

American Board of Physician Specialties Examinations Development, Administration, and Reporting

The American Board of Physician Specialties (ABPS) develops examinations for each of the medical specialties represented by their Boards of Certification. The examinations development process is a collaborative one, involving individuals with psychometric expertise and ABPS Diplomates, who serve as the subject content experts.

For each specialty ABPS develops a written certification examination consisting of multiple-choice questions. For some specialties ABPS also develops oral examinations that provide candidates an opportunity to demonstrate their knowledge and skills in diagnosing and treating patients by permitting the candidate to probe for information, develop differential diagnoses, and request laboratory tests or diagnostic examinations.

To be certified in the surgical specialties candidates must also pass an on-site, clinical examination. In collaboration with its Diplomates, ABPS has developed rating forms which are used by each member of the examinations team. The information recorded on these forms provides documented evaluations of the candidate's performance in performing surgical procedures.

ABPS also develops multiple-choice recertification examinations. ABPS Diplomates are required to recertify every eight years.

Valid Assessment Instruments

The validity of the examination process is paramount in addressing psychometric and legal standards for certification examinations. ABPS addresses the various components required for ABPS certification in assessing a representative set of appropriate knowledge and skills. Also, ABPS uses assessment modes that obtain responses from examinees reasonably approximating how physicians utilize the knowledge and skills in the practice of their specialty, with consideration for practical constraints of time and resources for the examinations process. Validation of credentialing tests depends mainly on content-related evidence. In addition to content validity, ABPS emphasizes "process validation," that is, the questions posed to examinees mirror the types of tasks, thought processes, and decisions that are actually a part of medical practice. For example, a question that addresses strictly content validity might require an examinee to demonstrate knowledge of units of radiation. A question that addresses both content and process might require an examinee to prescribe the units of radiation based upon information regarding the patient's condition and other relevant information.

By emphasizing the combination of content and process, every question on ABPS examinations represents an assessment of knowledge or skills clinically relevant to the practice of the medical specialty. ABPS emphasizes avoiding the inclusion of questions assessing trivia or recall of information that is not required as part of the actual practice of the medical specialty.

Practice Analysis

All ABPS examinations are based on practice analyses. Meetings with subject matter experts, e.g., ABPS Diplomates, lead to the identification of technical knowledge/skills (content) and tasks (processes) essential to the performance of a certified physician. This step is followed by an external validation of knowledge, skills, and tasks, typically through

reviews by a representative sample of professionals in the field, usually certified by the American Board of Medical Specialties (ABMS), The American Osteopathic Association (AOA), or ABPS. Through this process ABPS determines such things as the importance, criticality, and frequency of the identified knowledge, skills, and tasks required in practice. This information serves as the cornerstone of the content and process validity of the examinations.

Test Specifications (Examination Blueprints)

A set of specifications, or blueprint, for each examination or set of examinations is developed. For specialties requiring more than a written examination, ABPS takes advantage of the unique aspects of assessment available by any one mode, avoiding repetition of the same type of assessment for different modes of testing. Multiple-choice examinations do not lend themselves well to the demonstration of performing a task, e.g., performing surgery. Such a demonstration of performance is most likely to be observed through a clinical examination. An oral examination can serve as a pseudo-clinical examination, with the expectation that examinees will create answers or solutions, rather than merely identifying them as in a multiple-choice examination. On an oral examination an examinee is expected to explain a procedure but is not expected to demonstrate the procedure. Oral examinations focus on determining whether an examinee understands medical procedures, demonstrates appropriate thinking processes, asks for appropriate information at suitable times, and integrates complex sets of information. A unified approach to the development of the blueprints provides for a more valid, effective, and efficient assessment process.

The development of blueprints requires the collaboration of persons with psychometric expertise and those with the medical knowledge. The psychometrician challenges the physicians as to what skills, knowledge, and tasks can be assessed by each mode of assessment. The physicians provide examples of situations, scenarios, and questions that will apply to each examination mode.

The integrated blueprint approach allows the collaborating team to identify any "holes" in the assessment (knowledge, skills, and tasks identified in the practice analysis but not yet addressed by the assessment system) and determine in which mode assessment should occur.

The test specifications include details regarding the weighting of each content category. For example, for the written examinations, the test blueprint indicates the approximate number of questions for each category and subcategory, based upon the data from the practice analysis. The number of topics to be covered and the number of questions needed to provide reliable diagnostic information in each content category contribute to decisions regarding the number of questions that will appear on an examination.

For the written certification examinations the number of questions on an examination range from 150 to 350 with most of the examinations having 200 items. The oral examinations typically consist of two to six cases, in each of which the examinee is provided a scenario, i.e., a patient situation, and is asked how the physician wishes to proceed.

