

A Conceptual Model for Program Evaluation in Graduate Medical Education

David W. Musick, MA, PhD

Abstract

The author provides (1) a brief overview of the literature concerning program evaluation as applied to medical education, (2) a task-oriented conceptual model for use by residency directors in planning for program evaluation of graduate medical education training programs, (3) an explanation of the term "outcomes evaluation" including distinguishing between types of educational outcomes, and (4) a description of a five-step process of implementing the conceptual model.

Recent accreditation standards for graduate medical education programs require a shift from a process-oriented to an outcomes-oriented model of evaluation. Accordingly, residency program directors must ensure compliance by undertaking comprehensive program evaluation procedures that demonstrate educational outcomes. Such procedures include attention to the *need* and *focus* of the evaluation; the evaluation *methods* to be used; and the documentation and presentation of evaluation *results* to key constituents. Involving teaching faculty

and residents in developing a comprehensive evaluation program is vital to success. Regardless of philosophic debates pertaining to the appropriateness of the outcomes model for medical education, this approach appears likely to predominate in the foreseeable future particularly as related to the six general competencies of the physician. A practical, task-oriented approach will assist program directors in ensuring compliance with program evaluation standards.

Acad Med. 2006; 81:759–765.

Graduate medical education (GME) has entered into a new era, one that has been described as a "paradigm shift."^{1,2} Accreditation standards adopted in recent years by the Accreditation Council for Graduate Medical Education (ACGME) have stressed the importance of program evaluation as part of an overall shift from a process-oriented to an outcomes-oriented system of education. The measurement of educational outcomes as a substantive part of accreditation review has begun to receive greater emphasis, as evidenced by the following statement taken from ACGME accreditation standards: "The program should use resident performance and outcome assessment in its evaluation of the educational effectiveness of the residency program."³

This new emphasis on outcomes has many implications for how GME training programs function, and perhaps for how the process of program accreditation site reviews will function in the future. But what is meant by "educational

effectiveness" and how does one evaluate it? Is program evaluation the same thing as measuring educational effectiveness? And how can one ensure compliance with accreditation standards concerning program evaluation?

As a medical educator, I have served for several years in GME teaching and leadership roles (e.g., associate residency program director, designated institutional official for GME at a major academic health care center, teaching faculty member, and member of departmental and institutional GME policy-making committees). In these roles, I have provided consultation to directors of residency and subspecialty fellowship training programs on the subject of program evaluation. It is apparent that there is not a shared definition of the term *program evaluation*, particularly when it is associated with terminology referring to educational outcomes.⁴

Part of the confusion may be an issue of semantics. The term *evaluation* is often used interchangeably with the term *assessment*. And *evaluation* is used broadly within medical education, and can refer to several distinct processes: overall evaluation of training programs as a whole; of an individual resident's performance; of faculty teaching; or of a given educational lecture, conference,

rotation or other learning experience within a training program. Moreover, the terms *program evaluation*, *curriculum evaluation*, and, more recently, *outcomes evaluation* are also used interchangeably. In view of this lack of standardized evaluation terminology, there is an urgent need to clarify *program evaluation* in practical terms, so that training program directors and other educational leaders in GME have a better understanding of what is expected in this regard.

The purpose of this article is four-fold:

- to briefly review the literature pertaining to program evaluation, both in general terms and in reference to medical education,
- to present a task-oriented conceptual model of program evaluation,
- to discuss *outcomes evaluation* as one type of program evaluation (distinguishing between relevant institutional and program standards), and
- to provide a five-step process that will assist program directors and/or other medical educators in developing effective ways to use program evaluation data to improve GME training programs.

Dr. Musick is associate dean, Medical Education, and associate professor, Physical Medicine and Rehabilitation, Brody School of Medicine at East Carolina University, Greenville, North Carolina.

Correspondence should be addressed to Dr. Musick, 600 Moye Blvd., Brody 2N-72D, Greenville, NC 27834; telephone: (252) 744-2149; fax: (252) 744-3015; e-mail: (musickd@ecu.edu).

