Development and Growth of a Large Multispecialty Certification Examination: Sleep Medicine Certification—Results of the First Three Examinations

Stuart F. Quan, M.D.^{1,2}; Daniel J. Buysse, M.D.³; Sally L. Davidson Ward, M.D.⁴; Susan M. Harding, M.D.⁵; Conrad Iber, M.D.⁶; Vishesh K. Kapur, M.D., M.P.H.⁷; James A. Rowley, M.D.⁸; Michael J. Sateia, M.D.⁹; Michael H. Silber, MB.Ch.B.¹⁰; Adam J. Sorscher, M.D.¹¹; Bradley V. Vaughn, M.D.¹²; Manisha Witmans, M.D.¹³; B. Tucker Woodson, M.D.¹⁴; Phyllis Zee, M.D., Ph.D.¹⁵; Linda E. Mills, M.A.¹⁶; Brian J. Hess, Ph.D.¹⁶

¹Division of Sleep Medicine, Harvard Medical School, Boston, MA; ²Arizona Respiratory Center, University of Arizona, Tucson, AZ; ³Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA; ⁴Division of Pediatric Pulmonology, Children's Hospital Los Angeles, and Keck School of Medicine, University of Southern California, Los Angeles, CA; ⁵UAB Sleep-Wake Disorders Center, University of Alabama at Birmingham, Birmingham, AL; ⁶Fairview Sleep Center at the University of Minnesota Medical Center, Minneapolis, MN; ⁷Division of Pulmonary & Critical Care Medicine, University of Washington, Seattle, WA; ⁸Division of Pulmonary, Critical Care, and Sleep Medicine, Wayne State University School of Medicine, Detroit, MI; ⁹Department of Psychiatry, Dartmouth Medical School, Hanover, NH; ¹⁰Department of Neurology and Center for Sleep Medicine, Mayo Clinic College of Medicine, Rochester, MN; ¹¹Department of Psychiatry, Dartmouth Medical School, Hanover, NH; ¹²Department of Neurology, University of North Carolina, Chapel Hill, NC; ¹³Department of Paediatrics, University of Alberta, Edmonton, Alberta, Canada, and Stollery Children's Hospital, Edmonton, Alberta, Canada; ¹⁴Department of Otolaryngology and Human Communication, Medical College of Wisconsin, Milwaukee, WI; ¹⁵Center for Sleep & Circadian Biology, and Institute for Neuroscience, Northwestern University, Chicago, IL; ¹⁶American Board of Internal Medicine, Philadelphia, PA

This paper summarizes the results of the first three examinations (2007, 2009, and 2011) of the Sleep Medicine Certification Examination, administered by its six sponsoring American Board of Medical Specialty Boards. There were 2,913 candidates who took the 2011 examination through one of three pathways—self-attested practice experience, previous certification by the American Board of Sleep Medicine, or formal Sleep Medicine fellowship training. The 2011 exam was the last administration in which candidates who had not previously been admitted could take it without completion of formal Sleep Medicine fellowship training. As expected, the number of candidates admitted to the 2011 examination through the practice experience pathway increased, and the overall scores of these candidates were on average lower than the other candidates.

Consequently, the pass rate for all first takers of the 2011 examination (65%) was lower than that observed from the 2009 examination (78%) and the 2007 examination (73%). For each administration, candidates admitted through the fellowship training pathway scored the highest; over 90% of them passed the 2011 and 2009 examinations.

Keywords: Sleep Medicine Certification Examination, American Board of Medical Specialty Boards

Citation: Quan SF; Buysse DJ; Ward SLD; Harding SM; Iber C; Kapur VK; Rowley JA; Sateia MJ; Silber MH; Sorscher AJ; Vaughn BV; Witmans M; Woodson BT; Zee P; Mills LE; Hess BJ. Development and growth of a large multispecialty certification examination: sleep medicine certification—results of the first three examinations. *J Clin Sleep Med* 2012;8(2):221-224.

INTRODUCTION

The Sleep Medicine Certification Examination was administered for the third time in November 2011 under co-sponsorship of six member boards of the American Board of Medical Specialties (ABMS)—the American Board of Internal Medicine (ABIM), which is the designated administrative board, the American Board of Family Medicine (ABFM); the American Board of Otolaryngology (ABOto); the American Board of Pediatrics (ABP); the American Board of Psychiatry and Neurology (ABPN); and the American Board of Anesthesiology (ABA). The 2011 examination was the first time that ABA candidates were admitted.

