

# From Specialty-Based to Practice-Based: A New Blueprint for the American Board of Family Medicine Cognitive Examination

Thomas E. Norris, MD, Richard J. Rovinelli, PhD, James C. Puffer, MD, Jason Rinaldo, PhD, and David W. Price, MD

**The American Board of Family Medicine (ABFM) is committed to offering cognitive examinations that are both pertinent to the specialty of family medicine and psychometrically sound. This article reviews the history of the development of the blueprint of the ABFM certification and recertification cognitive examinations and describes the creation of a new one. The design of the new blueprint represents a significant change. The intention of the new plan is to create a continuously evolving approach that will assure family physicians that the content of their specialty board certification/recertification examination is relevant to their practices and to the discipline. The ABFM anticipates that assessments based on the new blueprint will assist family physicians in attaining and maintaining the knowledge required to practice high quality family medicine by focusing their certification and recertification examinations and, therefore, studies for those examinations on material that is relevant to their practices. (J Am Board Fam Pract 2005;18:546–54.)**

## Background

Construction of high stakes cognitive examinations that are psychometrically sound is usually dependent on the utilization of a blueprint or design for the examination.<sup>1–3</sup> The American Board of Family Medicine (ABFM) has, for over 20 years, used a “content blueprint” that represents areas of practice in the discipline of family medicine and that defined both the subject areas and the proportion of questions in ABFM certification and recertification examinations. Norcini and colleagues<sup>4</sup> have described the essential structure of this type of content-based definition of medical practice. The old basic ABFM content blueprint is shown in Table 1.

The development and implementation of Maintenance of Certification for Family Physicians (MC-FP) provided an impetus for the ABFM to

review and evolve the Family Medicine Board Examination. Maintenance of Certification for all specialty boards under the American Board of Medical Specialties umbrella includes a secure cognitive examination as one of its components. In early 2003, the ABFM, in consultation with a cadre of family medicine educators, determined that the cognitive examination Content Blueprint in use at that time required revision. The goals of the decision were to adjust the test to the maturation of the field of family medicine as more than the summation of parts of other fields, to connect board certification to the quality movement in the United States, and to maintain fidelity in a defensible process. The ABFM also anticipated introducing computerized examinations and offering the examinations multiple times each year, which would require generating additional test items to maintain security and validity. Both changes also required precise assignment of each test question to appropriate categories and sub-categories in “test item banks” to accurately retrieve questions and assign them to an examination based on the blueprint.

Equally important in the decision to revise the original blueprint was the planned move in 2006 from a classical test theory psychometric model to one that will rely on item response theory (IRT) (see definitions in Table 2). IRT offers advantages

Submitted 29 June 2005; revised 13 September 2005; accepted 5 October 2005.

From University of Washington School of Medicine, Department of Family Medicine, Seattle, WA (TEN); American Board of Family Medicine, Lexington, KY (RJR, JCP, JR); and University of Colorado School of Medicine, Boulder, CO (DWP).

Conflict of interest: none declared.

Corresponding author: Thomas E. Norris, MD, University of Washington School of Medicine, Department of Family Medicine, UW Box 356340, Seattle, WA 98195-6340 (e-mail: [tnorris@u.washington.edu](mailto:tnorris@u.washington.edu)).

**Table 1. Old Approach: Specialty-Based ABFM Cognitive Examination Question Distribution**

General Medical Discipline	Percentage of Representation
Internal medicine	36
Surgery	6
Obstetrics	7
Community medicine	9
Pediatrics	13
Psychiatry	7
Geriatrics	12
Gynecology	10

over classical test theory, including the ability to fashion an examination that focuses measurement on a selected level of knowledge, rather than a comparison of persons across tests and test items. This allows the administration of multiple tests each year with better assurance of comparability between the multiple tests. In the case of the ABFM examination, it would be very useful to have very precise measurement in the region of the cut score (ie, pass fail score). Because the current examination is a relatively random sample of items from the examination pool, its measurement precision is probably to be distributed normally around the average level of difficulty of the items in the pool.

This means that it measures most precisely at that level, rather than at the level of the passing threshold. This method of measuring knowledge is less effective than desirable or achievable.