Review of the Literature

The term *evaluation* is best defined as a process of decision making about the object being evaluated and how it compares to some standard of acceptability.⁵ Evaluation models have been developed for virtually every type of social endeavor, and are not limited to educational programs. Much of the evaluation literature is found within the realm of the social sciences. Evaluation as a distinct discipline gained impetus in the early 1970s as part of the Great Society movement. Existing evaluation literature provides a variety of models for social planners, heads of governmental agencies, and others who must decide whether a given program is effective in reaching its goals. Frequently, evaluation decisions are also closely tied to requests for funding. An early evaluation model described program evaluation as a process where participants agreed in advance on the purpose and design of evaluation procedures, and on how the results would be used.⁶

Program evaluation in the *educational* setting began to receive more emphasis during the late 1970s and 1980s as a result of increased governmental funding

of reform initiatives at all levels of education. A number of theoretical models emerged, the discussion of which is far too broad for this paper; excellent overviews may be found in books on the subject.^{7,8} Most often, evaluation in an educational setting was conceptualized as a process of making decisions about whether an educational program was meeting its goals and objectives. Thus, a given educational program was felt to be effective if its graduates appeared to have learned the stated objectives and successfully completed all requirements of the program. Within the realm of education in the professions, program evaluation operated in a similar fashion, but was more loosely organized and dependent upon the specific professional disciplinary context in which it took place.⁹

Regarding program evaluation in medical education, the literature pertaining to work done with medical schools and residency programs is less developed than for other fields of education and is largely descriptive.¹⁰ Evaluation models specific to GME programs have been nonexistent; there is simply no overarching theoretical base or consistent approach provided

whereby GME program directors can determine what is expected in this regard. For example, within individual specialty disciplines, many articles exist pertaining to the development and evaluation of courses, rotations or other educational experiences¹; but a unified approach consistently applicable to all GME training programs is lacking. Indeed, it is this lack of a theory-driven, structured approach within GME that may have contributed to the recent adoption of an outcomes evaluation model of program evaluation by accreditation policy makers.

Conceptual Model of Program Evaluation

List 1 provides a task-oriented conceptual model for evaluation within a GME training program. The central notion of the model is to identify the steps involved in planning and carrying out various types of evaluation, consistent with evaluation best practices and accreditation requirements. The evaluation *need* and *focus* represent the initial stages of determining why the evaluation is to be done, what or who is to be evaluated and what “rules” or standards will inform the evaluation. The evaluation *methodology* is the stage where procedures are established for how to collect and analyze evaluation data. Finally, the evaluation *results* stage represents the presentation of data to key stakeholders in an established forum (such as an annual program evaluation meeting). and also the written documentation of all steps taken in performing the evaluation as well as decisions made as a result. More explanation of these stages will be provided below.

While a comprehensive evaluation system would likely incorporate all of the aspects shown in the model, it is recognized that the breadth of a given evaluation procedure is influenced by real-world parameters involving time, resources and available expertise. Ideally, use of the conceptual model would foster prospective evaluation planning and implementation; however, it is also useful for the retrospective fitting of existing data into a rational evaluation framework.

List 1

Task-Oriented Conceptual Model of Program Evaluation in Graduate Medical Education

Step One: Determine evaluation need

WHY is the evaluation being undertaken and for whom? (Accreditation requirement; institutional requirement; specific project; research)

Step Two: Determine evaluation focus

WHAT entity is to be evaluated? (Overall training program, clinical rotation, didactic event, teaching faculty, residents/fellows)

Step Three: Determine evaluation methodology

WHEN is the evaluation procedure to be undertaken? (Planned clinical observation, end of rotation, end of year, after graduation)

WHERE are evaluation data to be collected? (Normal patient care settings, classrooms, other)

HOW are evaluation data to be collected? (Ratings of performance, written/oral examinations, attendance sheets, rotation objectives checklists, surveys, clinical skill examinations)

WHAT types of data analyses will be needed? (Reporting formats, data properties/psychometrics)

Step Four: Present evaluation results

WHO are the key stakeholders who must review the results? (Department chair, teaching faculty, institutional GME personnel, residents)

WHEN should results be presented? (Regular agenda item for faculty meetings; annual program evaluation meeting and/or educational retreat; education committee meetings)

Step Five: Document evaluation results

HOW are evaluation results documented and used for program improvement? (Content delivery issues, frequency with which outcomes are measured, program changes made as a result of evaluation data, resident input into program improvements)

What is Outcomes Evaluation?

