

An Evaluation System for Residency Training

Eric C. Jelly, MD, and Charles Friedman, PhD
Columbus, Ohio, and Chapel Hill, North Carolina

The authors distinguish between systematic and episodic methods of educational evaluation and review the numerous pressures necessitating incorporation of evaluation systems into house-staff training programs. Such pressures include needs to document quality and quantity of training experiences, often mandated by external agencies, as well as trainee desire for increased feedback to guide their personal development. In response, a house-staff evaluation model based on principles of participative decision making and collegial information sharing was developed and implemented at the University of North Carolina. The model features use of multiple and interacting sources of information about each resident and structured occasion for that information to be discussed in the resident's presence. This process leads to carefully documented decisions about strengths and weaknesses of each resident and about the program itself. Evaluation systems are beneficial to the degree that their design matches the organizational climate of the program they serve, and to the degree that they can evolve as the program evolves to maintain such goodness-of-fit.

Successful educational programs are growing, flexible entities responding to the needs of all participants. Programs respond in two ways, either *episodically* through individual initiative of students and faculty when need for change is perceived, or *systematically* through an organizationally mandated sequence of evaluation activities. Changes of the latter type will be fueled by data collected from many sources and may involve the process and/or the content of the program. This paper endorses the advantages of systematic over episodic evaluation and describes one such

system as applied to evaluation of a family medicine residency and the residents in it.

A system of evaluation seeks to integrate all parts of the process of data collection and interpretation into a unified whole. Such integration leads to greater economy of effort, more opportunity for all concerned to participate, and increased validity. Should the system also be simple, easily understood, and jargon-free, it is more likely to be accepted by all participants in a given program.

The model evaluation system presented here applies specifically to the evaluation of resident house staff, although a similar framework could also apply to evaluation of faculty or program effectiveness. The model emphasizes appropriate use of information collected and not specific techniques or instruments for collecting information. This approach reflects an assumption that

From the Department of Family Medicine and the Office of Medical Studies, University of North Carolina, Chapel Hill, North Carolina. Requests for reprints should be addressed to Dr. Eric C. Jelly, Family Practice Residency Program, Ohio State University, 456 Clinic Drive, Columbus, OH 43210.

data gathering methods such as in-training examinations, rating scales, and simulations have received the bulk of scholarly attention to date, leaving unexplored the equally important ethical, political, and educational issues surrounding use of the information.

New Pressures for Evaluation

Traditionally, evaluation has served a number of important functions at all levels of education in most fields. For example, a fundamental and basic requirement of most, if not all, medical education programs is to assess the degree of the learner's achievement and determine the advisability of "passing" the learner to the next training level or to graduate him/her from the program. This certification process often includes an externally imposed requirement to "grade" or rank the student in relation to his past and present peers.

Another traditional function of evaluation entails the use of information relating to resident competency to stimulate adjustments to the educational program as consistent deficiencies across many learners are identified and attributed to program elements. Such alterations may be applied both while a program is underway, for the benefit of present students, or at its completion, to benefit future learners. In this way programs can improve over time.

These functions usually have not been sufficiently demanding to require that evaluation in graduate medical education be conducted systematically. Episodic evaluation based on informally gathered data, often in combination with national standardized tests, has sufficed to meet the need to promote and certify residents and to guide whatever program adjustments might be considered necessary by its administrators. However, new pressures for improved evaluation have recently been generated. The consequent need for programs to respond appropriately almost demands that evaluation be conducted systematically using models such as the one to be presented.

The new stimuli include some demands externally imposed on residency programs. In family medicine, for example, program graduates increasingly need to document their training experience in terms of procedures performed in order to apply for and obtain hospital privileges, especially those requiring access to coronary care and obstetric

services.¹ So increasingly, it is not sufficient to certify the student as "ready to advance" or "not ready to advance"; it is necessary, in addition, to provide the more detailed information which undergirds such decisions.

