American Handbook of Psychiatry

New Methods for Assessing the Effectiveness of Psychiatric Intervention

Ernest M. Gruenberg

New Methods for Assessing the Effectiveness of Psychiatric Intervention

Ernest M. Gruenberg

e-Book 2015 International Psychotherapy Institute

From *American Handbook of Psychiatry: Volume 6* edited by Silvano Arieti

Table of Contents

NEW METHODS FOR ASSESSING THE EFFECTIVENESS OF PSYCHIATRIC INTERVENTION

The term “psychiatric intervention” is used in this chapter to refer to anything a psychiatrist does in response to the fact that any person is a patient for whom he has some degree or type of clinical responsibility. Treatment is ordinarily used to refer to actions taken by the clinician, or at his instructions, which are intended to relieve, ameliorate, or terminate a disordered state he believes the patient to have, or to affect favorably some distress associated with the disorder; we generally do not use the term “treatment” except when referring to certain classes of action—physical measures, such as drugs or shock or surgery, psychotherapeutic measures in a formal pattern. Other actions, such as hospital admission or release, ward assignment, counseling with relatives, charging fees (or not), making appointments, and so forth are generally thought of as being outside the concept of treatment and are regarded as having something to do with what is called“ management of cases” or of one’s practice or of the hospital or some part of it. It is not necessary to try to persuade the modern psychiatrist that this separation is not always sharp. In one sense, everyone will agree with the dictum that treatment begins the moment the psychiatrist meets the patient and asks about what brought him. The process of meeting and asking and listening to the answer begins a relationship that has some significance for the patient’s mental state and future. The term “intervention” is simply

introduced here to avoid the narrower meanings of treatment and to include a psychiatrist's decision not to take a particular action for one of his patients. The role of psychiatrist is the center of attention, but this does not mean that other professionals cannot intervene with both positive and negative effects on the patient's life, but only that this chapter is not attempting to deal with those events.

The effects of every medical intervention are assessed to some extent on each occasion. The patient when conscious of the intervention is making some assessment, and the observant physician is always keeping a watchful eye out for the effects of his treatment. Effectiveness, however, cannot be observed in such a neutral way. Effectiveness has to do with certain specific effects that it was hoped the treatment would produce. Giving an epileptic patient barbiturate tablets may have a large number of effects—gratitude for the attention, fear of the implied dependency, drowsiness—but the object of giving the barbiturate was to reduce the frequency and intensity of seizures and if this effect is not observed, the treatment is regarded as ineffective.

It will be helpful to face the fact at this point that the process of deciding what effect the treatment was hoped to have has nothing to do with the world of facts, but is directly related to the world of values. In the example above, seizures are assumed to be something undesirable, and the treatment intervention is introduced in an effort to abolish them (or some of them). This

intervention is not the act of a neutral observer, but of a person who is partisan—he is intervening to change the patient's functioning in a particular way. It is not possible to use facts to justify this attitude. Scientific methods can be used in an effort to understand how people come to adopt the attitude that seizures are bad things to be avoided if possible, and understanding the cultural, historical, and social forces involved in reaching this conclusion might or might not modify the attitude. But to the psychiatrist with clinical responsibility for the patient, this is not a debatable issue. As will be seen later, the person who seeks to assess the effectiveness of a treatment must believe he knows what effect is to be sought—his assessment will then depend upon whether that particular effect has occurred or not.

Hence, the first point to keep in mind is that since the dawn of modern medicine as a profession, doctors have been assessing the effectiveness of their medical interventions. The Hippocratic writers were forever assessing whether particular procedures would alter the course of events in the desired direction. Underlying the concept of medical intervention there is always a twofold assumption. One part assumes a knowledge of what would have happened if no action had been taken. The other part assumes that one or more of the things that will happen, if no one intervenes, is to be avoided, if possible. The fever, pain, diarrhea, convulsions, or whatever, will continue in the absence of the treatment; they are not desirable; therefore specific interventions are justified to avoid their continuing. Without going into all the

implications in this chapter, the reader should be reminded at the outset that a psychiatric intervention in the course of another's life is a form of medical intervention, surrounded by an ancient tradition and mystique and entailing complex and solemn obligations; the so-called Hippocratic oath is one of the most ancient formulations of the nature of an obligation one person bears to another or another group of people—older than any oath of allegiance, older than any currently used assertion of religious faith, older than any current marriage vow. Another ancient dictum of the nature of the physician's responsibility to his patient should also be recalled as we get into the modern methods for assessing the effectiveness of intervention: *primum non nocere*—the doctor's first duty is to do no harm.

Evaluation Research

Scientific methods have been developed in efforts to understand the nature of the world, to find some underlying principles that put the blooming, buzzing confusion of the world of facts into some sort of order capable of human understanding. The scientific methods are ways of systematically confronting an idea about what the world is like in a way that tests the idea's correctness. These scientific methods can also be applied to planned human actions. Most of us have applied them, more or less systematically, to the action of starting an automobile quickly enough on a cold morning to get the internal combustion motor going before the battery runs down. We

experiment with the number of times we pump the accelerator pedal before turning on the starter. We experiment with different positions of the accelerator. Those who have manual chokes develop hypotheses about the best manipulation of the choke. Some try racing the motor before turning off the ignition when stopping the car at night. We "learn" from experience (more or less systematically) that a car well-tuned starts more quickly than a car poorly tuned, that a thin oil in the crankcase impedes the starter less than thick oil, and so forth. We develop rules for each car at each season and these rules are the result of more or less systematic experiments regarding what seems to "work" better for that particular car.