Outcomes evaluation refers to a particular type of program evaluation. It is defined by the ACGME as

evidence showing the degree to which program purposes and objectives are or are not being attained, including achievement of appropriate skills and competencies by students.¹¹

The primary distinction between an outcomes-oriented approach and other approaches to evaluation is found in the word *evidence*. In GME, accreditation reviews have traditionally focused on the *process* of education. In other words, external reviewers periodically examined the program's documentation in an attempt to determine whether it was structured appropriately and whether the educational process showed sufficient potential to meet program requirements. This emphasis on process occurred because of the dominance of the Flexnerian model of medical education and because of the difficulty in defining competence in precise terms within a given discipline.¹²

In recent years, especially with a shift to an educational framework based on the "six general competencies" of GME, accreditation has now begun to emphasize not only the educational process but also its *outcomes*, i.e., a *demonstration* of how the program has met disciplinary requirements and produced competent physicians. This is a different approach because it compels programs to present *evidence* demonstrating (or certifying) that resident physicians have learned what they are supposed to learn and that, upon completion of the training program, they are competent to launch an independent practice of medicine.

The new emphasis on outcomes evaluation is illustrated by the following statement of the ACGME:

Assessing the actual accomplishments of a program requires a different set of questions: (1). Do the residents achieve the learning objectives set by the program? (2). What evidence can the program provide that it does so? (3). How does the program demonstrate continuous improvement in its educational processes?¹³

To illustrate this new emphasis on outcomes evaluation, in the following paragraphs, two common educational

components of all GME programs will be considered, one simple and the other more complex.

As a simple illustration of documenting an educational program outcome, consider that all ACGME-accredited training programs are required to provide didactic (i.e., lecture-based) instruction in an organized fashion. Under the previous process-oriented accreditation model, external reviewers conducting site visits would examine a given program's didactic schedule for thoroughness and comprehensiveness. If the didactic schedule appeared to cover all relevant topics shown in the program standards, this would likely suffice for accreditation purposes (although suggestions about organization or content would often be made). Today, however, programs site visited under the current outcomes-oriented model will be reviewed in more detail regarding didactic instruction. Site visitors will likely be interested not only in the fact that an organized didactic program exists; but will also want the program director to provide *documentation* (i.e., evidence) that the lectures were actually attended by residents and/or provided by attending faculty. The lecture schedule itself is viewed as part of the educational process, but the mere existence of an organized didactic program is not sufficient to comply with accreditation requirements. Site visitors must also consider the attendance rate of residents as partial evidence of programmatic outcome. While this example could be considered relatively trite, it is nevertheless a real-time example taken from an ACGME review of a subspecialty fellowship program; the residency review committee letter received afterward explicitly stated that the program must provide "documentation that conferences occur as scheduled, and documentation of conference attendance" (used with permission). Based on this simple illustration, how can a program director partially document the educational outcome (i.e., attendance) associated with resident didactics? Several possible steps could be taken to demonstrate that the didactic program is effective, including

- developing and disseminating a written attendance policy for all trainees, and having each resident sign a statement acknowledging their awareness of the policy,

- setting a minimum attendance rate for residents as a measurable indicator (many programs have a minimum attendance rate of 75% during a given year),
- monitoring attendance by collecting attendance data, with individual attendance data maintained in resident files,
- developing specific consequences for individual trainees that do not meet minimum attendance requirements (e.g., remediation or makeup work),
- requiring residents' ratings of lecture sessions (e.g., content, speakers' presentation skills), which are useful for speaker feedback but also for documenting resident participation, and
- stipulating that attending faculty must each give a minimum number of presentations and/or be present for a minimum number of didactic sessions.