Another externally imposed requirement relates to the perceived need by federal and private granting agencies to demonstrate the impact of funds allocated. This, in turn, leads to stipulations that all grants to support medical training include substantial evaluation components; some specifically recommend the active participation of an educational specialist in implementing such evaluation.² A related social trend has led to demands for increased accountability from the professions themselves.^{3,4} Evaluation of competence of both trainees and practitioners emerges as a mechanism for responding to this need.

Lastly, but no less important, is the increased need perceived by the trainees themselves for feedback to guide decisions about future training and practice. Residents share with other health care professionals engaged in advanced training the traits of adult learners capable of guiding their own course of study in self-motivated directions.⁵ Systematically collected evaluation data can blend smoothly with residents' own perceptions of personal competence to inform such self-guidance. This characteristic of residents, if reinforced by training programs, may generate "lifelong learning" behavior considered so essential to the professional well-being of future practitioners.

Organizing an Evaluation System

To obtain the greatest acceptability, an evaluation system is preferably designed by those who will contribute. Thus, the first and most important step in constructing a system is to provide an ongoing planning function in which all can directly or indirectly participate. This function is most appropriately carried out by a formally constituted planning body. This body should include representatives of residents, faculty, and administrators. If required or desired, educational specialists can be involved as well. Members of the planning body should be chosen in a manner congenial to the values of the organization, either elected or appointed as appropriate. It is extremely important for the program administrator to provide this planning body with a specific charge, especially to

