

## STRATEGY AND TACTICS OF EVALUATING A LARGE SCALE MEDICAL CARE PROGRAM

Howard R. Kelman, Columbia University School of Public Health and Administrative Medicine  
and  
Jack Elinson, Columbia University School of Public Health and Administrative Medicine

The emergence within recent years of a variety of large scale medical care programs which seek to extend traditional patterns of medical care services in new ways, or to drastically modify them, has created urgent challenges for those concerned with evaluation of their efficacy. The magnitude of some of these programs, the complexity of the services offered, and the peculiarities of their origin and development call for modification or new evaluation research strategies.

The material which follows reports some early experiences of an evaluation unit whose activities have been designed to provide continuous appraisal of the impact upon an urban ghetto community of a program through which the resources of a community general hospital center have been merged through affiliation with those of a University Medical Center. Evaluation in this effort has been conceived of as a built-in monitor of the program accomplishments and shortcomings, if you will, with the dual objectives of providing both long term as well as more immediate feedback of programmatic consequences. This evaluation program differs from the more usual types of evaluation studies, at least in its conception, in that appraisals are to be continuous and not one shot.

This report will illustrate some of the methodological issues and problems encountered thus far in developing a program of evaluation activities tailored to suit the specifics of the situation confronting us. First the setting and background will be sketched briefly.

The medical care program we seek to evaluate evolved against a background of growing professional and community concern for the declining levels of care that had become increasingly evident in the municipally operated hospital system of New York City. In an effort to halt this retrogression in care and to upgrade its quality, a program to affiliate the municipally operated institutions with medical schools and voluntary teaching hospitals was undertaken. Under this program, the operation and supervision of the medical and related professional services of the municipal institutions were to be assumed by the medical school or the voluntary teaching hospitals. These agreements, in addition to providing funds to obtain additional staff, equipment and other resources, provided also for necessary programmatic expansion and for related teaching and staff training activities. However, the program we will be reporting on is the only one of some twenty odd such mergers that provides an evaluation component in a general hospital setting as an integral part of its operation.

The history of the development of this vast affiliation effort has not proceeded without controversy as to its effectiveness and costs. Political interests as well as professional concerns, since the inception of the program, have served to keep public and professional attention sensitive to the various activities that have marked the program's development.

The hospital of our concern is a 700 bed physically outmoded institution - soon to be replaced by a modern structure - located in the heart of a Black ghetto community. It services an area containing some 400,000 persons and is the major health services center in the community.

In addition to the kinds of social currents and unrest that prevail in many such communities today, the University itself has for a variety of reasons become a convenient target for political groups in the community who are suspicious of its intentions. From time to time these groups have attempted to arouse community opposition around a number of issues including the health of the community and the operation of "their" hospital by the University Medical Center. It is against the background of these past and continuing social and political forces that the affiliation of the University Medical Center with the community hospital center of our immediate concern has taken place.

The affiliation process itself was accomplished over a period of several years on a department by department basis. More recently it has culminated in an overall institutional contractual agreement, except for one department. The evaluation component of the affiliation is now organized as a separate group within the affiliated community hospital and is linked to the University Medical Center by virtue of its being a function of the University's School of Public Health and Administrative Medicine, which in addition to the evaluation activities, carries responsibility for the operation of certain other services conducted at the community hospital site.

A wide variety of questions have been put forth as being salient to the evaluation unit's attempt to assess the impact of the affiliation program. The questions or issues cluster into two major areas of concern:

- 1) questions addressed to the community - its health levels, patterns of care, needs for care, and how these needs are met by the multiple providers of care, including the hospital, in the community; and
- 2) questions centering around the care given by the hospital - its appropriateness and adequacy and its consequences for the

recipients of care.

The first set of questions is being approached through a community based survey of representative households whose health status and related behavior, and changes in status and behavior, are being studied over time. Studies of the character and consequences of care in and around the hospital and its patients address themselves to the second set of questions. The two sets of studies, and questions, are viewed as complementary. In one - the Community is the "target," in the other it is the hospital patient which is the "target." Together, they ultimately will provide information concerning the hospital's programmatic impact.