This simple problem of a human action, which can be carried out in several different ways, can be used as our elementary paradigm to illustrate several important principles that distinguish evaluation research from research that applies a scientific method to the problem of understanding the nature of the world. In the first place, the goal of the action is clearly defined: get the internal combustion engine started with the least drain on the battery. In the second place, the means for achieving this goal are obvious: turn it over with the electric starter and give it a mixture of gas and air in the best proportions to get it exploding. In the third place, the situation limits the choices for action. Manipulation of accelerator, choke, and starter switch are the only options in the situation. The sequence of the actions and their timing is all that can be varied in any given "trial." Such things as tuning the motor

and having thin oil in the crankcase or the way of stopping the motor the previous night all had been done before the trial and form part of the context. It is remarkable that with these limited options and the crucial nature of the outcome in the lives of so many people, no one seems to even keep notes on how the trials go and to make systematic analyses; even trained scientists don't keep records on how these actions are carried out each morning and then draw inferences from analyzing a series of trials. This is a fourth important principle: When we think we can keep the relevant experiences clearly in mind and can "learn from experience," we do not design and record systematic experimental trials, which is not to say that we are not using scientific methods but only that the experiments and their results seem so obvious that we do not think of making records in order to gain the maximum information from the various trials. Fifth, there is very little in the way of "theory" in these experiments. The automotive engineer who has a refined sense of the process of carburation and ignition and the artist who has no idea what these terms refer to will carry out his uncomfortable cold morning experiments in pretty much the same way and will probably "learn" equally quickly what the most suitable pattern of action is. This fifth principle emerges from the fact that the goal is clearly defined and the alternative actions are clearly delimited so that the trial-and-error sequences proceed by rules of thumb rather than by a systematic deduction from the laws of physics. We are studying a man-made contraption designed to achieve the

goal by fiddling with gadgets. The experimenter is operating a device made by someone else for the experimenters and he is only trying to use the devices in the most efficient way. When none of his experiments are successful, he takes the device back to the shop to get it fixed so that it will work the way it is supposed to.

Each of these principles applies to evaluation research. First, the goal is defined and taken for granted; this means the statements of goals are statements about assumptions and are not part of the hypothesis to be tested by gathering data. The hypothesis deals with the choice of means to achieve goals, not with the value of the goals. Second, the variety of means available for achieving the goals are obvious and specifiable and the problem is to use the alternatives in the most effective way. Third, we ordinarily don't systematize our experiments in such situations. Fourth, the experiments are being conducted on man-made devices intended to achieve specific purposes and designed to be operated by using specific controlling devices.

The Clinical Trial

The "clinical trial" in medicine illustrates this set of principles. A group at the Mayo Clinic compared the efficacy of a group of analgesics in controlling pain among terminal cancer patients. The fact that aspirin was as good as or better than many of the synthetic compounds thought to be more

effective, and which are certainly more expensive and more dangerous, was surprising enough to make newspaper headlines. Here the third principle had obviously been working: no one had thought it necessary to keep careful records of their clinical experiments on this issue; they had "learned" from experience what systematic record keeping and study design revealed to be untrue! The goal was clearly the relief from subjective sensations of pain, the means, a limited number of analgesics at various doses—clearly man-made gadgetry with highly defined systems of controls and administration. The main benefit of systematic evaluation research is its ability to undo that perennial human state: "It's not ignorance that causes most trouble but what we know that's not so." Such clinical trials are the best established types of evaluation research in medicine. The principles of the definitive clinical trial were worked out less than four decades ago, in the middle 1930s, by teams associated with the Medical Research Council of the United Kingdom. Sulfonamide was the first new drug to be given systematic clinical trials at the first stages of its availability. Because systematic clinical trials were used right from the start with this drug, its effectiveness became established within a few months. This careful recording of the use of sulfonamide rapidly established the value of the systematic clinical trial at the same time that it established the value of the sulfonamides as a more powerful, safer, set of antimicrobial drugs than almost anyone had thought possible. In these trials the conditions treated by the sulfonamides tended to be rapidly fatal and the

sulfonamides were both much more effective and acted much more rapidly than the antisera previously available, so there was little room for misinterpretation of the results. Later, when the method was applied to nonfatal conditions with marginally better treatments than already available, the technology of the clinical trial became more sophisticated.

The systematic selection of patients thought suitable for the trial, the arbitrary assignment of these patients to treatment types, the independent objective recording of the course of events, both favorable and unfavorable, produced a systematic experiment in which scientific management and clinical management were related to each other in such a way as to produce the maximum information with the minimum risk to the patients.

Ethical objections to this procedure were present from the start; they were not regarded as the obstructionist arguments of clinicians whose favorite remedies might come in second best, but to a genuine concern for getting the best possible treatment to each patient as soon as possible. Clinicians were mistrustful of statistical interpretations as compared with their judgment based on intimate knowledge of each case. There is no doubt that the responsible clinician knows many things about his cases that the statistics of clinical trial can never reflect. Of course, whenever a treatment is given to any patient an experiment has begun—will this treatment given this way to this patient affect favorably the outcome of his particular condition?

The argument is not between those who wish to experiment in the treatment of patients and those who do not, but between those who want to conduct these experiments in a systematic way and those clinicians who prefer a less systematic way. Of clinical trials there will never be an end, but each one must be justified separately in the light of all the known facts at the time it is proposed. It is not ethical to start a clinical trial that requires withholding an established treatment for a dangerous condition in order to test the effectiveness of any new compound that also might be thought to be effective. Safety must be known to a certain degree before any clinical trial can be started. To withhold a drug requires grounds for substantial doubt regarding its effectiveness; to introduce a drug requires substantial reason to suspect its effectiveness and to think it sufficiently safe. In general there are two occasions most favorable in the history of a treatment: first, when it is beginning to gain acceptance among experts but is still suspect; and second, long after it has become well established and its value assumed but its value is coming under question by experts who begin to doubt whether the conventional wisdom was right. During the time that a treatment is well established and “everyone” knows it is useful, it cannot properly be withheld; but before it becomes well established some experts have begun to report good effects, while others remain doubtful; that is a good time for a systematic clinical trial. Before some experts have begun to give reason for thinking that the new treatment might be valuable, a clinical trial cannot be

justified because there is no reason to invest time and to subject patients to the treatment.

The clinical trial sets up a systematic experiment: it selects patients thought to be suitable for a particular treatment or set of treatments, assigns them arbitrarily to one treatment or another and then observes the outcome. It seeks to gain the maximum knowledge from the minimum number of research subjects through the neatness of its design and the care with which the trial is carried out. The technology of designing these trials was built on a generation of agricultural research into seed selection, fertilization, protective sprays, and so forth. It is from that type of research (systematic evaluation of planting, breeding, and cultivating actions) that we derive most of our statistical methods and principles of design. In evaluating any particular piece of evaluation research one should always keep in mind the methods these statisticians developed for evaluating the significance of a particular pattern of results. They asked themselves: If, in general, these two treatments had no different effect, and I were to take a sample and do what I did by applying treatment A to one part of the population and treatment B to the other part, what are the odds that in any one sample like this, those getting treatment A would do twice as well as those getting treatment B? And they give their answer by saying that they can figure out how many samples of the size used would show that much difference if in general there were no difference. They express this by giving a probability that the observed difference could have

occurred by chance: p = 0.05, which means that they figure 1/20 of such studies would show the As doing twice as well as the Bs even if in general there was no difference between the effect of A and the effect of B. This measurement only specifies the sampling risk of drawing the wrong conclusion. Obviously, the more evaluation studies you read and attend to the more likely you are to draw the wrong conclusion from one of them. Thus, if you read many evaluation studies, and each gives p = 0.05 there is a chance that one in twenty of them gives the wrong conclusion. You have no way of knowing which. So there is a danger in taking large-scale actions, if the only justification is a single study.