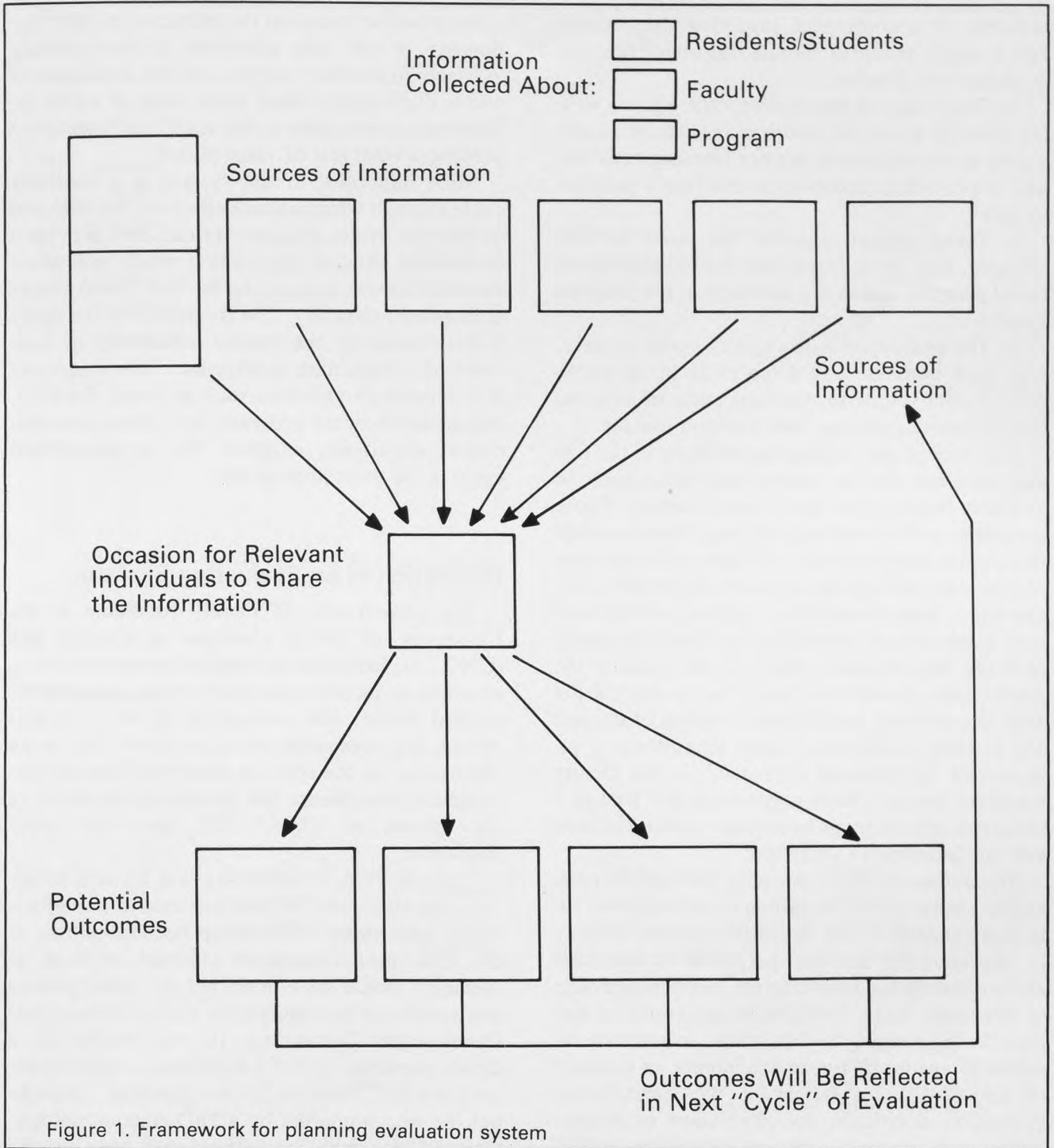


Figure 1. Framework for planning evaluation system

clarify the extent to which its duties are decision making or advisory.

A framework which can help a planning body begin evaluation system design is given in Figure 1.

This framework contains five major features considered essential to any evaluation system.

1. It suggests that information from multiple sources (such as self-assessment tests, clinical rating scales, and psychological instruments) be available for each resident.

2. It provides a structured occasion for *periodic sharing of the information* collected about each

resident, in a conference involving the resident and a small group of faculty members best acquainted with him/her.

3. The *results* of the conference are put in *written* form to guide the resident in making adjustments or corrections in his/her learning activities and to provide a record of the resident's progress to date.

4. These *written reports* are used to *fuel changes*, both in each resident's individual educational program and in the elements of the program itself.

5. The evaluation is an *ongoing cyclic process*, with each outcome of a given cycle being an important data source for the next cycle of information collection, sharing, and decision making.

The core of the evaluation structure is the formal occasion for the sharing and discussion of available information about each resident. Those present at such a meeting will vary from program to program but preferably will include the resident plus at least one representative of the faculty. The meeting is best scheduled at regular intervals and well in advance to allow time for those concerned to study the available data. At the meeting the participants should first feel free to discuss the data; the resident should have a feeling of support and security sufficient to allow acceptance or rejection of the opinions expressed by the faculty members present. Such negotiation will lead to a congenial agreement as to actions required to remedy the deficiencies identified.

The outcomes of the periodic meeting are preferably clearly stated in writing including plans for implementation of any decisions reached. It may be necessary for one staff person to be identified as responsible for following up recommendations or decisions made. Information generated at and prior to the meeting is then stored in a secure location to ensure its future availability as a source of information for letters of recommendation, promotion decisions, documentation of experiences, and, of course, future evaluation conferences.

The multiple sources of evaluation information that undergird the periodic meeting are added to the system through action of the planning body. Prior to addition to the system, each source needs to be considered carefully with regard to the incremental value of the information generated and its validity. Factors considered will be the eco-