IRT provides a better method of equating, or comparing and adjusting, the difficulty of different versions of the examination. This has practical importance especially for those candidates who take 2 or more consecutive examinations. The use of a more robust and well-grounded theoretical measurement model such as IRT improves the equating procedure. Therefore, a more precise blueprint was required to guarantee that test items were appropriately categorized by content as well as by difficulty as measured by IRT within specific content areas.

Throughout this article, a number of terms will be used that may not be familiar to the reader. Definitions of these terms are provided in Table 2.

In the remainder of this article, we will outline the process used to develop and define a new examination blueprint, elucidate the specific blueprint components, and explore implementation plans and future directions.

### The Development Process

The ABFM had conducted 3 validity studies of its specialty-based examination content blueprint over

**Table 2. Definition of Terms**

Term	Definition
Classical test theory	A theory about test scores, which postulates that a person's test score can be modeled as the sum of two unobservable variables, a true score and an error score where with certain assumptions a true score can be shown to be the expected score across parallel test forms. A limitation of the classical test theory model is that both person scores and item statistics are dependent on the test and the sample of examinees, respectively.
Item response theory	A modern statistical theory for test development and scoring that improves the understanding of both individual and item performance. It is based on the concept that both test items and individuals can be simultaneously described in terms of their standing on, or relationship to, one or more ability/knowledge level scales. An individual is characterized by his or her position on the scale, and an item can be characterized by the point on the scale indicating the trait or ability level at which a person would have a 50% chance of responding correctly to the item.
Examination blueprint	A table of specifications based on the knowledge domain of family medicine, which is weighted by a judged frequency of use and clinical importance. It is used to determine the scope and depth of material covered in the examination process.
Equating	A statistical process by which the score scale on one test is converted to the score scale of another test so that the scores are equivalent or parallel.
Certifiably competent	A description of a candidate who has scored at or above the cut score level on a certification examination.
Cut score	A specified point on a score scale at or above which a candidate passes and below which a candidate fails.
Dimensions/orders	A categorization of the domain of knowledge in family medicine, which serves as an organizing agent for the content blueprint.
Adaptive testing	A sequential form of testing in which the order of items presented is determined by the examinee's response to previous items.

the 35 years of its existence. The first study, conducted in 1982 by researchers from the University of Massachusetts, used a task analysis process that identified the knowledge, skills, and abilities of practicing family physicians. Burg and associates<sup>5</sup> described this approach in a 1976 article. In 1993, a validity study, including surveys of family physicians' patient mixes and clinical experiences, as well as a review of data from the National Ambulatory Medical Care Survey (NAMCS) was conducted. Unfortunately, survey return rates were low, and the results were questionable. In 1999, another patient mix study was conducted that, despite another low return rate, produced data consistent with the 1993 study.

In 2003, the ABFM Examination Committee was charged with the task of developing a new content blueprint and served as the steering committee for a larger task force that would consist of organizational representatives from the "family" of family medicine organizations. Organizations represented included the American Board of Family Practice (now ABFM), the American Academy of Family Physicians, the Society of Teachers of Family Medicine, the Association of Departments of Family Medicine, and the Association of Family Medicine Residency Directors. The task force served as an external validation expert consensus panel for redesigning, updating, and validating the content blueprint for the certification and recertification examinations. The ABFM anticipated that the blueprint would also help inform strategies for the lifelong learning and self-assessment components of MC-FP.

The ABFM Examination Committee and the Blueprint Task Force met in late February and early June 2003 to work on the project. Since the Task Force meetings, the ABFM Examination Committee has continued to evolve the blueprint by meeting in person, by phone, and by e-mail discussions.