Based on an outcomes-oriented approach to accreditation, a site visitor will review the training program's didactic program focusing on two things: the *process* of education (e.g., What lectures are provided and when? Lecture topics? Comprehensiveness? Organization?); and, the programmatic *outcome* pertaining to the didactic schedule (e.g., Did these lectures actually occur? Who showed up for them? Who presented them?). The point of this simple illustration is this: under an outcomes approach to program evaluation, *both* types of information (i.e., process and outcome) will likely be necessary to satisfy accreditation requirements pertaining to resident didactics.

A second, more complex example will also illustrate the difference between process-oriented and outcomes-oriented accreditation procedures. Again, let's consider the organized didactic series of a given training program and assume that the program carefully plans and monitors all aspects of the lectures as described previously, i.e., the didactic program is well organized, comprehensive, and attended by residents and faculty. Is this fact sufficient for accreditation purposes? More than likely, it is not. After all, residents could be attending didactics but still not learning! An outcomes-oriented approach to residents' education requires us to ask additional questions related to the educational effectiveness of the didactic program. How do we know that

residents' knowledge of given topics increased as a result of the didactic series? Did the didactic content actually make a difference in terms of residents' practice behavior (e.g., adherence to recommended clinical guidelines)? These types of questions get at the heart of the outcomes-based educational approach. Answering these questions satisfactorily results in what the ACMGE labels "evidence of how educational outcomes data is used to improve individual resident and overall program performance."¹³

Common methods used to measure clinical knowledge during medical school are often used less frequently during residency training (e.g., written and/or oral examinations). And, while clinical performance assessment (e.g., standardized patient exams) is designed to assess residents' skills in applying medical knowledge to patient care (including specific procedures and interpersonal communication skills), the available resources within GME programs to undertake these types of formal educational measurements are frequently lacking. The traditional "apprenticeship" model is more often the educational basis of residency training, where residents spend time with various attending faculty members (or more senior residents) and learn by observing and doing. With the advent of an outcomes-oriented approach, the ACGME has considerably raised the expectation level regarding teaching and measuring the competency of individual residents. Put simply, each training program must redesign its curriculum around the six general competencies and must put into place educational assessment procedures that will effectively document that residents' learning has taken place, and that such learning has positively affected patient care. This will require programs to institute measures of residents' knowledge, skills and attitudes in a more formal way than has previously been done. While the apprenticeship model will continue to be valuable from a teaching *process* standpoint, the presumption that residents have gained sufficient clinical competence by spending time with attending faculty over the course of the training program is no longer acceptable *for the purpose of documenting an individual resident's*

competency under an outcomes-oriented model of accreditation.

List 2 provides a list of categories of educational outcome measures that can be used by residency directors to document (i.e., provide evidence of) residents' learning and/or program success. Some of these measures are required by accreditation standards promulgated by the various residency review committees (RRCs); others are based on common educational practice in medical education. Program directors must be very familiar with the expectations for measuring individual resident competency as shown in disciplinary RRC standards. It is incumbent upon all program directors to institute explicitly designed and valid measures of competency (e.g., exams, skills assessment procedures) in a manner similar to those widely used in medical schools, so that questions about what a given resident has learned may be answered. Assistance with this task is also available from the ACGME, via the "outcomes toolbox" section of their Web site.^{11,13} The ACGME has stated that this modified approach to accreditation review, which consists of examining *both* the process and the outcomes of the GME training program, will result in stronger residency training, increased accountability for the "product" of GME training programs (i.e., the competent physician), greater continuity between various levels of the medical education continuum and, ultimately, a stronger medical profession.¹⁴

Outcomes evaluation is highly context-dependent, in that the expectations and needs of various constituents involved in the program being evaluated must be considered. So, for example, outcome evaluation of training programs in Surgery and in Family Medicine (while similar in some respects) would likely involve measuring different outcomes, with heavier emphasis within Surgery on procedural skills.² A key part of any outcomes evaluation system is to determine *what* outcomes are to be measured and *who* is to select those outcomes. In residency programs, measurable outcomes are gradually being added to disciplinary ACGME standards by the respective RRCs. National discipline-specific groups (e.g., specialty certification boards, residency program director associations) may also contribute

to discussions about appropriate educational outcomes for GME training programs.