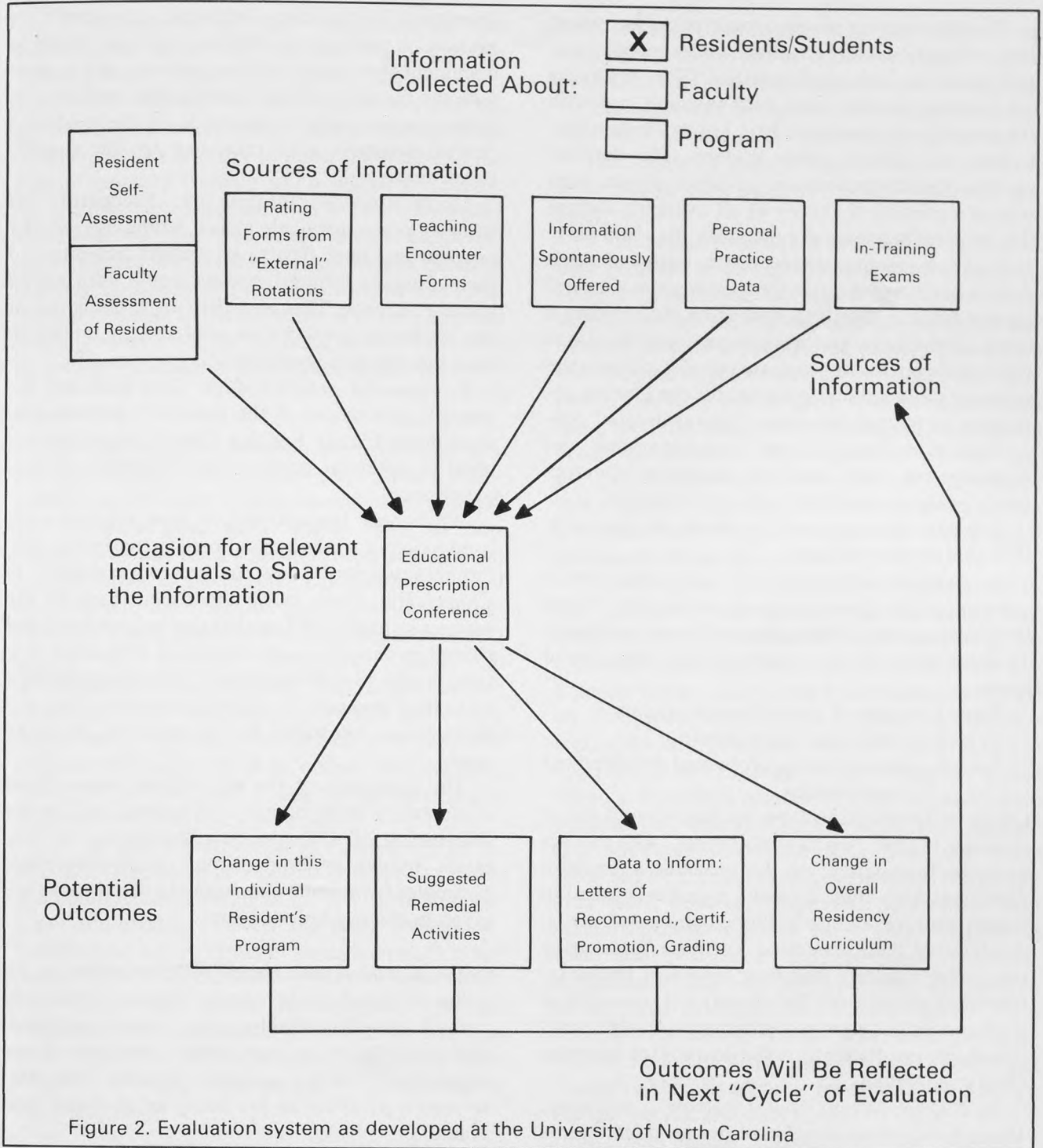
nomy of effort to obtain the information, the relationship of the data generated to that already available from other sources, and the avoidance of undue duplication. Small scale trials of a new information source prior to full scale implementation provide a valid test of value added.

Most important to the system is a feedback cycle through which all outcomes will be reflected in the next cycle. The process can then provide a continuous flow of information which will allow the educational process to be fine tuned rather than altered radically. The frequency of the cycle is determined by the regular scheduling of each resident's evaluation conference. This frequency will depend on variables such as level, duration, and structure of the program. In a three-year residency, experience suggests that a semi-annual cycle is the most appropriate.