## The New Examination Content Blueprint

### *A Practice-Based Design*

Based on discussion and the results of a study of new blueprints developed by the American Board of Emergency Medicine and others,<sup>6,7</sup> the Examination Committee and the Task Force concluded that the present specialty discipline-based structure of the current ABFM Content Blueprint did not

represent the current (and evolving) content of family medicine and was inadequate for creating a larger, better categorized, bank of test items. Interestingly, a new categorization of test questions was studied and considered by Pisacano (the founder of the ABFP) and others as early as 1986, but was not adopted.<sup>8</sup>

The Task Force endorsed a multidimensional approach to the content blueprint. Multiple data sources will be used to categorize, as precisely as possible, the content of family medicine. These will include surveys of practice content, data collected by analysis of International Classification of Diseases, Ninth Revision (ICD-9) and Current Procedural Terminology (CPT) codes derived from electronic medical records of practicing family physicians, National Ambulatory Medical Care Survey data,<sup>9</sup> as well as other means. This task will be continuous, and the "content" definition will remain relatively current because it is tied to the evolving practice patterns. The unit of analysis for this effort will be the practice of the individual family physician. At this time, the examination will be targeted at the average family physician, recognizing that over time there will be changes in the practice of the average family physician that will be reflected in the examination. Hopefully, at some point in the future, the examination might be tailored to the practices of individual family physicians who are taking the test.

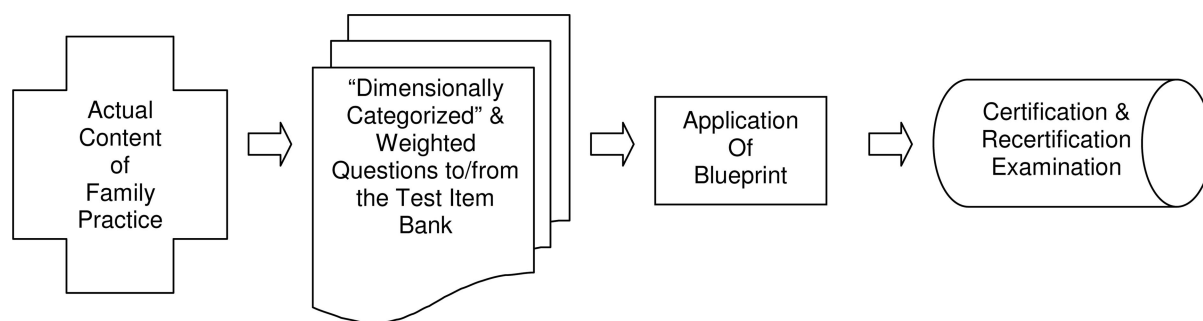
### The Dimensions/Orders System

Each examination question will be coded and placed in question banks. This will allow retrieval of questions based on categories needed to match the examination blueprint, and allow question writers to be given clear specifications for writing new questions that would fit into categories of need.

The categorization of questions that will be derived from the content of family medicine will be based on dimensions (also known as orders or domains) (Figure 1).

#### *First Dimensions/Orders*

Each question will be assigned to a first dimension category. The vast majority of questions will fall into an organ system category, with other first dimension categories for questions not amenable to categorization in an organ system category. In some ways, the organ system categorization ap-



**Figure 1. Conceptual diagram of the application of a content-based blueprint.**

proach is similar to that described by Pisacano et al in their 1989 article.<sup>8</sup>

At present, we anticipate that 90% of the test questions will fall into the organ system category. The other first dimension categories include population-based care and health systems (5% of the questions) and patient-based care and systems (5% of the questions).

The organ systems that will be included in the categorization are as follows

### **A. Organ Systems (90% of examination questions)**

- Respiratory
- Cardiovascular
- Musculoskeletal
- Gastrointestinal
- Special sensory (visual, hearing, etc)
- Endocrine
- Skin
- Nervous system (brain, spinal cord, peripheral nervous system)
- Psychogenic (psychological, behavioral, mental health)
- Reproductive (male, female)
- Renal/urinary tract
- Blood/immune system
- Nonspecific

### **B. Population-Based Care and Health Systems (5% of examination questions)**

- Health policy
- Bioterrorism
- Legal
- Epidemiology
- Biostatistics

- Evidence-based medicine
- Quality improvement
- Informatics

### **C. Patient-Based Care and Systems (5% of examination questions)**

- Physician-patient interactions
- Communication
- End of life care
- Palliative care
- Family issues
- Cultural issues
- Clinical decision making
- Evidence-based medicine
- Ethics