There are much more serious hazards in assuming that each evaluation study correctly interprets its findings. In fact, because evaluation studies use scientific methods to approach action problems, the conclusions of the authors must be scrutinized in terms of their methods just as carefully as one would any other scientific publication. Indeed, one should probably be a little more skeptical regarding the average evaluation study than about the average physics or chemistry paper.

A physics or chemistry study published in a scientific journal will automatically conform to certain basic standards of research methods and reporting. The application of scientific methods to planned action research is not so well established or conventionalized, however, and is likely to be

published in a journal that has a clinical or administrative readership unaccustomed to dissecting each scientific report.

The Preventive Trial

The preventive trial is another form of evaluation research. The action taken is supposed to prevent something from happening—it will lower the frequency with which a population becomes ill. The most dramatic and effective recent preventive trial accompanied the introduction of the Salk vaccine to prevent poliomyelitis. It was organized on a nationwide basis and used volunteers. The volunteers (and their parents) knew that half would get the new Salk vaccine and the other half were getting an inactive salt solution. They also knew that no one in the inoculating teams knew which was which; each vial had a number and the name of the child was entered on a list opposite the number of the vial. Many thousands participated in this study. The investigators had to wait for the next polio epidemic to find out their results. They were spectacular and only then were families notified as to which children had received the Salk vaccine and which the ineffective solution, so that those in the latter group could go out and get the Salk vaccine. This is the kind of preventive trial where a state of immunity is to be produced in each individual treated. There are more general preventive trials in which the water supply of a city is cleaned up and the typhoid and cholera death rates lowered. No individual is “treated,” but the whole population’s

relationship to its own feces is modified through engineering. This action too produced dramatic results.

The most important new developments in the methodology of evaluating psychiatric intervention effectiveness are in response to new ideas regarding the goals of treatment—what are regarded as the desired changes in the course of events.

The invention of the planned clinical trial moved medicine into a new era of planned innovations in treatment technology and the rapid, purposeful assessment of the value of each new treatment. Although that technology developed in the mid-193os, psychiatry only began to absorb its lessons and apply it to new psychiatric interventions two decades later following the discovery of the phenothiazines.

The Massachusetts Mental Hospitals Experiment

An Attempt to Rehabilitate Chronic Mentally Ill Patients

The modern clinical trial only became common in psychiatry after the introduction of the phenothiazines and reserpine in the mid-1950s. It was also the introduction of these drugs that brought intensive work on defining goals of treatment in terms of specified amounts of improvement regarding specific symptoms or disabilities rather than in terms of terminating

disorders. As will become apparent when we review four examples of modern investigations, there has been a move toward more exact specification of goals and also a move toward experimentation with the context in which a treatment is given, manipulation of the context itself becoming an "innovation" to be evaluated.

An excellent example of these transitions is a study designed to test the interaction of these new drugs with intensive milieu therapy and psychotherapy conducted by Milton Greenblatt, George Brooks, and a team of associates, beginning around 1960, using patients and staff in three Massachusetts mental hospitals. Schizophrenic patients who had been in the hospital on the current admission between five and ten years were located at Boston State Hospital and Metropolitan State Hospital. The 115 selected patients were divided into four groups by random selection in such a way that the addition of drugs to the treatment regime and their transfer to the active treatment program of the Massachusetts Mental Health Center could be assessed as separate and as combined interventions. Two groups stayed on the chronic wards of the large mental hospitals where they had been located, one received a tranquilizing drug regimen and the other didn't. Two groups were transferred to the Mental Health Center, one receiving the drug treatment and the other not.

Effectiveness can be assessed in terms of discharge within nine months

of beginning the regimen. Of the fifteen who were discharged out of the 115, half continued treatment at the day hospital. Twelve of the sixty-eight patients receiving tranquilizers (some in custodial and some at the mental health center hospital) were discharged and only three of the forty-seven who did not receive tranquilizers were discharged. The authors call this a "trend" but one that "does not reach significance." Without going into more sophisticated statistical techniques, the reader can form his own judgment by taking a pencil and paper and asking himself how many of the sixty-eight patients receiving tranquilizers would have been discharged if the tranquilizers, in fact, had no influence on the likelihood of the patient's being discharged. Simply divide 15 by 115 (13.0 percent) and multiply sixty-eight by this overall discharge rate. The result of 0.130 x 68 = 8.8 tells how many of those patients we would expect to have been discharged if the tranquilizer medication made no difference. This calculation is essentially the first step in all kinds of statistical significance tests; they all start by stating the null hypothesis in the form of a question: What would we expect if the two treatments made no difference? Then the data are examined to see whether the observed numbers differ enough from what one expects on the basis of the null hypothesis to justify us believing that the null hypothesis was refuted by the data. By simply calculating the expected numbers in this way, the reader will obtain a sense of how much weight he would be willing to place on the "trend." In this instance the excess number of discharges in the

tranquilizer group is just 3.2, i.e. 12—8.8. Common sense tells us that this is not a big enough difference to get very excited about.

If one looks at the effect of transferring the patient from the custodial to the mental-health center environment for six months, one finds that eleven of the sixty transferred patients were discharged while only four of the fifty-five remaining in the custodial hospital were discharged. By using the overall discharge rate again (13.0 percent), we can say that of the sixty transferred cases 7.8 would have been discharged if transfer made no difference, again too small a difference between observed and expected to draw any important conclusion. Hence by this criterion, discharge, no definite effectiveness could be attributed to either of the two treatment intervention patterns.