Institutional or Program Standards?

It must also be remembered that there is a difference between *institutional* and *program* standards related to evaluation outcomes. Educators who are brand new to GME may not realize that the ACGME has an institutional review committee (IRC) that has promulgated separate but related standards for institutions that sponsor GME training. The accreditation of the sponsoring institution will be linked to overall compliance with these IRC standards, and will have a direct effect on the accreditation status of all GME programs operating within the institution. And GME programs must comply with both program-specific and institutional standards. For example, participation by each accredited program in a well-defined procedure of internal review, at the approximate halfway point between official accreditation site visits, is a requirement found in both institutional and program standards.

There are certain institutional standards to which the sponsoring institution must ensure sufficient attention by the teaching hospital and all its GME programs. Examples include institutional resources devoted to GME; resident duty hours and learning environment, program and institutional affiliation agreements, resident supervision, resident employment agreements and benefits, and many others. Recent changes to ACGME accreditation standards also require residency programs (and certain subspecialty fellowship programs) to structure training within the framework of the afore-mentioned six general competencies (which replace the previously required "core curriculum" for GME). Institutions are expected to work closely with individual programs to develop educational objectives and methods of measuring educational outcomes in each of the six competency categories (i.e., medical knowledge, patient care, practice-based learning/improvement, interpersonal and communication skills, professionalism and systems-based practice). This dual accountability for educational outcomes means that a more centralized approach

List 2

Potential Educational Outcome Measures in Graduate Medical Education

- Percentage of didactic presentations or similar learning sessions (e.g., journal clubs, case conferences) attended by each resident
- Pass rates on local, program specific-knowledge examinations (written, oral)
- Pass rates on specialty in-training examinations
- Pass rates on specialty certification board examinations over a given period*
- Satisfactory completion of minimum numbers of identified clinical procedures
- Seeing minimum numbers of patients in certain categories or with certain diagnoses
- Percentage of educational objectives covered during clinical rotations ("objective checklists")
- Percentage of residents involved in institutional and/or community service committees
- Percentage of residents involved in research projects/publications/scholarly presentations
- Documentation and periodic formal review of resident performance data[†]
- Alumni/graduated resident surveys on quality of residency training
- Review of evaluation of specific rotations or other learning sessions by residents
- Annual program evaluation survey of residents by institution or program[‡]
- Ratings of evaluation of faculty teaching by residents
- Frequency and results of education planning sessions/retreats by program faculty

* While nearly all ACGME program standards mention this as a possible outcome measure of program quality, as of this writing only program standards in the following ten disciplines define a minimum acceptable percentage pass rate: anesthesiology, preventive medicine, neurology, orthopedic surgery, physical medicine and rehabilitation, radiation oncology, diagnostic radiology, vascular surgery, general surgery, and critical care surgery.

[†] Normally involves a variety of measures of residents' knowledge, skills, and attitudes compiled into a resident file or portfolio.

[‡] An annual, institution-wide survey of all residents and fellows toward the end of each academic year is highly recommended.

to managing GME will be needed to achieve compliance with all applicable ACGME standards. The IRC standards give authority and responsibility to the institution's graduate medical education committee (GMEC) and to the designated institutional official (DIO) for ensuring compliance with both institutional and disciplinary standards.

Applying the Conceptual Model of Evaluation

To establish a comprehensive program evaluation process, training program directors should work through each component of the conceptual model as follows:

Steps one and two: Determine the evaluation need and focus

Reference to the ACGME requirements makes clear the need for these steps: "Use of outcome data to facilitate continuous improvement of both resident and residency program performance. . . . Programs will be expected to show evidence of how educational outcomes data is used to improve individual resident and overall program performance."¹³ For residency program

directors, the two most obvious needs for program evaluation are compliance with ACGME institutional and program accreditation requirements; and continuous improvement of the training program itself.

