Communications

Core Content In-Training Assessment: Another Look

Marian R. Stuart, PhD, Samuel W. Warburton, MD, and Frank C. Snope, MD
Piscataway, New Jersey

Assessment of cognitive knowledge is an important and integral part of the Quality Assurance Program of the Department of Family Medicine at the College of Medicine and Dentistry of New Jersey (CMDNJ)-Rutgers Medical School.[1] The Core Content In-Training Assessment Examination was chosen in 1977 and 1978 as the instrument to supply baseline information about the amount of knowledge brought into the program by the resident, as well as to assess cognitive gains over time. This examination, prepared by a national group of family physicians under the sponsorship of the Core Content Review of Family Medicine, was designed to assess a resident's cognitive abilities as demonstrated by factual knowledge. Criteria for passing are established at two levels—one for the entering first year resident, and one indicating a knowledge base expected at residency graduation, which is also used as a comparison for the second and third year resident and for faculty members. Nationally, the examination was taken by 979 residents in 1977 and by 1,702 residents in 1978, a large percentage of whom were in training programs in family medicine.

When we first used this test in 1977, the scores received by the residents and faculty members were both unexpected and disturbing. Most of the second and third year residents, as well as the faculty, did not satisfactorily meet the expectations of the test developers for exit level competence, termed the minimum passing level (MPL). Moreover, their scores were not in agreement with other cognitive assessments of their performance, such as board examinations, peer audit, peer review, and self-review instruments. Needless to say, many residents responded with anger.

Despite the unfavorable results in 1977 in the CMDNJ affiliated programs, the decision to repeat the testing procedure in 1978 was made for two compelling reasons. First, it was essential to be consistent in the use of the measuring instrument. Secondly, it was felt that the development of an in-service examination for family physicians was an important project in which we wished to participate. Unfortunately, the results of the 1978 examination were no more satisfactory than those of 1977. In addition, the residents objected that the examination was not a learning experience. They

From the Department of Family Medicine, College of Medicine and Dentistry of New Jersey-Rutgers Medical School, University Heights, Piscataway, New Jersey. Requests for reprints should be addressed to Dr. Marian R. Stuart, Department of Family Medicine, CMDNJ-Rutgers Medical School, University Heights, Piscataway, NJ 08854.

0094-3509/79/061241-03$00.75

felt that the lack of specific cognitive feedback made it difficult to distinguish a cognitive deficit from a test taking problem (ie, did they not know the material or was the question ambiguous?).

The 1978 manual for the interpretation of the examination results describes the minimum passing level (MPL) and how it was obtained. "Multiple raters reviewed the appropriateness of each response to each item in an attempt to determine whether a 'minimally competent' resident should be able to identify the correctness or faultiness of that response (the Nedelsky technique)." However, the examination was graded simply on the basis of whether a question was right or wrong, and although the MPL takes into account whether an answer is close to being correct or not and penalizes less for the better choice than for the obviously wrong answer, this was not reflected in the actual scores. The validity of comparing a score which is derived on one basis to a criterion (minimum passing level) derived on a different basis must be questioned.

A review of the national averages shows that there is very little increase in raw scores over the three resident years. While first year residents seem to bring more knowledge into their programs than would be expected, second and third year residents nationally fall far short of the MPL. This demonstration of little growth appears to be a serious shortcoming of the examination, raising many questions. Does it reflect the absence of cognitive learning during the residency training or is the test sampling earlier learning? Perhaps the test emphasizes inpatient knowledge while the resident is increasingly emphasizing outpatient knowledge. Perhaps the examination should be given at the end of the third year rather than at the beginning, or perhaps third year residents should take the examination at the beginning of the year to define weak areas and at the end of the year to measure the end product. Will the residents accept this? Our affiliate residents will not. Is the test a valid measurement of the cognitive aspects of family practice?

We were most concerned with the difficulty of comparing a resident's performance on this year's examination to that of last year. Since the MPL changed the second year, it was difficult to determine whether a few points one way or the other on a subtest was in any way indicative of increased or decreased cognitive knowledge.