The elements of the mental status were appraised by psychiatrists attached to the research team at the beginning of the experiment and again six months later; those who were "much improved" six months later can be regarded as examples of success, according to the authors. Though the diligent reader of such reports can feel distressed at the fact that the published report does not give sufficient information for one to attempt to reproduce this criterion—"much improved"—this discomfort should not lead one to discard the data for that reason. One must ask oneself whether, taking all the published evidence together, it seems reasonable to assume that a reproducible criterion was applied. It is more important to estimate whether

one thinks that the research psychiatrists were capable of making unbiased ratings of "much improved" in the context of the study; the important bias is of course with respect to which study group the patient had been in. Presumably the first mental status was made before study group assignment, so that fact could not affect the examiner. The six-month examination, however, was made after the patient had been in the study group six months. The examinations were made at the locus of treatment so the examiner could not help but know where the patient had been treated—the "double blind" procedure was not possible. No effort was made to keep the examiner blind with respect to the drug regimen either. Therefore, if the examiner were inclined to be biased, the reported data could reflect such a bias.

The report is inadequate in the information it gives regarding methods used to reduce these biases and is faulty in failing to give us the investigators' own estimate as to whether bias was present and if so, of what kind. They are justified in stating that to make these appraisals "blind" would have involved great expense, and they may be right that at that point such a great expense would not have been justified.

It is worthwhile diverging for a moment here to ask, how could blind appraisals of whether the patients had improved or not been obtained? Inasmuch as the data being used are essentially clinical observations, not laboratory impersonal data, the problem is one of putting good clinical

observers into contact with the patient's manifestations of disordered functioning at two points in time; for this purpose there is no substitute for the interview. It may seem at first sight to be absolutely impossible to think of a way to gather this kind of information with no possibility of bias, but there are conceivable plans. For example, the research interview could have been videotaped on the two separate occasions. These videotapes could have been played back one after the other to clinical observers asked to rate the degree of improvement in each element of the mental status. These observers need not be told which group the patient was in. In fact, they need not be told that the patients were part of an evaluation research program. The data gathering for this purpose could be combined with another study—one which focused on the development of reliable ratings regarding videotaped mental status interviews. This second objective could be the one explained to the clinicians who rated the videotaped interviews. This procedure might appear to be too "tricky" and too deceptive to the clinicians providing the ratings. However, if the trouble were taken to generate a genuine serious investigation into methods of obtaining reliable ratings that would produce their own independent findings and these study subjects were included in the material used to conduct that study, the failure to inform the observers of the additional use of the ratings they provide could be defended. It is obvious that though the clinicians rating the interviews from the video tapes could be kept "blind," even of the existence of the study being conducted, the interviewer on

the scene could well be biased and the way in which the taped interview was conducted could be affected by his knowledge of the study and the subject's group assignment. This also could be avoided in principle if the context of the interview could be similarly modified so that the interviewer's focus is made irrelevant to the evaluation research. For example, outside interviewers could be employed and they could be provided with the second study's frame of reference—a good interview for mental status appraisal. Other situations occur in which trained clinicians are asked to conduct an especially careful and comprehensive interview. For example, clinicians taking their specialty examinations, or clinicians conducting examinations that will be used for the purpose of teaching other clinicians how to conduct an interview. To describe these devices briefly is also to indicate how complicated an undertaking is needed sometimes to get rid of bias in the data that is going to be used.

If one is willing to take on faith that the data gathered in the Massachusetts Hospitals Experiment did not emerge from some systematic bias in data gathering, one can then look at the findings as indicating the effects of the two treatments. The improvements were not equally distributed over the different elements of the mental status. Those who received drug treatment improved in their social behavior, more so among those who also received the shift to milieu treatment. Appearance, activity, and speech were the main areas of improvement. On the other hand, mood, ideation (content), and grasp improved not at all. This is an important clue as to what is most

readily improved in the chronic mental hospital patient, either through use of changed milieu or through the application of tranquilizer drugs, or better yet the two at the same time. Behavior improves but subjective symptoms and thought processes remain relatively untouched. Around 1960 a large number of evaluation studies ensued that examined the rapidly changing way in which mental health programs were approaching the problems of the seriously ill mental patient. There was intense activity in clinical settings in these years, which preceded the introduction of the National Mental Health Center Act in 1963. Investigators who had never before engaged in evaluation research as such became involved in the challenge of turning their research skills toward evaluations of the newer innovations being started in so many places. The result was a number of new methods for assessing psychiatric interventions that arose from ever increased efforts to find a way of distinguishing forward motions in psychiatric care from backward motions or useless motions. Four of these studies are selected for more or less close examination in this chapter because they are good illustrations of the varieties of new developments. The first was an effort to rehabilitate chronic custodial care schizophrenic patients. The second was an effort to prevent chronic hospitalization, first by cutting the admission rate through a screening clinic in the community and later by introducing alternate modes of care after hospitalization. The third was an attempt to avoid hospitalization altogether through a form of home treatment. The fourth sought to avoid

chronic deterioration by using all types of services to facilitate community care, using short episodes of hospitalization as a means of postponing family and community rejection of the chronically handicapped patient.

It is almost impossible to conceive of a planned experiment to test the effectiveness of moving from a custodial to a total push-type of special institution with different relations to the community that would be on the double blind model. What error in the assessment could be introduced because of this weakness? The effect of moving to another staff group in another locus is not controlled for. Change in environment might produce as much effect. In some cities, it might be possible to arrange group transfers in the guise of solving some administrative problem where there was no intention to improve the patient's treatment. Such moves do occasionally occur for administrative reasons, such as when hospitals are redistricted. In practice, it would be difficult to use these opportunities to study the effect of that kind of move as a neutral move to contrast with the move to the specially designed program. Administrative moves do not usually have completely neutral effects on the staffs involved. The receiving staffs tend to screen suspiciously the patients being transferred as illegal lemons being dumped onto them. The receiving staffs' reactions can be one of resentment leading to a slowdown at work and let the administration suffer the consequences or, in contrast, they can take the attitude that they will be stuck with these newly transferred patients and do their best to get them well enough to be

transferred to a different service. Nonetheless, comparable data regarding comparable groups of patients simply moved from one institution to another would be helpful in interpreting the data. Another feature of this trial, which cannot be assessed because of the absence of anything like a blind control, is the enthusiasm of the special staff and the effect it has on the social grouping of the patients as they become formed into the new context. The program may be less important than the fact that it exists and is an experiment. This phenomenon is known as the "Hawthorne Effect" first described by Roethlisberger and Dickson at the Hawthorne plant of General Electric. The fundamental principle elucidated in those studies was that among a group of industrial workers performance levels improved with every environmental change, whether the change was in the direction of improved or worsened conditions from a long-range point of view. Thus, stepping down the level of illumination led to a short-term improvement as much as improving the illumination. The fact that the staff is in something "new" could have as much effect on what happened to the patients as the particular nature of the new program. This effect can only be controlled by having the operation become routinized. Some pharmacology professors have enjoyed telling medical students that when a new drug comes on the market the doctors should hurry up and use it while it is still effective. One must beware of a similar phenomenon regarding new types of psychiatric intervention even when no new drug is involved.