In regard to the evaluation *focus*, it is imperative that residency programs pay careful attention to their disciplinary program standards concerning evaluation. As a general rule, there will be paragraphs or sections within each set of standards that are labeled with the headings "Evaluation," "Resident Evaluation," and/or "Program Evaluation." These standards will provide guidance as to what types of evaluation procedures must be undertaken by the training program. Familiarity with both institutional and program-specific standards is vital for program directors. And there is also another source of information for programs: the ACGME's "Common Program Requirements."³ This document summarizes what all training programs have in common across all the various sets of accreditation standards. All accredited training programs are held accountable for the

contents of this set of requirements as well.

In regard to the performance of residents, there is evidence that involving trainees in the process of developing an evaluation system results in greater participation in that system.¹⁵ This approach is consistent with adult learning principles whereby trainees help design the overall educational program and are made fully aware of how their progress through the program will be evaluated. The ultimate responsibility for program outcomes rests with the teaching faculty; their participation in understanding the evaluation process and working with the program director to develop a cogent evaluation system is critically important.

In addition to internal evaluation conducted by a specific training program, faculty and/or residents may also be expected to participate in an evaluation process promulgated by the sponsoring institution. This may include some type of annual overall program evaluation survey or other institutional data-collection exercise.¹⁶ The local GME Committee (GMEC) is charged with the responsibility for this task.

Step three: Determine the evaluation methods to be used

This step is concerned with data collection and analysis procedures. Each training program must invest sufficient resources into a consistent process that measures program outcomes and demonstrates competency at the level of the individual resident. The resulting data will inform judgments and decisions about program improvement. Again, some of the most common categories of educational outcome measures used by GME training programs can be viewed in List 2. Once decisions have been made about which outcomes must be measured within a given training program, a systematic approach to data collection can then be designed and implemented.

A number of possible procedures can be used to determine how a given institution, program, faculty member, resident or educational experience will be evaluated. Some methods are well-established and need not be expounded here (e.g., ratings of residents' clinical performance by various people that interact with them such as attending faculty, patients and other health

professionals, end of rotation surveys, attendance records, objectives checklists). The purpose of this article is not to enumerate all possible evaluation methods; but rather to point out that attention must be given to *how* evaluation data will be systematically collected; *when* it will be collected; *where* (or in what setting) it will be collected; and *how* it will be used to improve the program itself. Program-specific accreditation standards, the "Outcomes Section" of the ACGME Web site^{11,13} and the evaluation literature pertaining to medical education may be consulted for further guidance. Individual institutions may also have educational specialists who can lend expertise to the process of data collection.

Step four: Determine how and when to present evaluation results

Once data have been collected and analyzed, the next step is to decide who should review the data and when. Formal presentation of data on at least an annual basis to faculty, residents, key committees, and other groups is expected³ and will lead to the facilitation of discussion and problem-solving. Data will also likely be reviewed as part of the sponsoring institution's periodic internal review of the training program.

This step is well described in the ACGME Common Program Requirements for all training programs:

The educational effectiveness of a program must be evaluated at least annually in a systematic manner . . . representative program personnel (i.e., at least the program director, representative faculty, and one resident) must be organized to review program goals and objectives, and the effectiveness with which they are achieved.³

While ongoing, informal review of evaluation data should occur, this is *not* sufficient to meet accreditation standards. As required by ACGME standards, at least once per year a *formal meeting* of key decisionmakers within the training program should be held where all evaluation data are presented.³ Some programs call this annual event an "education retreat" or a "program evaluation meeting" and involve the entire faculty along with residents and other key educational personnel. Other programs appoint an education

committee and charge the group with responsibility for this annual meeting.