Another way of defining the relationship between the two study efforts is by viewing the community-based study as providing information about the "denominator," while the hospital based studies will tell us something about certain "numerators."

In developing our evaluation activities we have had to take the following factors into account:

- 1) Our entry into the hospital affiliation program was at some point after its inception. Thus, we do not have the possibility of a before-after evaluation strategy. Our approach is probably more akin to a "during-during-during" strategy.
- 2) The identification of the beginning, the midpoint or even the end of the care program is impossible to define in a general hospital which was already an ongoing operation prior to the affiliation of the institution with its counterpart University Medical Center.
- 3) The Hospital Care program is not a readily definable package of the same order as a drug, a public housing project or even a rehabilitation effort. What we are confronted with is a dynamic, complex, extremely busy hospital in which affiliations have developed over a span of several years with counterpart departments at the medical center on a service by service basis. Thus some services, or sub-programs, in the overall hospital program may be as much as three to six years in the making, others perhaps one year old and still others are on the drawing board or in various states of planning and development.

An advisory committee, which meets monthly, serves to keep us abreast of program developments and provides also an opportunity for the hospital directors of services to spell out their departmental goals, problems and accomplishments.

- 4) We have not found it possible to capture or to reconstruct more than fragmentary base line information - relevant to

evaluation of programmatic accomplishment - and have had to plan ourselves for the creation and gathering of basic descriptive information relative to the community, the sociodemographic character of the population, health resources and so on. These sorts of information can then be used as base line measures to assess change in denominator and numerator populations in utilization patterns, health levels, and related questions concerning health care in the community.

- 5) The creation of a reliable body of information at the hospital site useful for evaluation purposes has also had to be undertaken by us. Our need for reliable information relating to the treatment of patients that would also be amenable to electronic data processing merged also with a need on the part of the hospital for systematic utilization review. Currently in its early stages of tryout - after nearly a year or more of development - is a systematic record abstract which contains a large number of items of information concerning the ingredients of treatment. It is meant to be completed on all patients discharged from the hospital. It is anticipated that such information on the nearly twenty thousand discharges per annum will be a valuable source of information for purposes of medical and record audits, or more broadly put professional audits of care and for utilization review. It may also serve other hospital needs for statistical and other reportage of effort. Conceivably follow-up and other kinds of special studies for purposes of evaluation can in the future be designed and executed. It is also anticipated that this activity may have the effect of "beefing up" other record keeping systems in the hospital.
- 6) Another set of issues we have had to confront has been the differing definitions of evaluation by hospital personnel on a variety of levels, and flowing from this, their differing demands or expectations of evaluation activities.

One view of the purposes or functions of an evaluation unit sees its role as being akin to that of an inspector general in a military establishment. Evaluation activities defined thus would concern themselves mainly with checking up on the manner and mode of the detailed delivery of services to patients. The evaluation unit would find ways of uncovering the shortcomings of care, bring them to the attention of the providers and check on their change or correction.

Another conception of evaluation would define its functions as akin to a statistical and research service bureau

of the hospital. The function of the evaluation unit in this view would encompass such activities as the gathering and compilation of statistical reports on hospital activities and on the community and its population. Also the unit would provide technical consultation in the design and conduct of research, record keeping and data processing and assume responsibility for providing computer support to ease the problems of administrative bookkeeping.

Still another conception of evaluation defines it as a kind of follow-up service operation. This view held largely by those with direct patient care responsibilities looks to the evaluation unit to provide them with information on the fate of individual patients. A variant of this viewpoint also looks to the unit as a resource for resolving administrative and programmatic problems and questions.