Another weakness of these studies is that they may understate the effect of the new program because the atmosphere of change could affect the staff at the institution from which the patients came. A competitive atmosphere in the old hospital could make the staff there wish to show that they can do at least as well as the experimental staff. Keeping the program secret from the staff of the hospital from whence the study population was drawn is not absolutely impossible, but would require a great deal of organizational preparation, a slow drawing off of selected patients whom only the investigators knew were associated with left-behind controls. In most mental hospitals it would be almost impossible for no rumors to leak from one locus to the other. In the Massachusetts Hospitals Experiment all eligible patients in Metropolitan State Hospital were used in the study and the staff there was very conscious of the experiment; in Boston State Hospital the experimental activity was not widely known. This is not the same as saying that the exact nature of the experiment would become widely known in the larger institution. It only means that the larger institution staff would develop an ideology about what was going on and would presumably react to it. It would be an extremely difficult and expensive undertaking to monitor these changing perceptions of what the study was about. Ordinarily, it would hardly be justified. It is important, however, to take all these factors into consideration when reading reports of this type of study—the authors should give the readers enough information to let the reader form a judgment as to

how much of the observed effect might be attributed to these uncontrolled factors. The intervention is planned to produce a change in the course of events within the mental and behavioral life of the selected patients; one can only know what would have happened if there was no intervention if one can observe a comparable group of patients for whom that intervention did not occur. That is what the control or comparison group is for—to tell us what difference the intervention made. A study with no control is like a compass without a needle—it can lead you anywhere because it has no sense of direction.

The Worthing Experiment

Preventing Institutional Neurosis by Preventing Hospital Admission

Another type of innovation occurred in assessment methodology when some programs were developed to prevent hospital admission as a means of preventing institutional neurosis or institutionalism. In England, beginning in 1958, Sainsbury and Grad studied the service organized by Joshua Carse and John Morrissey, which was referred to as the Worthing Experiment. The Graylingwell Hospital served several districts, one of which was Worthing. An outpatient service was established and a rule made that no patient would be admitted from Worthing to the hospital without a full assessment by the outpatient staff—which led to a rapid drop in the annual number of

admissions. Then another experiment was set up for the Chichester district, served by the same hospital. The first experiment had been simply to reduce admission rates through a screening clinic. The second experiment undertook to achieve that objective too, and to reduce institutionalization on a long-term basis, to facilitate care in the community, and to make the optimal disposition of the individual referral. This experiment, of course, contained a much broader set of objectives than that of simply reducing the annual admission rate. Sainsbury and Grad did an elaborate study contrasting the experience of psychiatrically referred patients from Chichester with those from Salisbury, a third district of the same hospital in which no experiment was undertaken.

The data gathered in an attempt to assess the effectiveness of the modified psychiatric intervention in Chichester is well worth reading as an exercise in the difficulties of getting the relevant data. The main lesson to be derived from the experiment was probably stated by Cecil Sheps in the published discussion of Grad's and Sainsbury's report: "I feel that if you are really going to evaluate services, the single most important prerequisite is specificity of objective, and this has been lacking. . . ." That statement assesses the effectiveness of the investigators' assessment of certain psychiatric interventions. The lesson is particularly telling because of the extraordinarily high level of scientific work the investigators carried out. My own view is that they were caught at a moment in the transformation of goals in innovative psychiatric intervention patterns when it appeared that a specific set of goals

had been defined, but by the time the experiment was over, other experiences in other locations had made it appear that the beginning goals of the Chichester experiment were not sufficiently specific. This is a hazard which all social experiments face; one does not know what is going to happen in the broader world while the study is being executed that will make one's own study look quite different from the later perspectives. No social experiment can count on staying an island unto itself, which is not an argument against such social experiments, only an argument for reasonable caution and the courage to risk being outpaced by events. These investigators exercised both. There is also a lesson to be learned from their study in that the utilization rates for psychiatric services in Chichester turned out to be consistently higher than for their control community Salisbury, which increased the difficulties of interpreting the differences between them.

Why did they not use a standard preventive trial design instead of two whole communities, one for the experiment and one for the control? It seems clear from the description of the experiment that they could not have done so because the psychiatric intervention in Salisbury was to be carried out by a single team of clinicians who provided the pre-care screening and consultations, the hospital unit's inpatient care, and the aftercare in the community and referral to the appropriate community agencies and close working relations with the local general practitioners. It is hard to see how a psychiatric service of this kind can develop smooth consultative relationships

with the local general practitioners regarding a random half of the case load while maintaining a distant intermittent contact with the same practitioners regarding the other half. If this is a crucial characteristic of the pattern of psychiatric intervention being tested, the random case assignment design simply will not work.

What were the events that overtook this study's statement of objectives? The notion of finding the best disposition for each psychiatric referral was overtaken by the spread of the notion that the best way to promote community care for people with chronic severe mental disorders was to maintain a continuing watching brief, regarding each current arrangement for each patient as the best for the moment and being prepared to change it on very short notice and without waiting until something very unsatisfactory forces a change. Francis Pilkington in the discussion of the Chichester findings said that an appraisal of the family's burden one month after the first psychiatric contact was unrealistic—what is not too hard on a family for one month can become intolerable a few months later.

A mixture of purposes was involved that also made assumptions about means and ends. One of the main motives for the two experiments was that the hospital was becoming overcrowded and it was assumed that the most efficient way of reducing hospital census was to reduce the number of annual admissions. This apparently logical approach, turned out not to be the means

by which a drop in mental hospital census has occurred in general. Actually, the assumption that a rising mental hospital census is bad was never fully examined.