It is advisable to include others in this annual event who play significant roles in administering the training program and who will be affected by decisions concerning the program (e.g., department chair, program coordinator, chief residents). The focus of the meeting should be on discussing the results of all evaluation data; discerning how well the program is meeting its educational objectives; and deciding what specific steps should be taken to improve the program. The annual educational retreat or meeting will demonstrate that the program has taken seriously its obligation to "show evidence of how educational outcomes data is used to improve individual resident and overall program performance."¹³

Step five: Documentation of the evaluation results

Consistent with the outcomes approach that requires *evidence* of educational activity and accomplishments, it is vitally important to produce written documentation of the program's plan for evaluation, all evaluation data, and procedures used to collect those data. These documents should be continually updated and made available for examination at the time of program accreditation site visits. It is strongly suggested that a formal documentation system be implemented, to include written minutes of all education-related meetings, annual retreats, and/or other similar discussions. Such minutes should include a written agenda, copies of all data sets examined during the annual meeting, detailed descriptions of all decisions or recommendations agreed upon by meeting participants, and a sign-in sheet or other record of attendees. It is also recommended that regular faculty meetings have a designated time for education-related items, with details of those discussions preserved in the minutes of the meetings.

Conclusion

In this article, I have discussed how the outcomes approach to education has influenced program evaluation procedures in GME by reviewing the literature; by offering a conceptual model of evaluation that emphasizes systematic, rigorous attention to evaluation *need*,

focus, methods, and results; by distinguishing between programmatic and institutional outcomes; and by outlining a stepwise process of evaluation design and implementation. It should be obvious to the reader that program evaluation has become increasingly important *and* formal in recent years, as a transition has been made to an outcomes-oriented philosophy of medical education. This formality may easily catch some program directors off guard, particularly those who administer small subspecialty fellowship programs with only a few trainees. Nevertheless, the expectation for a formal approach to program evaluation is now clearly present within ACGME accreditation standards.

With this new theoretical approach has come a need for additional expertise in how evaluation procedures are applied to the unique setting of GME. Program directors today must either gain this expertise themselves, or have access to it. And the DIO at each institution that sponsors GME *must* have training in educational theory, curriculum development and educational evaluation if she or he is to effectively assist program directors in complying with accreditation requirements pertaining to program evaluation.

The outcomes movement in education is not without its critics. For example, some have questioned whether the outcomes approach is compatible with an emphasis on self-directed learning and principles of adult education. That question is beyond the scope of this article, but needs to be discussed further. There is philosophical tension between the need to extensively document educational outcomes via the methods I have described here, and the need to encourage resident physicians to be self-motivated, independent learners. The previous example of documenting lecture attendance is germane here. How can we encourage self-motivated, self-directed adult learning among resident physicians while we take attendance at every lecture as if our learners are still in grade school? GME policy leaders must continue to discuss these issues, so that an appropriate balance can be maintained between program documentation requirements and residents' progressive responsibility for their own education.

It appears likely that the evidence-oriented approach represented by the

educational outcomes movement will remain predominant for years to come. Indeed, the outcomes approach (with particular emphasis on the use of the six general competencies) has also been adopted as a substantial part of continuing medical education¹⁷ and physician certification and recertification¹⁴ processes. It is also prominent within the accreditation process for medical schools, as evidenced by language contained in standards promulgated by the Liaison Committee for Medical Education: "Educational objectives state what students are expected to learn, not what is to be taught . . . student achievement of these objectives must be documented by specific and measurable outcomes."¹⁸

One GME leader has stated that "the ACGME is interested in the competency of the training program and whether the program has demonstrated a pattern of graduating individuals who are competent."¹⁴ This is indeed an important goal and one that deserves support. However, achieving this goal will require two additional strategies. One, institutions that sponsor GME must recognize that the outcomes framework (whereby training programs are expected to formally measure and document individual resident competency) represents a major educational paradigm shift within residency training. Achieving the goal of increased competency of graduating resident physicians will require additional resources devoted to faculty development, curriculum planning and competency measurement. Two, as the ACGME continues to develop its final procedures for determining whether training programs are adequately measuring competency in its graduates, it must respond to criticism by many GME program directors that the expectations in this regard are open to subjective interpretation. As a DIO, I frequently saw instances where programs in certain disciplines were held to higher documentation standards than programs in other disciplines; or where various

programs within a given discipline seemed to receive different accreditation results in spite of similar approaches taken to program evaluation issues. Flexibility among various specialty disciplines in choosing which outcomes to measure (and methods used to measure them) is desirable; but arbitrariness in accreditation decision making based on a lack of consistent understanding of expectations is not. While some variation is to be expected as part of developing a new outcomes-oriented approach to GME, increased training of all site visitors and RRC members regarding educational measurement issues will be necessary to increase the consistency with which final accreditation decisions are made. With these necessary adjustments to the GME landscape, I believe the educational outcomes movement will result in a more highly trained, competent physician workforce.