Illustration of an Evaluation System

The Department of Family Medicine at the University of North Carolina at Chapel Hill (UNC-CH) instituted an evaluation system which attempts to incorporate many of the features described above. The particulars of the UNC-CH system are presented not as an ideal but as an illustration, as the specific structural features and organizational climate that generated the details of the system at Chapel Hill may not apply elsewhere.

Early in 1976, a committee was formed to initiate and then steer all resident evaluation efforts in the department. This group became known as the Evaluation Committee although most of its members would have preferred the more general and seemingly less pejorative title of Educational Development Committee. It was chaired by a family physician (E.C.J.) with interest and experience in educational theory and practice. Throughout, he was supported by a PhD educational specialist (C.F.) who was also a part-time faculty member in the department. In addition to the committee chairman, the clinical faculty was represented on the committee by a physician interested in and responsible for research in the department, thus providing liaison between the two activities. Residents nominated a representative from each of the three years, and service on the committee developed over time into a popular



duty. The committee always had at least two resident representatives; and at one time the competition for places resulted in four residents attending the monthly meetings. Representatives from associated programs were regular attendees at the committee meetings and occasionally guests were

invited to address specific issues. For example, a hospital administrator was asked to attend when documentation of hospital experience was discussed.

The system developed at UNC-CH is shown in Figure 2.

The information sharing conference, known as the "educational prescription conference,"⁶ was scheduled for each resident to last 1½ hours every six months. At this time, each resident met with the two faculty members best known to him/her, usually the clinical team leaders. The detailed agenda of each conference varied; however, each was preceded by a survey of all available data in the resident's personal evaluation file. The early part of the meeting involved discussion of these data in review of the previous six months. General agreement was then reached as to the resident's areas of strengths and weaknesses, and decisions were made on any appropriate modifications to the resident's educational program for the coming six months. A written document, the resident's "educational prescription," was then generated, and arrangements were made to implement the recommendations specified in the prescription.

The data sources used to inform the prescription conference included:

A. *Resident self-assessment instrument and a companion faculty-assessment instrument.* These locally produced forms contained items relating to 33 competency areas organized into three broad types.

Type 1—classical clinical competencies

Type 2—behavioral competencies

Type 3—practice and professional development competencies

On the self-evaluation form, each resident is asked to assess his/her own "comfort level" with respect to each competency; on the separate companion document each clinical faculty member is asked to assess each resident's ability in the same areas. Analysis of data generated by these forms over two years suggests that they represent useful information sources for the educational prescription conferences but were not sufficiently reliable, and, therefore, insufficiently valid to rank or formally certify individuals.⁷

B. *Rating forms from "outside" rotations.* These forms were developed by the departments hosting the rotations and did not necessarily convey the type of information optimal to the family medicine department. Nonetheless, owing to the difficulty that would result in trying to alter the forms used in the external experiences, this information was accepted as the best available.

C. *Teaching encounter forms.* Each resident had a "chart review" with a faculty member

scheduled immediately following a patient care session. A method was devised and trial tested to document the content of the sessions and to provide the faculty members conducting each review with an opportunity to comment on the resident's clinical progress with reference to the specific cases discussed.

D. *Spontaneous information.* Frequently off campus preceptors, physician attendings on the hospital inpatient floors, as well as patients and their families offered spontaneous information about residents. This was filed in anticipation of the upcoming prescription conferences, at which time the resident would be made aware of it.

E. *Personal practice data.* This included the quantity and nature of the resident's patient contacts in the Family Practice Center; such data can identify potential gaps in the resident's clinical experience.

While only the above five data sources were used as input to the educational prescription conferences when the system was implemented in Chapel Hill, there is no theoretical limit to the amount or nature of data that can be generated and shared as part of a comprehensive evaluation system. Profound practical limits, in terms of the time and effort required to collect evaluation data and the amount digestible by the resident, naturally apply.