The Louisville Experiment

An Attempt to Avoid Hospitalization

Another key experiment in the prevention of hospitalization was conducted in Louisville, Kentucky, by Pasamanick, Scarpitti, and Dinitz from 1961 to 1964. It throws further light on the issues involved in assessing the effectiveness of psychiatric interventions designed to prevent hospitalization. In this experiment the clinical trial design was used, that is, a stream of patients thought to be suitable for a potentially better but unproved new form of intervention were randomly assigned to the older treatment and the new treatment. A stream of patients being inducted into the ordinary mental hospital treatment program for Louisville were randomly split into a group that would receive drug treatment and outpatient care with intensive visiting from a corps of specially trained public health nurses. Actually, this second group was twice as large as the hospital intervention group because a second trial was built into the design. The investigators not only wanted to know whether the intensive home attention with drugs was more effective than the hospital type of intervention, but also whether the home care with drugs

given to the patients was better than the same home care without drugs, that is with a placebo. Thus, there were two experimental groups and one control group getting "ordinary" care.

When we assess the effectiveness of these three types of psychiatric intervention we must keep in mind Sheps's comment quoted above that "the single most important prerequisite is specification of objectives. . . ." The objective here was to prevent hospital admission very specifically. Thus it is easy to assess the effectiveness. While those permitted to use the hospital form of intervention were of course 100 percent hospitalized, less than one in four of the drug home care group were hospitalized and about one-third of the placebo group were hospitalized. This is a definite index of success in accomplishing the stated objective. The report merits careful study because of the meticulous detail with which the relevant facts are recorded and the crucial nature of the investigation, which shows that intensive home care can prevent hospitalization and also that the currently available drugs really did increase the effectiveness of the home care intervention. It is the first good data about the effect of drugs on schizophrenics on home care, which indicates their definite effectiveness in helping to prevent hospitalization. The data further indicate that these drugs also helped reduce the prevalence of certain troublesome behaviors.

These are the two major studies based on the notion that prevention of

hospital psychiatric intervention is a worthwhile objective. In both instances substantial evidence exists that appropriate psychiatric intervention outside of the hospital with a variety of supplementary services can prevent a substantial proportion of hospitalizations.

The Louisville experiment has a major weakness. The objective in this instance seems clearly enough stated, but examination of the study method reveals that the three samples were selected *after* the "patients had arrived at Central State Hospital and were placed on the admission ward with no immediate treatment." This is quite a different concept of "preventing hospitalization" than the Worthing and Chichester experiments. In fact, it was an attempt to prevent hospital intervention as the initial treatment plan for the patients, but all study subjects began the current episode of treatment with a hospitalization. The publication reporting the Louisville experiment repeatedly uses the term "hospitalization" as interchangeable with "institutionalization," which gives us the needed clue to understand what their stated objective meant to them. "Hospitalization" is an aspect of that process which leads to "institutionalism," so the object of avoiding hospitalization and finding an alternate mode for psychiatric intervention is related to the more distant objective—to prevent institutionalism. It is then necessary for the student of methods to assess the effectiveness of particular types of psychiatric intervention to take another look at the Louisville experiment: Is the student willing to grant the assumption that prevention of

hospital intervention at the beginning of a particular episode of psychiatric intervention is the best way to prevent institutionalism? If this assumption is granted, then the effort to provide intensive home care certainly achieved a step toward that objective, but the reported data do not include information as to whether chronic institutionalism was prevented or not. Nor in the five-year follow-up is there information on the frequency of long-term institutionalization or institutional neurosis. The groups were probably too small to appraise this type of phenomenon in any case. But the authors' summary states that "eventually no differences in psychological or social functioning could be found. This indicates a need for the structuring of community mental health services on an intensive aggressive basis. . . ." This result followed a phasing out of the specialized services designed for home care— which indicates that the program was effective in avoiding hospitalization for a period, but that in the long run the patients did just as poorly as those who had received hospital care initially.

These findings suggest that the anti-hospital type of intervention did not produce the long-term effect of preventing deterioration in personal and social functioning. The other side of this statement is apparently also true: that initial hospitalization did not worsen the long-term course of those patients who were hospitalized as compared with those who received the intensive home care in the first three years, whether or not they received drugs at that time. The long-term success or failure of the patients, then, does

not seem to be crucially affected by which type of care is given during one short period of their chronic disorder. Consequently one must ask whether preventing hospitalization or facilitating community care is the more worthy objective. Each showed some success in terms of the stated goals, but the issue that the student of effectiveness assessment must face is whether either really stated its ultimate objective.

We can infer that the stated goals can be seen as intermediate rather than ultimate. Avoidance of hospitalization need not be seen as an end in itself. Community care need not be seen as an end in itself. Without attributing unstated views to the authors cited, we can ourselves conclude that seeking these effects— preventing hospitalization and facilitating community care—can be seen as part of a strategy not fully made explicit for accomplishing another objective: preventing chronic deterioration in patients with severe chronic mental disorders. Leaving the Chichester experiment (1959-1963) and the Louisville experiment (1961-1964) behind, let us ask what kind of psychiatric intervention would be appropriate if the effect being sought was to prevent chronic deterioration?

The Duchess County Experiment

An Attempt to Prevent Chronic Deterioration in the Severely Mentally Ill

In 1958 Robert C. Hunt advanced the following propositions:

1. The disability associated with psychotic mental illness is enormous.

2. The illness and the associated disability are not necessarily homogenous or synonymous.

3. Disability is only in part intrinsic to the illness.

4. Disability is in large part an artifact of extrinsic origin.

5. Since the disability is an artifact it is not inevitable and something can be done about it.

6. The factors which produce disability are multiple.

7. The multiple extrinsic factors have a common origin in traditional attitudes toward the mentally ill in our culture, [p. 10]

With that set of notions, he proposed to initiate an experiment to prevent chronic disability among all the severely mentally ill people of Duchess County, New York, which I shall refer to as the Duchess County Experiment.