Acknowledgments

The author gratefully acknowledges Judy Shea, PhD, associate professor, University of Pennsylvania School of Medicine, Philadelphia, PA, for her thoughtful review of the text and contribution to development of the conceptual model of program evaluation.

References

- 1 Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from flexner to competencies. *Acad Med.* 2002;77:361–67.
- 2 Long DM. Competency-based residency training: the next advance in graduate medical education. *Acad Med.* 2000;75:1178–83.
- 3 Accreditation Council on Graduate Medical Education, Common Program Requirements, Section VII. C. 1&2 (<http://www.acgme.org/DutyHours/dutyHoursCommonPR.pdf>). Accessed 27 April 2006.
- 4 Henry RC, Mavis BE. A strategy for developing educational evaluations for learner, course and institutional goals. *J Vet Med Educ.* 2002;29:147–52.
- 5 Green LW, Lewis FM. *Measurement and Evaluation in Health Education and Health Promotion.* Palo Alto, CA: Mayfield Publishing, 1986. Cited in Kern DE.

Curriculum Development for Medical Education: A Six Step Approach. Baltimore: The Johns Hopkins University Press, 1998.

- 6 Weiss CH. *Evaluation Research: Methods of Assessing Program Effectiveness.* Englewood Cliffs, NJ: Prentice-Hall Publishing, 1972.
- 7 Stufflebeam DL, Webster WJ. An analysis of alternative approaches to evaluation. In: Madaus GF, Scriven M, Stufflebeam DL, eds. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation.* 10th ed. Norwell, MA: Kluwer Academic Publishers, 1993.
- 8 Isaac S, Michael WB. *Handbook in Research and Evaluation.* 3rd ed. San Diego: EdITS Publishing, 1995.
- 9 Kassebaum DG. The measurement of outcomes in the assessment of educational program effectiveness. *Acad Med.* 1990;65:293–96.
- 10 Sheets KJ, Anderson WA, Alguire PC. Curriculum development and evaluation in medical education. *J Gen Intern Med.* 1992;7:538–43.
- 11 Accreditation Council on Graduate Medical Education, Outcomes Project, Frequently Asked Questions, #3 (<http://www.acgme.org/outcome/about/faq.asp#a3>). Accessed 13 October 2004.
- 12 Klessig JM, Wolfsthal SD, Levine MA, et al. A pilot survey study to define quality in residency education. *Acad Med.* 2000;75:71–73.
- 13 Accreditation Council on Graduate Medical Education, Outcomes Project, Slide Presentation (http://www.acgme.org/outcome/project/OutIntro_fnl1.htm). Accessed 27 April 2006.
- 14 Leach DC. The ACGME competencies: substance or form? *J Am Coll Surg.* 2001;192:396–98.
- 15 Lane DS, Ross V. Consensus on core competencies for preventive medicine residents. *Am J Prev Med.* 1994;10:52–55.
- 16 Heard JK, O'Sullivan P, Smith CE, Harper RA, Schexnayder SM. An institutional system to monitor and improve the quality of residency education. *Acad Med.* 2004;79:858–64.
- 17 Final Report from the ACCME Task Force on Competency and the Continuum (<http://www.acme.org/>). Accessed 27 April 2006. Accreditation Council for Continuing Medical Education, April 4, 2004.
- 18 Functions and Structure of a Medical School (LCME Accreditation Standards). Liaison Committee for Medical Education, Standard ED-1 (annotation) (<https://www.lcme.org/functionslist.htm#educational%20objectives>). Accessed 27 April 2006.