EVALUATION OF PROGRAMS IN GRADUATE MEDICAL EDUCATION*

HAMILTON SOUTHWORTH, M.D.

Professor of Clinical Medicine
Columbia University College of Physicians and Surgeons
New York, N.Y.

COMPETITION between the many programs in graduate medical education in this country is keen and, in general, very healthy. The director of each program may, however, have difficulty in deciding just how good his program is in relation to that of other institutions. He gets a rough idea of its popularity by the number and quality of the applicants for his residencies and fellowships, and by whether and with what ease he may get government support for a training grant. But there should be better ways of directly evaluating the performance of the residents and fellows going through the programs.

As a former but fairly recent member of the American Board of Internal Medicine (ABIM) I believe I can contribute best to this symposium by talking about problems in the evaluation of young internists, for the performance of recent graduates before a specialty board is a valid method of judging the quality of a program.

The object of the various boards has been to identify the physicians with particular competence in each specialty field and to recognize this with a diploma. But the examination process is expensive and time-consuming, and to narrow the field of competitors the boards have been forced to formulate training requirements and thus, to a varying degree, have become involved in setting the standards for training programs. How deeply to become involved in this is a problem being debated at present by the ABIM.

Let us first, however, consider exactly what one can evaluate. It is relatively simple, even in a field as complicated as internal medicine, to examine candidates for knowledge of facts. The ABIM written examinations in recent years have had an internal reliability coefficient

*Presented in a panel, Graduate Medical Education, as part of a *Symposium on the Education of Tomorrow's Physicians* held by the Committee on Medical Education of the New York Academy of Medicine on October 12, 1972.

of 0.93 to 0.95. If knowledge of facts was all that was required the boards could formulate written examinations for certification without bothering about training requirements. But I believe there are two other requirements that are necessary for a good internist: problem-solving ability and a proper approach to the patient. Problem-solving ability is now being tested by the patient-management problem (PMP) or "rub-out" examination in which the examinee works his way through a clinical problem and is penalized for poor management or unnecessary tests and is also rewarded for correct decisions. How effective this technique will prove to be is not yet fully known, but it has already been shown that the ability to handle PMPs increases significantly during an internist's training.

A proper approach to the patient is the third requirement for a board-certified internist. For years an oral examination was used to judge this faculty, for it made possible an appraisal of the candidate's ability to take a history, his skill in doing an accurate physical examination, and his general ability to establish rapport with patients. The oral examination, however, was not reproducible, both because of the differences between examiners and because of the element of chance in the type of case assigned. It is now being increasingly abandoned in board examinations but a satisfactory substitute has not yet been found. Professorial reviews in the home institution of history taking, physical examinations, and case presentations are now under trial by the ABIM, but this in large measure depends on the program director. It may not prove too satisfactory, but it is making the teachers pay attention to how well their candidates have acquired bedside skills.

I have listed all these techniques of examining young graduates because if they represent the best means of evaluation they also indicate the skills that the program directors of graduate education must strive to impart during internship, residency, and fellowship. These techniques also point out the need for an integrated approach and the close personal supervision of all members of the house staff. Cooperation between the specialty boards and the program directors has been suboptimal in the past but at least in internal medicine has improved in recent years.

Another aspect of specialty-board examinations that should help training in various programs across the country is the exact level of performance decided on as acceptable. Determining this by what is

called "setting the cutting score" is a matter of judgment, often without any generally accepted criteria. It was formerly held that, based on a distribution curve of performance, a certain percentage should flunk. But what constitutes that percentage? In recent years the ABIM has tried setting up as a test group the candidates from certain recognized training programs that should be expected to perform both well and consistently. The cutting score may then be set at the level at which the large majority of these candidates would pass. This is a compliment to university and university-affiliated programs and also a challenge to them to maintain consistent leadership.

In 1970 the ABIM went a step further and made it possible for program directors to find out how the percentage of their candidates that had passed compared to that of rival institutions. The results of a compilation for the years 1962 to 1968 were sent to every program director, giving him a clear-cut yardstick by which to evaluate the quality of his program. Figures for 1964 to 1970 are now being prepared. The subject is a delicate one because the performance of each candidate is a confidential matter known only to him and to the board and not even revealed to his program director. But the over-all performance of a group of candidates is a most valuable statistic that I believe will become increasingly available from more and more of the boards.

Some of the specialty boards, such as those for neurosurgery and orthopedic surgery, have gone one step further and set up in-service examinations similar to qualifying written examinations that can be taken on a trial basis by residents one or two years before they come up for their qualifying examinations. The ABIM gave one of these in 1971 to determine what examinable attributes increase during the course of residency training, but the results helped the young physician completing his first year of residency to find out how close he was to a level of competence that would probably mean he could pass his boards. The program director who also received the individual scores was able to help his poorer residents plan further clinical training before they were already committed to practice or to subspecialty training. Most important, he was able for the first time to get a current evaluation of his individual trainees and thus of his whole program. It is hoped that these in-service examinations will be continued in medicine and that more boards will adopt them.

Finally, I should like to say a few words about subspecialization, the great centrifugal force based on the expansion of knowledge that is splitting internal medicine and also other specialties into fragments. The ABIM aims to keep the medical subspecialties under its over-all surveillance, and if possible not to go the way of the separate surgical subspecialties. It has also struggled manfully to find a way of giving special recognition to the man who is able to develop reasonable competence in all internal medicine. Such a trainee has the greatest potential of becoming chairman of a department or a program director. So far it has not proved feasible to set up an advanced examination for all internal medicine without derogating the qualifying examination. The main thrust has been the attempt to prevent early subspecialization by insisting on minimum requirements of broad clinical experience before concentrating on any individual subspecialty. Even now this can be reduced by the special permission of program directors. It is earnestly hoped that these men will use their authority to maintain solid training in the broad field.

In the last few years the specialty boards have evolved from small certifying groups into organizations that in the process of properly evaluating candidates have had to set standards and are having an increasing influence on graduate training at the residency and fellowship level. The future is hard to assess but it is fairly certain to become more complicated. I feel it is of the utmost importance that the program directors of all accredited hospitals work as closely as possible with the specialty boards, make use of the evaluation techniques the boards have developed in judging their own programs, and help the boards set up proper requirements for the development of competent physicians.