Identical to the previous two administrations of the examination in 2007 and 2009, there were three admission pathways to the examination available to prospective candidates. Pathway

A was self-attestation, subject to possible audit, of the equivalent of 12 months of full-time, post-training practice experience providing clinical care to patients with sleep disorders, accumulated over a maximum of five years prior to examination application. Candidates also had to attest that they (1) evaluated a minimum of 400 individual patients with sleep disorders, (2) interpreted and reviewed raw data of at least 200 polysomnograms, and (3) interpreted and reviewed raw data of at least 25 multiple sleep latency tests. Pathway B required that candidates hold a valid American Board of Sleep Medicine (ABSM) sleep medicine certificate.¹ Pathway C required successful completion of a one-year fellowship in Sleep Medicine. As previously described,² the ABMS approved certification of Sleep Medicine in 2005. At that time, Sleep Medicine practitioners consisted of either individuals certified by the ABSM (Pathway B) or those who were not certified by any formal process (Pathway A). Few

Table 1—Numbers of takers, pass rates, and mean scores for the 2007, 2009, and 2011 Sleep Medicine Certification Examinations

	2007				2009				2011			
	N	% Pass	Mean	SD	N	% Pass	Mean	SD	N	% Pass	Mean	SD
First-taker group	1882	73%	-	-	2140	78%	500	100	2457	65%	466	103
Repeater group	-	-	-	-	371	66%	452	81	456	50%	433	72
Total group	-	-	-	-	2511	76%	493	97	2913	62%	461	100
Standard (minimum passing score)		442				434				440		

The 2011 data exclude a few candidates whose examination outcomes were pending at the time this article was being prepared. Mean scores are not reported for the 2007 candidates because that examination was scored using all 240 questions, whereas in 2009 and 2011 candidates' scores were computed using 200 of the available 240 questions. Modern test theory was used operationally for the first time in 2009 so comparisons between 2011 with 2009 scores are possible, but these cannot be compared with 2007 scores.

had received any formal training in Sleep Medicine, and none of the training programs had been accredited by the Accreditation Council for Graduate Medical Education (ACGME). During initial discussions to plan for the new examination, Pathways A and B were made available in order to allow for current Sleep Medicine practitioners to become certified. However, in 2004, ACGME accreditation requirements for Sleep Medicine fellowships were approved, and shortly thereafter, a number of Sleep Medicine fellowship programs received ACGME accreditation. Thus, it was decided that during these preliminary planning meetings to develop the Sleep Medicine examination, Pathways A and B should be made available for only a limited amount of time—specifically, three examination cycles. This would allow sufficient time for current Sleep Medicine practitioners who had not undertaken fellowship training to become certified by passing the new examination. The 2011 administration marked a milestone in the history of Sleep Medicine because it was the last time that new candidates were admitted to the examination through either the self-attested practice experience pathway or through previous certification by the ABSM. Beginning with the 2013 examination, only physicians who have completed an ACGME accredited Sleep Medicine fellowship (Pathway C) will be admitted to take the examination. However, candidates admitted through one of the other pathways and who were not successful in passing the examination will still be eligible to retake the examination without additional formal training.

EXAMINATION METHODOLOGY

An overview of the process for developing and scoring the examination was previously described.² The test blueprint defining the primary medical content domains for the Sleep Medicine Certification Examination was the same for the 2007, 2009, and 2011 administrations. The ABIM website contains further information about the test blueprint (http://www.abim.org/pdf/blueprint/sleep_cert.pdf) as well as the test development process (<http://www.abim.org/about/examInfo/developed.aspx>). For each of the three administrations, the Sleep Medicine Certification Examination contained 240 single-best-answer multiple-choice questions and was administered by computer. However, beginning with the second (2009) administration, candidates' overall scores were computed using 200 of the 240 questions. This was done to help ensure that each examination was parallel in content and psychometric characteristics to the

prior administration, thus maintaining high-quality measurement characteristics. The additional 40 questions were considered experimental questions, a common practice in the testing industry to identify which questions performed well enough to be retained for future test takers. Because of the large number of candidates who registered for the examination, 2009 also marked the first time that multiple forms of the examination were administered, and that examination was given on two separate days. Modern test theory was used to ensure that candidates' overall scores had the same meaning regardless of the form taken; scores were standardized and reported on a scale with a mean of 500 (SD = 100).