In these several viewpoints there are obvious elements of compatibility with our own role definitions and activities as they have developed. There are also obvious pitfalls inherent in each view which we have thus far successfully avoided. To play the role of the inspector general would, we believe, quickly embroil us in entanglements with staff responsible for care. The role of the overseer we feel is not compatible with objective evaluation of care and its consequences. To become a sort of statistical or research service bureau while undoubtedly helpful to the overall operation of the hospital would, we believe, deter us from the task of appraisal of care and relegate us to the status of a bookkeeper or accountant. Finally, we do not define ourselves as being competent to provide clinical and program administrative services even assuming that the potential rivalry of an additional service unit could be avoided.

We have contended with these tugs and pulls through constant repetitive interpretation of our role and through providing limited consultation on research, record keeping, data processing and related topics. Members of our staff also sit as members of ongoing hospital administrative, review, and research committees. We are coming to be viewed as a member of the hospital family - a new and different one perhaps - but still not as outsiders who may threaten exposure.

As was indicated earlier, the overall design of our evaluation activities provides for studies of in-hospital care and its consequences for patients.

The particular approach we have been

developing evolved against the background of both general considerations concerning evaluation of hospital care as well as specific features of concern, centering around our particular hospital program.

Approaches to the evaluation of hospital or medical care are varied, and include:

- 1) follow-up or "outcome" studies of selected groups of patients,
- 2) studies or audits of the quality of medical and surgical care,
- 3) studies of the structure and functioning of the treatment environment or system
- 4) experimental program assessments and/or clinical trials, and
- 5) statistical compilations of hospital or health activities.

It is clear, in considering the multi-functional properties of an acute general hospital, that each of these approaches (and others not cited) has or could have a contribution to make in providing information as to how a hospital functions, and in gauging the effects or impact upon patients of changes in hospital treatment program. It is also clear that each approach, or strategy, though it would be contributory would also be incomplete when measured against the complex task of evaluating a total hospital care program. As yet there is no single, all inclusive, research model or methodology that can be readily applied to the hospital situation we face or to any community general hospital for that matter.

For example, counts of hospital or health activities (e.g., number of admissions, mastoidectomies, visits to clinics, board certified medical doctors provide information concerning program effort or the achievement of program activities. But, they say not very much about patient benefit - effects or outcome. The same is true in part of the "medical audit" approach. Clinical trails tend to be either diagnostic or treatment-specific and require highly rigorous field control conditions.

As we have sought to develop and tailor our own method we have leaned heavily on two approaches which have been successfully employed in prior medical care research. The first is the more or less traditional medical audit whereby judgements of quality or effectiveness of care on selected diagnostic categories of patients are made by some expert or peer group based on information contained in the hospital record. Such studies have employed either standardized pre-selected criteria or have asked for overall, less rigorously defined ratings of care from expert groups.

In our own work we have broadened the concept of medical audit and are attempting to secure ratings which go beyond the technology of medical

or surgical treatment to include nursing as well as social work services.\* We have been somewhat arbitrary in deciding to go this far and no further because conceivably all facets of the hospital and how it functions can and do influence patient care. Though we may alter our view at some later point, these seem to us at this time to be the most consistently crucial elements.

We also are endeavoring to increase the information base on which these judgements of care are made by having the members of the assessment team not only review the record, but also see the patient and clinical staff responsible for the patient's care. These reviews of care will be made while the patient is in the hospital and available for study.

In addition to these judgements of care the team is also rating "outcome." Outcome is defined in terms of health and social functioning along a variety of dimensions including clinical course, life threat, self care, ambulation, productive activity and discomfiture to self and to others. The team makes ratings of patients' levels of function on these dimensions at or close to the point of discharge and it reconstructs them for status at admission. In this way we hope to identify change in patient status occurring during the hospital stay.

An additional set of judgements or ratings asked of the team involves judgements of expected outcome, again employing the same criteria. That is to say the team is asked not only to judge patient status at different time points, but also whether the observed change in status is different or similar to that which this patient ought to have achieved were optimum care available and applied. Where a discrepancy occurs between achieved and expected status, the team is asked to identify, if possible, the reasons for the discrepancy. These may include specific factors relating to the content or delivery of care in the hospital, untoward and unforeseen events affecting the patient's clinical course, or factors not intrinsic to care or the system of care.