#### ***Other Dimensions/Orders***

In addition to the first dimension categorization for all questions, all questions will eventually also be tagged or defined based on a number of other dimensions. Not all questions will have other dimensions, and some may have multiple dimensions. It will be possible to have a “not applicable” tag for a specific question if another dimension does not fit. Tagging of questions in a computerized searchable database will allow a very precise level of definition of both existing questions within the question pool and help identify areas in which additional questions should be written. Other dimensions conceptualized at this time are shown in Table 3. We anticipate these dimensions will eventually be used to more precisely refine the specific questions that fall into each primary dimension. For example, in the cardiovascular primary dimension, the age and gender dimensions might be applied to assure that the most commonly asked questions deal with patients whose ages and genders

**Table 3. Other Dimensions in the Categorization System**

Age dimension	Etiology/origin dimension
Fetus	Autoimmune/allergic
Newborn	Behavioral/lifestyle/cognitive
Infant	Degenerative
Child	Environmental
Adolescent	Genetic
Adult	Iatrogenic
Young	Idiopathic
Middle aged	Infectious
Senior	Neoplastic
Old	Nutritional
Very old	Pregnancy
Sex/gender dimension	Toxicologic
Male	Traumatic
Female	Undifferentiated
Populations dimension	Vascular/ischemic
Immigrant/refugee	Anatomic/obstructive/ surgical
Ethnicity/race	Normal
Disabilities	Anatomy
Gay/lesbian/bisexual/transgender	Physiology
Socioeconomic	Growth and development
General	Diagnosis and management dimension
Continuum of care dimension	Diagnosis
Prevention	History
Primary	Physical
Secondary	Laboratory
Emergent	Imaging
Acute	Diagnostic procedure
Chronic	Other
Early	Management
Advanced	Pharmaceutical
	Procedural/surgical
	Behavioral/psychotherapy
	Complimentary and alternative medicine
	Other

match those actually seen in patients with cardiovascular disease.

### Weighting

The old content blueprint has assigned percentages of the questions into discipline-defined areas. For example, 36% of the questions currently deal with internal medicine topics. These weightings were obtained from past studies of the proportion of the average family medicine practice consisting of internal medicine, pediatrics, and so on. Different

data are needed to use the new “multidimensional” blueprint. It will be possible to use actual practice content data and to assign the proportions of questions in each blueprint area more accurately. In an effort to begin compiling practice content data, the ABFM has conducted content of practice surveys of diplomates who took ABFM examinations in 2003 and in 2004. In addition, the ABFM has obtained family medicine data from the recent NAMCS surveys and compared that data to the data obtained in the 2003 and 2004 surveys of family physicians. NAMCS data were selected for this comparison because of its focus on ambulatory care, the most frequent site of family medicine encounters. This comparison required some realignment of the categories. The data from the 2003 and 2004 surveys for organ systems are shown in Table 4. The percentage of the patient visits seen by the family physicians who were surveyed are listed by organ systems, from most frequently seen to least frequently seen. Data from population-based categories and patient-based categories from the ABFM surveys are displayed in Tables 5 and 6.

To provide a cross check of the accuracy of the practice content surveys, we interpolated the practice data from the NAMCS study,<sup>9</sup> the organ system categories in the ABFM surveys and blueprint, and compared the data with the ABFM surveys. To make the data more comparable, the 2 ABFM survey results were averaged and the mean percentage of practice content by organ system was compared with the percentage of practice content by organ system found in the NAMCS data. This information is displayed in Table 7. Based on the overall similarity seen in the ABFM survey data and the NAMCS study data, the Board concluded that the “percentage of practice content” findings are reasonable.

