Leigh TM, Young PR, Haley JV. Performances of family practice diplomates on successive mandatory recertification examinations. *Acad Med.* 1993;68:912–919.
- 9 Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: The relationship between clinical experience and quality of health care. *Ann Intern Med.* 2005;142:260–273.
- 10 Green LA, Puffer JC. Family medicine at 40 years of age: The journey to transformation continues. *J Am Board Fam Med.* 2010;23(suppl 1):S1–S4.
- 11 American Board of Family Medicine. Examination pass rates. <https://www.theabfm.org/ThreeColumnTemplate.aspx?id=25769804146&terms=pass+rates>. Accessed February 13, 2013.
- 12 American Board of Family Medicine. Candidate Information Booklet. <https://www.theabfm.org/cert/2012MCFPCandInfoBook.pdf>. Accessed February 13, 2013.
- 13 Norris TE, Rovinelli RJ, Puffer JC, Rinaldo J, Price DW. From specialty-based to practice-based: A new blueprint for the American Board of Family Medicine cognitive examination. *J Am Board Fam Pract.* 2005;18:546–554.
- 14 Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests, 2nd edition. Chicago, Ill: University of Chicago Press; 1980.
- 15 Xierali IM, Rinaldo JC, Green LA, et al. Family physician participation in

- maintenance of certification. *Ann Fam Med*. 2011;9:203–210.
- 16 Ericsson KA, Krampe RT, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev*. 1993;100:363–406.
 - 17 Ericsson KA, Prietula MJ, Cokely ET. The making of an expert. *Harv Bus Rev*. July–August 2007;85:114–121.
 - 18 Ericsson KA, Roring RW, Nandagopal K. Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Stud*. 2007;18:3–56.
 - 19 Rhodes RS, Biesten TW, Ritchie WP, Malangoni MA. Continuing medical education activity and American Board of Surgery examination performance. *J Am Coll Surg*. 2003;196:604–609.
 - 20 Eva KW. The aging physician: Changes in cognitive processing and their impact on medical practice. *Acad Med*. 2002;77(10 suppl):S1–S6.
 - 21 Day SC, Norcini JJ, Webster GD, Viner ED, Chirico AM. The effect of changes in medical knowledge on examination performance at the time of recertification. *Res Med Educ*. 1988;27:139–144.
 - 22 Turchin A, Shubina M, Chodos AH, Einbinder JS, Pendergrass ML. Effect of board certification on antihypertensive treatment intensification in patients with diabetes mellitus. *Circulation*. 2008;117:623–628.
 - 23 Brennan TA, Horwitz RI, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification status in the quality movement. *JAMA*. 2004;292:1038–1043.
 - 24 Holmboe ES, Lipner R, Greiner A. Assessing quality of care: Knowledge matters. *JAMA*. 2008;299:338–340.

Teaching and Learning Moments

Debriefing After Catastrophe

Bill, one of my active 78-year-old patients, underwent minimally invasive coronary artery bypass grafting in 2008. His first two postoperative weeks were complicated by shock, sternotomy with additional bypasses, a multilobar pneumonia, renal failure requiring hemodialysis, fungal empyema, clostridium difficile diarrhea, and chest compartment syndrome precipitating a mediastinal “washout.” Amidst this catastrophic cascade, I called a family meeting, told Bill’s wife and children that even if he were (against great odds) to live, he would be a shadow of his former self, then asked them to make him “comfort measures only.” The surgical team disagreed with my assessment and the family’s opinion was split, so we continued aggressive care.

Bill survived, although he spent two months in the surgical intensive care unit, three additional months in a subacute rehabilitation facility, and was admitted twice more to the hospital. He currently lives with his wife, continues to spend time with his family, and winters in Florida. He struggles with depression and chronic chest wall pain, has gone blind from causes unrelated to his surgery, and has ongoing nonischemic dyspnea on exertion. He can walk perhaps one block and is understandably less engaged with his family business and his community than he had been prior to his surgery.

In 2012, Bill, his wife Elizabeth, and their daughter Margaret participated in our

section’s morbidity and mortality grand rounds. I presented his history to the audience, and Bill, his family, and I sat as panelists while a moderator led us in a debriefing during which we reflected together on this episode, how it molded our physician–patient relationship, and what each of us learned from having come through it. Bill and his family were generous, the audience was engaged, and the experience was profoundly educational.

When asked if he had suffered in the hospital, Bill replied that he remembers almost nothing of his initial postoperative weeks. While many in the audience were surprised, Bill’s answer mirrors the literature suggesting that patients frequently have little or no recall of stays in the intensive care unit.

The audience asked Elizabeth and Margaret whether they had ever discussed end-of-life care with Bill. When they admitted they had not (a shortcoming of primary care), they described what it had felt like making such consequential decisions without a clear vision of what he would have wanted. Margaret spoke powerfully of having to “choose from a buffet of medical options about which she knew very little,” and to the difficulty of being torn between diverging medical opinions. She was unsurprised when a member of the audience cited that surrogate decision makers incorrectly represent patients’ wishes in up to one-third of cases and that no more

than 10% of Americans hold advance directives.

Finally, the moderator initiated a conversation between Bill and me, asking me how I had felt when Bill survived, and asking Bill what advice he would give physicians. Fortunately, he and I had had this conversation before, so our exchange was comfortable. I described my relief that Bill had lived, my shame at having offered his family an incorrect prediction that could have led to the withdrawal of care, and my frustration over Bill’s vastly diminished quality of life. Bill asked that physicians be more aware of the potential risks of the treatments and procedures we offer our patients and that we have realistic discussions with them about expected courses of recovery.

These rounds were a staggering reminder of the trickiness of predicting outcomes during catastrophic illness, the difficulties of surrogate decision making, the problem of disconcerted care, and the value of debriefing with our patients after trying episodes.

Author’s Note: The names in this essay have been changed to protect the identity of the patient and his family.

Michael Stillman, MD

Dr. Stillman is associate program director, Department of Internal Medicine, University of Louisville School of Medicine, Louisville, Kentucky; e-mail: michael.stillman@louisville.edu.