The outcomes of the educational prescription conferences might be minor in nature, such as the fine tuning of a resident's experiences, or they might involve major decisions or modifications. Examples from some educational prescriptions are given in the examples below.

Dr. A. (third year resident) In an effort to develop efficiency in the Family Practice Center and in the hopes of maintaining a high standard of medical care, Dr. A. is sometimes intolerant of the shortcomings of the paramedical staff. This is interpreted by them as his being an abrasive individual.

"Dr. A. and I have had several discussions concerning the minor problems listed above and he has agreed to try a different approach in dealing with the paramedical staff, and in two to three months' time I will obtain some feedback from them concerning the effectiveness of his efforts."

Dr. B. (first year resident) Dr. B. stated his personal discomfort with inadequate time to do in-

depth reading about the conditions he has encountered during his first year residency. He feels he learns better by in-depth study in a quiet area which allows him to concentrate. He expressed a need to have ample understanding of the disease process in order to feel comfortable in the role of primary treating physician. The specific uncertainty he expressed was a fear of missing something that might then come back to haunt him in the future. We exchanged ideas about different learning situations and began to discuss ways that he might be able to get more out of his experiential learning, which will be the main avenue by which he will be exposed to information in the future.

The data available for and reports generated by the prescription conferences were stored under conditions of strict confidentiality and retained after resident graduation. These materials thus remain available for discussions on promotion and overall rating as necessary. Upon completion of the program, information related to clinical experiences is useful for letters of recommendation and access to hospital privileges. Since all prescriptions were available for review by the chairman of the evaluation committee and by the department chairman, regularities of patterns of resident deficiency could be noted. Such recognition could, in turn, lead to discussion and changes in the overall residency program.

Use of the prescriptions and outcomes to inform the next cycle of evaluation serves to maintain continual monitoring of the residents and program, and produces strong and viable evaluation. The cyclical nature of the evaluation system is apparent to the resident and works to increase confidence in its fairness and thoroughness. Through examination of the prescription documents, it is evident that the faculty become, over time, more adept at providing useful evaluative information to the residents.

Suggestions for Implementation

Implementation of the system just described was a gradual process which took place over two years. Some significant methodological and organizational problems were overcome as the Evaluation Committee—and in particular the two faculty members carrying out most of the implementa-

tion—responded to advice, help, and support from individuals representing different facets of the program. The systematic features previously described appeared to significantly influence acceptance of and trust in the information generated.

The following specific factors seem to be conducive to positive outcomes of an evaluation system.

1. The basic foundation of an effective evaluation program is the establishment of a planning body, or committee, through which all are represented, and all wishing to do so can be heard and even participate to the desired degree. This forum can become a place where enthusiasts are tempered, doubters encouraged, and all participants learn the intricacies of developing an educational program. Explicit representation of the departmental administrative structure will ensure that decisions made by this committee are not inconsistent with other developmental efforts under concurrent implementation.

2. Patience is the quality most fundamental to the successful expansion of the evaluation program. Increasing momentum can come with increasing confidence, but a fine touch is required to time the introduction of each new data source. As an illustration, early in the evolution of the system, all faculty, residents, and staff were invited to complete a Myers-Briggs Type Indicator, a personality inventory commonly used in health care professional settings. This brought considerable and open resistance during a time still characterized by distrust of the evaluation effort. One year later, however, this same instrument was administered to many of these same individuals with less resistance and, in some quarters, enthusiasm. Each new instrument usually needs to be seen and modified in a pilot trial before being generally introduced. In this way, each gradually becomes reliable, valid, and accepted. Faculty and trainees alike are adept at recognizing and rejecting the "quick and dirty."

3. Once firmly established, the "system" should not become a barrier to further change in evaluation or educational methods. It serves the learners and faculty, not vice versa. Existence of a forum for explicit consideration of evaluation issues can lead to effective investigation and concerns outside the committee's original purview. For example, the department—with planning assistance from the Evaluation Committee—

sponsored evening dinner sessions to facilitate evaluation of the orientation program for new residents. A similar session was held to define the characteristics of the ideal family physician as a step toward producing both objectives and methodology for ranking residency candidates.