Although the majority of the new blueprint is based on analysis of practice content, some decisions have been made by consensus opinion of the examination committee due to lack of data. The ABFM Examination Committee has determined that 90% of the questions will be organ system based, and 5% each will be assigned to patient-based and population-based care systems. Therefore, to determine the actual percentage of organ system-based test questions to assign to each category, the totals found in Table 7 were adjusted from 100% to 90%. This information, illustrated in Table 8, shows the actual percentage of questions

**Table 4. Frequency of Presenting Problem by Organ System\***

	N	Mean	SD	Relative Percentage of Practice	
				2004	2003
Cardiovascular system	807	4.28	0.943	16.7	17.0
Musculoskeletal system	807	3.66	0.895	12.2	13.0
Respiratory system	807	3.59	0.815	11.7	10.8
Psychogenic (mental health/behavior)	807	3.46	0.953	10.7	11.3
Endocrine system	807	3.34	1.034	9.8	8.0
Gastrointestinal system	807	3.29	0.720	9.5	10.4
Reproductive, female	807	2.75	1.048	6.1	5.3
Integumentary system	807	2.58	0.724	5.3	5.5
Nephrological/urological	807	2.39	0.607	4.4	4.7
Neurological system	807	2.35	0.615	4.2	4.0
Nonspecific	807	2.20	0.935	3.4	3.5
Hematological/immunological	807	2.14	0.521	3.1	2.4
Sensory system	807	2.09	0.548	2.9	2.4
Reproductive, male	807	2.04	0.523	2.6	1.7
Total	807			100.0	100.0

\* Higher scores indicate higher frequency of exposure to issue.

related to each organ system that will be assigned on the cognitive examination starting in 2006.

Several additional decisions will need to be made by the ABFM Examination Committee and the Board regarding the content of questions within each organ system. It is clear that frequency of occurrence alone cannot drive the weighting process. Criticality, complexity, and other factors will play a role. For example, Table 7 suggests that 13% of the questions on the examination should be based on the respiratory system. 2002 NAMCS disease-specific data showed that nearly one quarter of the respiratory disease seen by family physicians

consists of viral upper respiratory tract infections (URIs). It seems imprudent to have 25% of all respiratory questions (nearly 4% of the overall examination) focus on the topic of viral URIs. Therefore, some adjustments will probably be made to cover important conditions, which, although not as frequently seen by family physicians, are important areas for recognition, diagnosis, and management or referral.

In addition, at some point in the future the Board will need to make policy decisions about whether the “weights” will be reflective of a standard or average family medicine practice, or

**Table 5. Population-Based/Health Systems-Based Issues\***

	N	Mean	SD
Evidence-based practice	807	2.07	1.049
Quality improvement	807	2.59	0.884
Geographic (urban/rural)	807	2.82	1.065
Informatics	807	3.04	1.075
Health policy	807	3.09	0.941
Legal	807	3.13	0.893
Epidemiology	807	3.27	0.852
Biostatistics	807	3.52	0.801
Bioterrorism	807	3.88	0.414

\* Lower scores indicate higher frequency of exposure to issue (1 = daily, 2 = weekly, 3 = monthly, 4 = almost never)

**Table 6. Patient-Based Care and Systems Issues: Highest to Lowest Strength of Endorsement\***

	N	Mean	SD
Clinical decision-making	807	1.09	0.428
Communication	807	1.17	0.533
Doctor-patient interactions	807	1.21	0.599
Family issues	807	1.56	0.730
Evidence-based practice	807	1.80	0.962
Cultural issues	807	2.00	0.991
Ethics	807	2.39	1.013
End-of-life care	807	2.62	0.868
Palliative care	807	2.71	0.895

\* Lower scores indicate higher frequency of exposure to issue (1 = daily, 2 = weekly, 3 = monthly, 4 = almost never).

**Table 7. Percentage of Practice by Organ System NAMCS (2002) versus ABFM (2003/2004)**

Organ System Category	NAMCS Annual National Frequency	NAMCS 2002 1st Diagnosis Percentage	ABFM 2003 and 2004 Survey of Practice Percentage
Cardiovascular	20,749,525	10.4	16.7
Endocrine	17,363,399	8.2	9.8
Gastrointestinal	12,387,729	5.1	9.5
Hematologic/immune	4,625,960	2.8	3.1
Integumentary	15,068,757	7.0	5.3
Musculoskeletal	31,860,955	13.9	12.2
Nephrologic/urologic	4,680,860	2.1	4.4
Neurologic	3,341,277	2.0	4.2
Psychogenic/behavioral	8,140,406	4.4	10.7
Reproductive, female	6,291,831	3.4	6.1
Reproductive, male	1,537,156	0.6	2.6
Respiratory	36,844,808	16.9	11.7
Special sensory	8,221,433	1.1	2.9
Nonspecific	35,958,364	16.9*	3.4*
Not applicable to ABFM's content blueprint	1,348,820	0.8	
Total	208,421,280	100.0	
Correspondence between NAMCS and ABFM survey of practice	$r = 0.78^*$		

\* The nonspecific category was not included in the correlation between NAMCS and ABFM percentages.