All candidates were held to an absolute content-based standard for passing the examination, rather than a relative standard that is dependent on their performance in comparison with other candidates. A standard was set for each examination year because the composition of the admission pathways changed from the previous administration (e.g., there were fewer candidates admitted to the first 2007 examination through the fellowship training pathway). The standard was established by the committee using the modified Angoff method, a validated and established method for examination standard setting.³ In short, for each examination, the committee discussed the characteristics of minimally qualified or borderline candidates. Next, the 200 questions used for scoring the examination were reviewed, and group members individually identified the expected performance of borderline candidates for each question. Finally, these judgments were systematically combined to derive a minimum passing score for all examinees on the standardized score scale.

EXAMINATION PERFORMANCE AND RESULTS

A total of 2,457 candidates took the 2011 examination for the first time (**Table 1**). This represents a 15% increase in the number of first takers compared with the number of first takers in 2009 (N = 2,140), and a 31% increase compared with the number that took the first 2007 examination (N = 1,882). Taking the 2011 examination for the first time were 1,512 (62%) ABIM candidates, 527 (21%) ABPN candidates, 234 (10%) ABOto candidates, 108 (4%) ABFM candidates, 71 (3%) ABP candidates, and 5 (< 1%) ABA candidates. Except for the inclusion of ABA candidates, the relative proportion of candidates from each Board is consistent with the proportions observed in 2009 and 2007.

Table 2—Number (%) of first takers, pass rates, and mean scores for each admission pathway—2007, 2009, and 2011 administrations

	2007				2009				2011			
	N (%)	% Pass	Mean	SD	N (%)	% Pass	Mean	SD	N (%)	% Pass	Mean	SD
Practice Experience	1034 (55%)	59%	-	-	1215 (57%)	67%	472	102	1638 (67%)	53%	439	105
Certified by ABSM	683 (36%)	93%	-	-	573 (27%)	91%	535	79	526 (21%)	85%	512	74
Fellowship Training	165 (9%)	82%	-	-	352 (16%)	93%	543	83	293 (12%)	91%	531	76

The 2011 data exclude a few candidates whose examination outcomes were pending at the time this article was being prepared. Mean scores are not reported for the 2007 candidates because that examination was scored using all 240 questions, whereas in 2009 and 2011 candidates' scores were computed using 200 of the available 240 questions. Modern test theory was used operationally for the first time in 2009 so comparisons between 2011 with 2009 scores are possible, but these cannot be compared with 2007 scores.

Table 1 presents examination performance data for the first-takers, the repeaters, and the total groups for each of the three administrations. The pass rate in 2011 is lower than the first-taker pass rates observed in 2009 and 2007. Pass rates for the repeaters and the total group were also lower. Compared with the pass rates in 2009, the lower pass rates in 2011 can be attributed primarily to the lower ability of those candidates.

Table 2 presents the number (and percent) of first takers admitted through each pathway and their pass rates and mean scores for the first three administrations. Both in absolute numbers and proportionally, there were fewer first takers admitted through the fellowship training pathway in 2011 (293, 12%) compared with those in 2009 (352, 16%). This is somewhat surprising given that the number of fellowship positions accredited by the ACGME has increased steadily over the past six years. One possible explanation is that the number of candidates classified as fellowship trained in 2009 was inflated by including those who had completed non-ACGME accredited fellowships. Not unexpectedly, there were numerically and proportionally fewer first takers previously certified by the ABSM in 2011 (526, 21%) compared with 2009 (573, 27%) and 2007 (683, 36%). In contrast, the absolute number and proportion of first takers from the self-attested practice experience pathway was much higher in 2011 (1638/67% compared with 1215/57% in 2009 and 1034/55% in 2007). This was expected because 2011 was the last year candidates could be admitted to the examination without fellowship training. Compared with the pass rates in 2009, pass rates decreased for each pathway in 2011, although minimally, for those who completed a fellowship. This was largely due to the lower ability of these first takers compared with first takers in 2009. Specifically, the ability and pass rate of candidates admitted through the practice experience pathway were the lowest; conversely, the ability and pass rate for the candidates admitted through the fellowship training pathway were highest (and similar to those observed in 2009). The pass rates for the practice experience pathway were expected, given that many of these candidates had no formal training and only minimally met the examination entry criteria. The slightly lower pass rate for ABSM-certified candidates was surprising and not readily explainable, but does indicate such candidates who delayed taking the exam, on average had slightly less ability. Nevertheless, ABSM candidates as expected performed markedly better than practice pathway candidates. The strong performance of fellowship-trained candidates not only provides