4. The issue of access to evaluation information and confidentiality is best *explicitly* considered and clarified. These considerations were frequently addressed by the committee, leading to a specified policy approved generally by the faculty and residents. The data were kept secure in the evaluation files by the secretary to the chairman of the committee. Access to such files was strictly limited to the resident him/herself and the clinical faculty of the department, a group defined by the committee for each academic year to conform to personnel turnover. Each resident was asked to sign a letter granting the department permission to retain this material after the resident left the program and to use it for specific purposes, including recommendation letters, documentation of experience, and research using aggregate data only.

5. To ensure success, overevaluation and underevaluation are best avoided. A common critique of the system addressed the numerous and rather lengthy forms. Conversely, a complete evaluation was considered preferable over one that involved less effort but was necessarily inconclusive.

6. Since the evaluation program potentially affects all aspects of the residency program and departmental organization, it should be sensitive and matched to the goals, values, and norms prevalent at any given time. Mismatch to this organizational "climate" can result in ineffective, intrusive, and often unacceptable evaluation procedures.

Comment

Although it shares many features with other approaches to house staff evaluation described in the literature,^{8,9} the system herein described differs in its emphasis on an overall process guided by a few common-sense principles. The system comprises, above all, a framework for planning and evaluation compatible with virtually all training programs. It is not tied to a particular model of education and/or evaluation, and can be implemented without reference to a lexicon of social scientific terminology.

The system is based on an overriding premise which views evaluation more as an organizational than a psychological issue. From this premise it is deduced that the manner in which evaluation is conducted—how the information is collected and used—dominates the data collection mechanisms themselves in determining success of an evaluation effort. Furthermore, the wealth of research on validity and reliability of particular evaluation approaches ensures the existence of numerous useful data sources to incorporate into an evaluation system; this plethora also argues for more specific investigative attention in medical education to evaluation's organizational overtones and consequences.

The system described here reflects a distinctly democratic ethos congenial to the personal values of the authors and considered appropriate to an advanced level of professional education. An alternative ethos would likely generate alternative system designs.

References

1. Hansen DV, Sundwall DN, Kane RL: Hospital privileges for family physicians. *J Fam Pract* 5:805, 1977
2. Program Brochure: Grants for Faculty Development in Family Medicine. Department of Health, Education, and Welfare, Public Health Service, Bureau of Health Manpower (Hyattsville, Md). Brochure issued November 1977 for grant award in fiscal 1978
3. Gurel L: The human side of evaluating human services programs: Problems and prospects. In Guttentag M, Struening EL (eds): *Handbook of Evaluation Research*, vol 2. Beverly Hills, Calif, Sage Publications, 1975, pp 11, 27
4. Competence in the Medical Professions—A Strategy. In Bureau of Health Manpower (Hyattsville, Md): DHEW publication No. (HRA) 77-35. Government Printing Office, 1977, p 2
5. Stein LS: The individual curriculum in nursing leadership. *J Cont Educ Nurs* 2(6):7, 1971
6. Shahady EJ, Hoff DL, Aluise JJ: Performance evaluation of family practice residents using an educational prescription. Presented at the Annual Workshop for Directors of Family Practice Residencies, American Academy of Family Physicians, Kansas City, Mo, June 7-9, 1976
7. Friedman CP: Self-perception vs faculty perception of family medicine resident competency. Presented at the Annual Meeting of the North American Primary Care Research Group, Toronto, April 12-15, 1978
8. Corley JB: *Evaluating Residency Training*. Charleston, SC, Medical University Press, 1976
9. Emery JLR: *An Evaluation System for the Family Medicine Residency Training Program*, mimeographed booklet. Galveston, Tex, University of Texas Medical Branch at Galveston, 1976