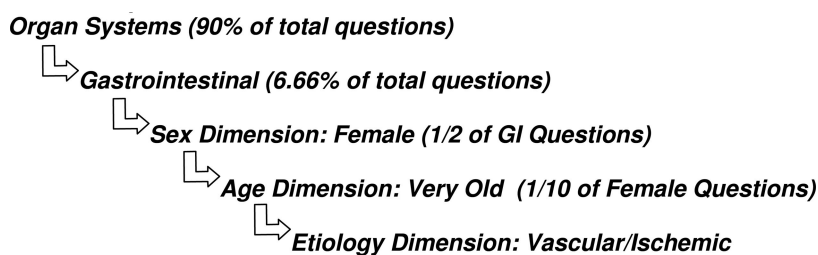
whether to be more tailored: for example, to have a different weighting for the examination of a family physician whose practice does not include prenatal or antenatal care. It is conceivable that eventually, using data generated from electronic health records systems, we might be able to tailor individual examinations to reflect individual practices (recognizing that some conditions are important for family

physicians to recognize and evaluate, even if they are not seen frequently). Even if this is technically feasible, there are many policy questions and testing validity questions that must be answered first.

An example of the specific weighting that will be possible in the new system is shown in Figure 2—note that with the exception of the organ system and gastrointestinal percentages, those used

**Table 8. Percentage of Examination Items per Organ System, Based on 90% of the Examination Being Composed of Organ System Categories**

Domain	NAMCS	ABFM Survey Averages	NAMCS/ABFM Averages	90% of Average
Cardiovascular	10.4	16.7	13.55	12.36
Endocrine	8.2	9.8	9.00	8.21
Gastrointestinal	5.1	9.5	7.30	6.66
Hematologic/immune	2.8	3.1	2.95	2.69
Integumentary	7	5.3	6.15	5.61
Musculoskeletal	13.9	12.2	13.05	11.90
Nephrologic/urologic	2.1	4.4	3.25	2.96
Neurologic	2	4.2	3.10	2.83
Nonspecific	16.9	3.4	10.15	9.26
Psychogenic/behavioral	4.4	10.7	7.55	6.88
Reproductive, female	3.4	6.1	4.75	4.33
Reproductive, male	0.6	2.6	1.60	1.46
Respiratory	16.9	11.7	14.30	13.04
Special sensory	1.1	2.9	2.00	1.82
			<b>98.70</b>	<b>90.00</b>



**Figure 2.** Example of use of the blueprint to select a question based on content of practice data.

are purely for the sake of example—they are not based on actual practice content data, because at this time, we do not have actual data for other than the first orders or dimensions.

### Complexity or Depth of Knowledge

Each ABFM examination question is currently rated on complexity and the depth of knowledge required to answer the question. One concept used in the discipline-based blueprint was that family physicians should possess the level of knowledge required to manage patients “up to the point of referral to specialty care” in the discipline to which the question was assigned. This discipline-based approach will not work well with the new practice content-based system. The new paradigm adopted by the ABFM Board is to adjust the difficulty of the examination to the level that would help better differentiate a baseline level of knowledge possessed by “certifiably competent” family physicians compared with those who are not certified. Using this concept with an item response theory approach to the analysis of the level of difficulty of individual questions and the entire test, Board editors and psychometricians will be able to construct examinations that reliably differentiate family physicians with a certifiably competent level of knowledge from those whose level of knowledge is not sufficient for board certification.