independent evidence for validity of the examination scores but also supports the relevance of clinical training experienced during an ACGME fellowship.

Two important considerations when assessing the psychometric characteristics of an examination are the reliability of scores and the reproducibility of pass/fail decisions. Score reliability was assessed using the coefficient α . Coefficient α provides an estimate of the amount of variability in candidates' scores that is due to true differences in ability rather than random influences such as guessing. The consistency of the pass/fail decision, which is related to score reliability, is an estimate of the proportion of candidates who would receive the same pass/fail decision if repeatedly tested with equivalent examinations. The coefficient α for each of the three examinations exceeded 0.90, which meets testing industry standards.⁴ This value indicates that the variability in scores is largely due to differences in the true abilities of the candidates. Pass/fail decision consistency for 2011, 2009, and 2007 was 0.88, 0.89, and 0.89, respectively, which indicates that approximately 90% of the candidates that took any one of the three examinations would receive the same pass/fail decision if retested with an equivalent examination.

CONCLUSION

The data presented here indicate that the first three ABMS Sleep Medicine Certification Examinations performed very well. Candidates' scores were reliable, and the resulting pass-fail decisions were consistent. The differences observed in the performance of the candidates from the three pathways provide some evidence for the validity of the examination scores. Furthermore, the overall pass rates observed from each administration were judged to be reasonable, and the performance of each of the three admission pathways was generally as expected.

The administration of the 2011 ABMS Sleep Medicine Certification Examination marked a milestone in the history of the specialty that began with the first Accredited Clinical Polysomnographer examination in 1978. In the 31 years that have elapsed, the practice of Sleep Medicine has evolved to incorporate new knowledge, diagnostic paradigms, and treatment. The practice of Sleep Medicine will continue to change. The challenge in the future for the Sleep Medicine Test and Policy committee will be to continue to make the examination relevant to the practice of Sleep Medicine by expanding and improving the

pool of questions and by eventually incorporating high-fidelity testing methods that better simulate the practice of the specialty, such as using multimedia to include actual polysomnograms in the exam rather than using static illustrations.

REFERENCES

1. Shepard JW Jr, Buysse DJ, Chesson AL Jr, et al. History of the development of sleep medicine in the United States. *J Clin Sleep Med* 2005;1:61-82.
2. Quan SF, Berry RB, Buysse D, et al. Development and results of the first ABMS subspecialty Certification Examination in Sleep Medicine. *J Clin Sleep Med* 2008;4:505-8.
3. Angoff WH. Scales, norms, and equivalent scores. In: Thorndike RL, ed. Educational measurement. Washington, DC: American Council on Education, 1971:514-5.
4. Subkoviak MJ. A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *J Educ Meas* 1988;25:47-55.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication February, 2012

Accepted for publication February, 2012

Address correspondence to: Stuart F. Quan, M.D., Division of Sleep Medicine, Harvard Medical School, 401 Park Dr. 2nd Floor East, Boston, MA 02215; Tel: (617) 998-8842; Fax: (617) 998-8823; E-mail: squan@arc.arizona.edu

DISCLOSURE STATEMENT

This was not an industry supported study. Dr. Quan is Editor-in-Chief of Journal of Clinical Sleep Medicine. Dr. Buysse has consulted for Merck, Philips, and Transcept and has participated in speaking engagements for Servier and Astellas. Dr. Iber has consulted for Apnex medical. Dr. Zee has consulted for Sanofi-Aventis, Merck, Philips-Respironics and Purdue. Dr. Hess and Ms. Mills are employees of the American Board of Internal Medicine, the organization responsible for the development, administration, and scoring of the Sleep Medicine Certification Examination. The other authors have indicated no financial conflicts of interest.