In spite of these shortcomings, we felt a valid comparison of performances could be made from year to year and that some reliable inferences could be drawn from the test results through a score conversion procedure. To do this, we first compared each individual score to the national averages by the year of residency and converted them to standard scores. Standard scores are obtained by subtracting an individual's score from the group mean, thereby determining how much an individual does either better or worse than the average of all the people in the class taking this particular examination. When this difference is divided by the standard deviation of the group from the mean (indicating the distribution of scores), a Z score is obtained. The Z score in turn indicates how many standard deviations an individual falls either above or below the mean of the group, showing whether he or she has more knowledge or less knowledge compared to others in the class. By multiplying the Z score by 10 and adding 50, a standard score or T score is obtained which shows an individual's performance in relation to the group. (These scores multiplied by 10 again are perhaps more familiar as SAT results.) The range of the T scores generally is from 20 to 80. The mean is 50 and the standard deviation is 10.[2] Any score that falls between 40 and 60 can be seen as average. Anything above 60 is very good and above 70 is superior. Anything below 40 is cause for concern, and below 30 indicates an area that definitely needs remediation. With a small programable electronic calculator, eight hours were required to convert all the scores for the 85 residents and faculty who participated in the testing program at CMDNJ-Rutgers Medical School.

Having converted the results of the Core Content In-Training Assessment to a standard score through this simple mathematical procedure, it is now possible for us to make individual assessments and to determine strengths and weaknesses of individual residents in relation to the national averages as well as to other residents in the same year and the same residency program. We have been able to get clear indications of which of our programs do better in certain subareas, and which areas will require remediation. Results for 1977 and 1978 are easy to compare. However, while it is possible to interpret scores in relation to national group means, individual growth over time

can only be inferred. It is hoped that, in time, a cognitive written examination can be devised that reliably samples the domain of medical knowledge learned in family medicine residency programs, and that valid criteria can be developed to show growth over time. In the meantime, it would be helpful if every test were to give immediate feedback and references to further information sources, thus providing a learning experience as well as a more relevant self-assessment for the residents.

Acknowledgement

The Quality Assurance Program is supported by a grant from the Robert Wood Johnson Foundation.

References

1. Sadler GR, Snope FC: Quality assurance in graduate education in family medicine. J Med Soc NJ 74(1):50 1977
2. Walker HM, Lev J: Elementary Statistical Methods. New York, Holt, Rhinehart and Winston, 1958

Evaluation of Clinical Skills: An Asset-Oriented Approach

Jon K. Sternburg, MD, and Barbara S. Brockway, PhD
Madison, Wisconsin

Evaluation techniques may influence or model a physician's approach to patients. Since medical schools and postgraduate training programs often model punitive or deficit oriented evaluations, it is not surprising when medical students and residents use the same approach with patients. They often ask only "What is wrong?" (What is the pathological condition? What is the deviation from the norm?) This disease oriented approach is not a primary care model, which is prevention oriented, and therefore, needs to identify skills necessary to maintain health.

During the past two years the University of Wisconsin Department of Family Medicine and Practice has experimented with an asset oriented approach to clinical skill evaluation. It focuses the evaluations primarily on positive characteristics including talents, accomplishments, skills, and abilities.

Specifying and praising the student's skills does not rule out demonstrating and correcting inadequacies. Appropriate skills, however, should not be taken for granted because they are "expected." Unless "expected" behaviors are clearly specified and periodically reinforced, they may begin to decrease in frequency.

An asset oriented approach also redirects the resident's attention to patients' behaviors. For example, some patients are called "turkeys" or "crocks": pejorative labels identifying deficit characteristics (traits we do not like and want to reduce). Unfortunately, that labeling may act as a perceptual set for the next visit, and this negative bias is difficult to change once established. By concentrating on deficits, we lose sight of skills or talents. Attending to deficit behaviors (eg, whining, demanding, non-complying, complaining) is frustrating for the physician and dysfunctional for

From the Department of Family Medicine and Practice, University of Wisconsin, Madison, Wisconsin. Requests for reprints should be addressed to Dr. Jon K. Sternburg, Department of Family Medicine and Practice, University of Wisconsin, 777 South Mills Street, Madison, WI 53715.

0094-3509/79/061243-03$00.75