### Implementation

To establish a valid measure of family medicine knowledge, the definition of what is to be measured (represented by the blueprint) must be used to generate specific items that can be reasonably expected to comprise interpretable subscales. Our historical, post hoc classification of items into subscales results in a group of weakly related items being lumped together and treated as if they are measuring a coherent sub domain (usually based on

a keyword appearing in the item stem), when they were not created to do so. Thus, our plan at this time is to follow standard test development practice with subscales implemented at the front end during item writing.

### Future Directions

The new blueprint offers opportunities to pursue adaptive testing and physician-specific testing formats. Adaptive testing is an examination format using a minimal number of representative questions that become progressively more difficult, until the candidate begins to answer most questions incorrectly, at which point the test is terminated. This type of testing could allow the creation of psychometrically valid cognitive examinations that would require far less time to complete. “Physician-specific testing formats” refers to tests created specifically for individual physicians, based on the actual content of their practices. The ABFM will need to decide if the relative weighting of future examinations will be reflective of a “typical” family medicine practice, or if all or part of the examination will be somewhat tailorable to reflect practices emphasizing some content areas but excluding others. When electronic health records become widely accepted, we may well have the ability to “data mine” specific physician practices and create individualized examinations. However, even if this becomes technically possible, many policy and testing validity questions would need to be considered before any implementation.

### Conclusion

Brennan and colleagues<sup>10</sup> assert that a physician’s certification status should be among the key evidence-based measures used in the quality movement. Their work demonstrated that most patients would change physicians if their current physician did not maintain board certification. Specialty cer-



tification should inextricably link quality, specialty board certification, and the actual practice of medicine. Thus, it is imperative for the certification and recertification process to retain its relevance to practice and to the care that is actually being provided to patients by practicing physicians. In an effort to build a dependable, flexible, and enduring bridge between the cognitive certification/recertification examinations and the actual practices of family physicians, the ABFM Blueprint Task Force proposed a new content blueprint for the ABFM certification and recertification examinations. This new blueprint has been formally accepted and adopted by the ABFM Board of Directors. Beginning in July of 2006, this new approach will be used to provide the design structure for future ABFM cognitive examinations. Among the many strengths of the new blueprint is its ability to continuously evolve in structure, based on the actual and ongoing evolution of family medicine practices.

Charles Darwin pointed out that, "It is not the strongest of the species that survives, not the most intelligent, but the one most responsive to change." Family medicine has shown responsiveness to change, and our certification examinations must mirror these changes to maintain their relevance to practice, the public, and their links to the quality of care.

## References

1. Bridge PD, Musial J, Frank R, Roe T, Sawilowsky S. Measurement practices: models for developing content-valid student examinations. *Med Teach* 2003; 25:414–21.
2. Billi JE, Gruppen LD, Boxer GJ, et al. The development and evaluation of new versions of the written examination for the American Heart Association Advanced Cardiac Life Support provider course. *Ann Emerg Med* 1994;24:26–31.
3. Ropka ME, Norback J, Rosenfeld M, Miller C, Nielsen B. Evolving a blueprint for certification: the responsibilities and knowledge comprising American professional oncology nursing practice. *Oncol Nurs Forum* 1992;19:745–59.
4. Norcini JJ, Lipner RS, Benson JA, Webster GD. An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. *Ann Intern Med* 1985;102:385–9.
5. Burg FD, Brownlee RC, Wright FH, et al. A method for defining competency in pediatrics. *J Med Educ* 1976;51:824–8.
6. Raymond MR. A practical guide to practice analysis for credentialing examinations. *Educational Measurement: Issues and Practice* 2002;Fall:25–37.
7. Raymond MR. Job analysis and the specification of content for licensure and certification examinations. *Appl Measurement in Edu* 2001;14:369–415.
8. Pisacano NJ, Veloski JV, Brucker PC, Gonnella JS. Classifying the content of board certification examinations. *Academic Medicine* 1989;March:149–54.
9. Woodwell DA, Cherry DK. National Ambulatory Medical Care Survey: 2002 summary. *Adv Data* 2004;26:1–44.
10. Brennan TA, Horwitz RI, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification in the quality movement. *J Am Med Assoc* 2004;292:1038–43.