The effect he proposed to produce through a modified form of psychiatric intervention was to reduce the amount and severity of chronic disability due to serious mental disorders. The modification of psychiatric intervention he proposed had its roots in observations he had made

regarding the effects of certain psychiatric service programs he had studied. It is necessary to recapitulate those observations briefly so that the reason why he picked this effect can be properly understood. His notion that chronic disability could be prevented arose when observing psychiatric intervention programs that did not have that objective. A series of "improvements" in mental hospital management had, apparently, produced the unexpected effect. These improvements began shortly after the Second World War in three different British mental hospitals and were initiated more or less in parallel and independently by three British mental hospital directors. They all began in the same way. Each director was impressed that the hospitals used entirely too many locks and they started, conservatively, to unlock some of the wards, at first only in the daytime. They found that the patients not only did all right, but they seemed to do a little better. So they went on. G. M. Bell at Dingleton Hospital in Melrose, Scotland, was probably the first to unlock all wards. Not only did the patients like it and do better, but the staff had come to like the greater responsibility they were carrying. T. P. Rees at Warlingham Park Hospital, Croyden, was pushing along in the same direction. He was particularly impressed by how much the staff gained from the changes. He developed the theory that the more responsibility staff and patients were given, the better they performed. He said that all the patients could handle more responsibility than they had had and that when they took responsibility, they began to improve. Duncan Macmillan at Mapperley Hospital,

Nottingham, had been impressed during the Second World War, that when a wall of the hospital collapsed during the bombing the patients did not run away, but pitched in, helping clean up the debris and comforting those who had been hurt, not very differently from the way the nonhospital residents of the city did after a bombing. He was particularly impressed by the effect on patients of locking the door when they were admitted to the hospital; he felt that the patient responded by a major loss in self-confidence and optimism that interfered enormously with his ability to mobilize his resources as the psychotic episode began to recede. Macmillan was also particularly impressed by how his gradual unlocking of wards and permitting greater responsibility and freedom for the patients had changed the staff-patient relationships. "The staff have to use their personalities to deal with situations which were formerly dealt with by the locked door . . . ," he said.

By 1953, when word of these developments first reached the United States, all three hospitals had been totally unlocked for several years. Macmillan had gone so far as to use the first postwar funds he received to improve the appearance of his hospital's wards, to have all the ward doors removed and installed double swing pantry-type doors at the entrance to each ward, which could not be locked under any circumstances. Hunt met Rees when the latter was on a visit to America and was deeply impressed with his accounts of how the hospitals had become transformed. Patient behavior had improved radically, the worst forms of aggression, soiling, self-

neglect, mutism, refusal of food, and so forth had become much rarer. And new cases hardly ever developed these patterns of withdrawal and self-neglect. In the three hospitals the census of mental patients had dropped dramatically, almost 50 percent in less than a decade. The hospital staff had moved a large proportion of their work into the community where they took care of as many former hospital patients as they did in the hospital. It all sounded very nice and a bit too good to be true. The reader should remember that these reports arrived in this country in 1953; Henri Laborit had discovered chlorpromazine in 1951, and the first psychiatric meeting at which its results were described was in 1952. All of these advances were made before any of the so-called tranquilizing drugs were available commercially. To anyone with experience in mental hospital work, Rees's reports were unbelievable and Hunt had worked in New York State Mental Hospitals for seventeen years. But it was an interesting story even if only half true and when a year or so later the World Health Organization offered Hunt a traveling fellowship to study these programs, he accepted with alacrity. His report stated that the programs did work exactly as specified and with no fakery. The tranquilizing drugs were available by then and were being used, but he was impressed with the relatively low dosage, the rapid census drops, the absence of disturbed and deteriorated behavior of the worst sorts, which still existed in other mental hospitals despite much higher drug dosage patterns.

After a series of steps recorded elsewhere, Hunt determined that he would try to copy the fundamental principles involved and to organize a county service for the immediate area of Hudson River State Hospital of which he had become director. He had to modify the hospital organization because it was clear that the British successes had occurred in small hospitals, (starting with censuses of under 3000) with small catchment areas near the hospital.

In 1958 Hunt proposed an experiment in which he would allocate a proportional part of his large state hospital (more than 6000 patients) for the residents of Duchess County (10 percent of his catchment area's population). One-tenth of the plant and personnel would be used to set up a sub-hospital with comprehensive responsibility for providing the indicated psychiatric interventions for all of the seriously mentally disordered residents of the county. In addition to providing all needed acute and long-term inpatient care, this staff was charged with providing all after-care, social service, and family care, and all local facilities and related professionals were encouraged to cooperate in any possible way to maximize the community care of the patients. A day hospital unit already existed at the hospital. The only new service introduced was "pre-care," which meant consultation by hospital psychiatrists about patients in the community who were thought to be in possible need of mental hospital admission. The British pioneers had insisted that this was a crucial feature not only to prevent misuse of hospitals, but also

to help the patient accept voluntary hospital care. A slight increase in personnel was introduced (I estimate less than 10 percent of the annual cost to the state of running the mental hospital county unit). This extra personnel was financed by the Milbank Memorial Fund and met a few elementary needs: (1) a unit director at the rank of what might be considered associate director of the mental hospital had to be brought in, as none of the existing staff had the range of competences needed to run such an enterprise and the state would not underwrite an extra assistant director category; (2) an extra secretary for the unit director at a higher rank than the state budget provided for in a service chief; and (3) extra stenographers and social workers because of the increased work involved in communicating outside the hospital.

Five points are worth special attention in this description: (1) the innovation in psychiatric intervention was not a new kind of treatment but an innovation in the organization of personnel and resources and of policy; (2) the idea that the changed pattern of using existing resources would reduce the amount of chronic deterioration emerged out of observing that efforts to improve and humanize mental hospitals had led to rapid drops in census and an apparent major reduction in the frequency of chronic deterioration; (3) in contrast to the Chichester and Louisville experiments, this pattern of psychiatric innovation was an attempt to improve the way in which the hospital was used in the treatment of long-term patients, not a way of avoiding the hospital; (4) the assumption was that the hospital would be

made *more* available to those in need of its services, that admission rates would rise, and that this increase would be more than compensated for by a radical decrease in the duration of hospital stay and much earlier return to home living before complete restoration of all functions; and (5) Hunt no more planned any systematic data gathering to assess its effectiveness than had Bell, Rees, or Macmillan.

But when Hunt asked the Milbank Memorial Fund to provide the small, extra, flexible financing to implement his ideas, they responded by saying that the idea was too important potentially to justify starting it without some systematic evaluation research. I was assigned the duty of locating the research worker, but in the end was assigned the job of organizing and executing the research with a special team.