Item Development

Developing the test questions is an iterative, collaborative process. Using the test blueprints as a map, a variety of questions are developed for the examinations. Questions are developed to meet content and task categories. The initial item is designed by a subject matter specialist, typically an ABPS Diplomate but may also be a physician certified by

ABMS or AOA, or have particular knowledge or expertise relevant to the examination being developed. Items are reviewed by an editor, who revises the items with attention to item construction rules, readability, grammar, and potential bias. The items are entered in an electronic item bank from which they can be drawn to construct test forms. ABPS uses a banking system that allows for the storage of the text of item, corresponding graphics, descriptive information, and item statistics.

Examination Forms Development

Using the test specifications and the items available in the item bank, ABPS staff configures a form of the examination for each specialty. Subject matter experts identify items that are too similar to appear on the same test for and items that might provide clues to the answer for another item. These items are electronically flagged in the item bank so that they do not appear on the same form of the examination. Additionally, an editor reviews the examination form for consistency of style, embedded directions, arrangement of items, and inclusion of all necessary components for a complete form. The answer key is reviewed to verify that there is an approximately equal distribution of each answer choice (equal number of A's, B's, etc.). As needed, items are moved to different positions on the form or replacement items are selected from the item bank. The test form is then printed with quality control checks performed by ABPS staff to verify the accuracy of the printed product.

Administration of the Examinations

Examinations are administered under standardized conditions. For the written examinations ABPS produces an Administration Manual containing specific instructions and scripts for examination administrators and proctors. For oral examinations, specific scripts following a standardized outline are developed for the examiners. ABPS trains the examiners regarding appropriate administration procedures. For the on-site, clinical examinations, members of the examinations teams follow a structured rating form.

Although written examination items and tests forms undergo continuing reviews, the accuracy of an item can immediately change due to a change in accepted medical practice. To identify such changes, ABPS has instituted a formal procedure by which examinees can comment on the items. As they are taking the tests, examinees have the opportunity to identify items about which they have a concern, question the accuracy of a correct answer, or suspect multiple correct answers.

Examinees are advised that all items so identified and their comments will be reviewed before ABPS does the final scoring and reporting of results.

As with the written examinations, the content and scoring of the oral examinations undergo reviews prior to each administration to verify accuracy and currency.

Review of Comments and Item Analyses

After administration of the written examinations, the examinees' responses are initially scored. Item statistics are generated from the data. The results of the item analysis and a compilation of examinees' comments are provided to the Examinations Committee. The examinees' responses are rescored based upon the recommendations of the reviewers. The results are used to develop the score reports for examinees and examination analysis reports.

Establishing Passing Scores

ABPS uses variations of the Angoff method to establish passing scores on its written and oral examinations. The Angoff process requires judgments from subject matter specialists. Again, ABPS relies on its Diplomates. For example, for the written examinations, participating Diplomates are asked to specify the percent of minimally proficient physicians in the specialty that would know the answer to each question. The average value or consensus of the participating physicians is obtained for each item. The average of the Angoff values for each item used on an examination provides an estimated passing score. The distribution of scores is analyzed. If the examination form is being used for the first time, the examination analyses, score distribution, and resulting outcome (pass/no pass rates) are reviewed by the Board of Certification to determine whether any variation from the initial passing score should be made. Boards of Certification review data, maintaining the anonymity of the examinees.

For the oral examinations the results are reviewed by the Examinations Committee to verify correct administration and scoring, e.g., comparing the individual ratings and notes of the examiners. The data are statistically analyzed before examination results are released. For example, for the Emergency Medicine oral examinations, the difficulty of the cases presented is analyzed to verify that each examinee has received a set of cases of comparable difficulty to those received by other candidates. The analyses are reviewed by the Examinations Committee to determine whether a reduction in the passing score for a case is warranted.

Scoring and Reporting

All responses by examinees to items on written examinations are processed by electronic scanning. The data are processed independently by two different scoring programs to verify accuracy. The response form for an examinee is visually reviewed if data indicate the candidate provided multiple-marks for a multiple-choice item. If the multiple-marking flag was a result of a poor erasure or a smudge on the answer document, the candidate's record is manually changed to reflect the examinee's intended response and the candidate's responses rescored.

Manual rescoring of an examinee's responses is available for a fee. If the rescoring results in a change in the outcome of the examination results, the ABPS refunds the fee to the candidate.

Follow-up audits of scoring procedures and examination results may be conducted. If an audit indicates an error in scoring resulting in a change in the results of a candidate, ABPS will rescore and provide the candidate the revised results.

Recertification Examinations

The written examinations required for recertification every eight years are drawn from the same bank of items as the certification examinations. Because the item banks contain questions that are relevant to the medical practice in that specialty, the items are clinically relevant and applicable to those who have possibly had more years of experience than physicians who are typically applying for initial certification.

The recertification examinations are assessing skills and knowledge applicable to acceptable medical practices in the specialty. These examinations do not address "cutting-edge" changes in medicine but rather address established practices that remain current. Realizing the need for competent physicians to be abreast of the latest developments in medicine, the ABPS has adopted a phased-in set of new recertification requirements in which physicians

are also required annually to complete self-assessments in their specialty. Physicians may select self-assessments to address their own determined needs and are encouraged to select areas that address current day issues, medical research, and possible changes in medical practice.