We have observed thus far two kinds of discrepancies. In the first instance, expected outcome exceeds achieved outcome. Here one would look to see whether and to what extent upgrading in kind or quality of care could or would be beneficial. We have also found instances where achieved outcome exceeds that of expected outcome. Here one could say that the patient's level of functioning was not "disabling" enough for his or her well being.

---

\*These ratings are being made by an assessment team consisting of a physician, nurse and medical social worker.

Where there are no discrepancies between expected and achieved outcome, we would argue that upgrading in kind or quality of care is not necessarily indicated in order to improve function, or if it is provided, benefits should not be sought for in terms of the criteria or outcome measures that we are using. (Upgrading in care may be entirely justified for other reasons of course.)

We are coupling this appraisal of hospital care and its consequences for patients with studies of the patient at three months and also at twelve months post-discharge. Again, employing the same outcome measures, patients' functional status will be reassessed. Additional information concerning after care utilization, social and economic circumstances and patients' views of care will also be collected. For the development of this phase of the work we have leaned heavily on the methodology of post-discharge follow-up studies.

Our decision to use the functional status of patients as the yard stick to "measure the results of care is based upon the following considerations:

- 1) Particularly in a municipal hospital situation, a wide variety of forces - of which the technical competencies or performance of physicians and other professional staff is but one - may affect patient care and influence outcome.
- 2) Outcome is also likely to vary and to be affected by the patient's past bio-social history and experiences - his potential to respond to treatment - whatever the character of treatment it self is.
- 3) The patient's post discharge status is also likely to be influenced by still another group of factors - social situational, post discharge services, etc., all of which may influence his clinical and functional state.

To select out of these different matrices of factors one or two treatment variables only for study (e.g.: Medical Doctor's performance, length of stay) means we also need to be able to hypothesize about their presumed consequences for outcome of care. This, to say the least, can be extremely hazardous. It is possibly less hazardous and maybe more relevant to our overall evaluation goals to define and describe changes in the character and status of the patient population over time.

It should also be noted that we do not propose to measure the specific ingredients of care or services (which have been delivered or received by the patient) per se. Nor do we propose to measure the quality of care per se. We do ask that the assessment team address themselves to these questions in so far as it is possible for them to define and analyze their consequences for, or effects upon, the patient's clinical and functional status.

The primary focus then of the judgements of the assessment team is the patient - his status and changes in status during the course of hospitalization and subsequent to it. It is in relation to evident, expected, and predicted, status changes that other judgements concerning patient care practices (clinical administrative, social) and their influence on outcome (patient status) are to be made. Through these means, if our efforts are successful, and coupled with study of the patient following discharge, we hope to trace and link up ultimate changes in patient status to patient care factors in and out of its hospital. It is our expectation that the information produced through these efforts will enable us to address ourselves to the following questions believed salient to an evaluation of the hospital care program.

- 1) In socio-medical terms, what is the character of the population now receiving hospital care and what is its potential for change?
- 2) To what extent, and for what kinds of patients, are care needs being met or not, and with what consequences for the patient and others around him?
- 3) To what extent, if at all, could or would upgrading in kind or amount of hospital care be expected to achieve different "results"?

It is possibly worth while noting here our belief that the validity of the procedures we are developing rests upon the following set of assumptions. The first assumption is that care currently being offered in certain areas is something less than optimum. This is so, either because of an insufficiency in kind or quality or overwhelming demand, or because of a poor potential for favorable response on the part of patients, or some combination of these factors. Secondly, it is the hospital's job to favorably alter the health status of patients where it is possible to do this. In addition we assume that the hospital's job extends beyond responsibility for the care of the patient while he is inside the institution and includes provision for after care where indicated. Finally we believe the dimensions that we have chosen to measure or to determine patient status are salient.

The decision to focus initially on the in-patient hospital stay is not purely an arbitrary one. While ideally in studying medical treatment it would be best to begin at or near the origin of an illness episode, to do this at this point in the development of our overall evaluation program is not feasible in terms of establishing an "at risk" but presumably "well" population. The desirability of having a population which can be defined clinically and diagnostically ruled out, again on the basis of feasibility, using the hospital admission as the entry point of study. (A useful approach if one is studying a numerically restricted group of patients with very well defined and established medical conditions.)

Using a discharged population permits a greater latitude of choice and selection of representative patient populations and, also relevant to objective evaluation, obviates the possible bias of observed treatment. Another advantage of a "soon to be" discharged population resides in the background of past experience of studies of quality of care which have focused similarly on the hospital stay of discharged patients.

Just as the origin of the illness episode may not coincide with hospital admission, the termination of an illness episode may not conclude with hospital discharge. In these assessments the discharged population will be studied three months after leaving the hospital and again at 1 year post-discharge. This is planned for several reasons. First, it is desirable and important to relate the outcome of care to estimates of the character of care received in terms of immediate consequences (soon after discharge) and then again after some time lapse (residual effects). Secondly, it is of importance to get the patients' side of things. (events around the need for hospital care) and their views of care and services received as well as other situational factors affecting post-hospital status (after-care services, etc.). Third, a prospective view of patients' function and behavior will permit study of how the illness episode is "concluded" (if in fact it is), how and if other arms of the hospital are employed (Out-Patient Department, Home Care, Emergency Room); how other health and social agencies are called upon by providers of care or by the patient for continuing care; and of equal importance it permits observation of "new" illness episodes which may emerge during the 12 month post-discharge observation period.

It is to be noted that these evaluation efforts go beyond available evaluation studies in two important respects:

- 1) They extend the observation period well beyond the hospital stay.
- 2) They seek to "measure" both care and outcome of care and propose to relate them to one another as well as to other factors.

To conclude we might summarize by saying that our efforts thus far might properly be classified as being in the pre-evaluative, baseline data gathering, and methodological developmental stage. In addition to the complexities of the program and the difficulty of defining the relevant evaluative questions we also find ourselves working within a system which is not geared to the demands and requirements of evaluation research. The system is itself being retooled for more efficient, up-to-date and more humane patient care. In this process the role of an evaluation group is subject to varying definitions and strains. To a large extent the pace of its development proceeds at a rate determined largely by the ability of the medical care system to respond to the total demands upon it for patient care in which evaluation is but one relatively

minor component.

Our experience to date suggests that built-in, long-term, evaluation of the work of complex large-scale health programs and organizations presents in its aggregate, if not in its specifics, some very special theoretical as well as practical problems, in contrast to more limited, less complex "fewer faceted" service programs. Some of the reasons for this are:

- 1) The "target" is a community or some other geographically defined population, itself embedded into a larger population matrix.
- 2) The program objectives cannot be neatly or parsimoniously defined because perhaps, like "sin" the program has always been with us, or if new, some of the goals are global and diffuse and the essential worthwhileness is self-evident.
- 3) The application of an experimental design is either inappropriate or impossible.
- 4) The research models we have at hand - crude as these may be when measured against laboratory or other exquisite standards - were derived from the pilot,

contrived demonstration or experimental trial "era." While no doubt the need for, and utility of, such endeavors will continue, they do not seem to fit the "new" demands for evaluation that we, for example have encountered. In short, as we have "discovered," there is no single inclusive, well developed methodology that can be readily identified, and applied.

- 5) As a result of these factors and others, such evaluation programs are likely to be long term and "high risk" in terms of their immediate pay-off, and eventual yield. This has obvious implications, career and otherwise, for evaluation research personnel as well as for the sponsoring or "host" agency. It also has administrative implications with respect to the structural relationships between the evaluation and service programs that are required to assure the achievement of both service and evaluation goals. All of these problems assume different proportions in long term as opposed to single shot evaluation projects.