Assessment of effect in terms of preventing deterioration would present no special methodological problems in itself. It would be necessary to locate or develop some objective criteria for recognizing chronic deterioration and apply these criteria to a suitably selected population at some risk of developing chronic deterioration, a random half of which was given the new form of psychiatric intervention. But the nature of the intervention made this impossible. The reasons were similar to those operating in the Chichester experiment. It is not possible to organize a double blind random assignment psychiatric service highly integrated with the nonpsychiatric community

services for only half the patients from the community while the same professionals remain unintegrated in their work patterns for the other half. It was thought that perhaps another county's patients from the parent hospital's district would make a suitable control population. This turned out to be unrealistic because it was well established that hospital utilization rates fall off with the distance of people's homes from the hospital.

Because this experiment introduced a new way of organizing psychiatric intervention at a point in time, a before-and-after design was considered to be of some help. A before-and-after system has certain inherent weaknesses, however, the main one being that no one knows what unplanned changes will occur over a given time span if no planned changes are introduced. One is studying a phenomenon—in this case the frequency of chronic deterioration—about which remarkably little was previously known. There could well be short-term fluctuations of major size in its frequency that had never been carefully enough observed. It was known that severe chronic deterioration was not the common sequel of a psychotic episode, but there were very little data on which to base estimates of how frequently and how shortly after the first hospitalization it was to be expected.

The first step was to specify the criteria for identifying a person as being severely disabled in the presence of a serious mental disorder. The criteria to be used were developed with considerable caution and special efforts were

taken to see that they would be as relevant as possible to assessing whether the stated goals were achieved. The research workers generated intermittent intensive interactions with the administrative innovators and pushed for highly specific formulations of exactly what would be less common. The expectation was that the main changes would be in behavior, a finding that surfaced several years later in the Massachusetts Mental Hospitals Experiment. It was anticipated in the Duchess County Experiment because of careful observations made at the three British pioneering programs in community care. With some trepidation, it was decided to ignore completely the patient's subjective mental symptoms because Hunt and his associates did not expect much change in these phenomena. Self-care, participation in work and recreational roles, and freedom from dangerous or troubling behavior became the main areas of inquiry. The data gathering focused on those specific changes in behavior that were expected by the innovators. They set the fundamental criteria and the research team through pretests decided which changes could be objectively ascertained with reliability. This was first done on all the county's chronic ward patients in 1959 in order to implement the testing of a subordinate hypothesis: Those chronic patients who were already seriously disabled when the new program started would improve but slightly. A great deal was learned in this process about the techniques of ascertaining whether any given patient met the criteria set in a particular week. In one week more than 18,000 printed data forms were filled out by

ward personnel of Hudson River State Hospital, were edited within eight hours, and the informants were questioned about any discrepancies or missed item the next time they reported for work. That was done before the reorganization of the services had started.

This gave the investigative team confidence in its ability to make the necessary assessment within the hospital at any one point in time. The particular pattern of disabled behavior that was being assessed was given a new name so as to avoid confusion with related concepts such as institutional neurosis and chronicity. The new term was “Social Breakdown Syndrome” (SBS). Though this technology provided a means for ascertainment at one point in time, the major effect anticipated was that new cases of chronic SBS would be reduced in frequency.

To assess whether the new program was reducing the frequency with which new cases of chronic SBS was developing required additional techniques. The onset of the SBS had to be determined for each individual and a mechanism developed to monitor cases of SBS to see how long they continued after they were located. These techniques were also developed.

The most difficult problem was to decide who had to be studied. This was not a case control study in which one could observe two groups of patients, those who had been exposed to the new pattern of intervention and

those who were exposed to the old pattern of intervention. Each of the three previously described evaluations used that method. In this case the entire population of people with severe mental disorders in Duchess County was exposed to the old pattern of intervention until 1960 and thereafter was exposed to the new pattern. It might appear that those whose first entry to the hospital was before 1960 could be compared with those whose first entry was after 1960, but this ignores the well-known fact that chronic SBS often does not develop until several years after the first admission and those who had first been admitted, say in 1958, were provided with exactly the same sort of intervention after 1960 as were people first admitted subsequent to 1960. In addition, the mechanism for reducing the rate at which chronic deterioration develops included making the hospital more readily available for readmissions and this was expected to make the hospital more available for first admissions too. Therefore, admission rates were expected to rise and a later cohort might include people at less risk of developing chronic SBS than earlier cohorts of admissions. Inasmuch as the new pattern of intervention included a plan to release patients to community care very early in their recovery— keeping the hospital prepared for any needed repeat short-term admission—there was a significant risk that the chronic SBS cases might develop in the community without the knowledge of the clinical teams.

The method used to meet these problems was to create a register of all Duchess County residents who had had psychiatric treatment after 1955 and

to keep it up-to-date regarding new entries to treatment and transfers or discharges or deaths. This register population was then looked upon as the population at risk of developing chronic SBS. The problem then was to find a way of determining whether long-term episodes of SBS began more frequently in the years prior to 1960 than in the years subsequent to 1960. The concept "long term" was defined as an SBS episode lasting a full year or longer.

Specially trained data gatherers were sent out to locate those members of the registered population who were living outside the hospital and to determine whether at that time the individual was in an SBS episode. If the answer was yes, then a research clinician went to the individual's location and took a careful history to determine when that episode started. The technique of locating the individuals, gaining cooperation for data gathering, developing reliable onset date determination techniques, and continuing monitoring techniques took more than two years to develop. By that time, the new form of intervention had already been going on for two years. Despite this tardiness in developing the capability to gather the needed facts, it was possible to get data that provided a means for determining whether the number of new chronic SBS episodes starting in later years was smaller than the number that started in earlier years. This was made possible because the hypothesis only referred to chronic cases, and these chronic cases—by definition—would have to continue to be SBS cases for more than a year. By

screening the population at one point in time, one picked up a mixture of new and old

SBS cases. All chronic cases that had started in the previous year would still be cases on the day they were screened. By monitoring cases until they terminated or passed their anniversary, one would get a complete count of one year's onsets of chronic SBS cases. Of course, some of these cases would have started a week or two before the day they were screened, and to find out which of those were going to become chronic it would be necessary to monitor the whole group for a full year after the screening date. By April, 1965, on the basis of many thousands of interviews it was possible to report tentatively that for every two cases of chronic SBS that had begun during 1963 somewhere between three and four cases must have started in 1960. By 1969 more detailed analysis of twice as much data and greater familiarity with the epidemiology of SBS, both acute and chronic, made it possible to report that each year the new form of intervention prevented at least forty person years of chronic severe deterioration per 100,000 general population in the age group sixteen to sixty-four.

The level of disability reflected in SBS is on the whole pretty severe. A person who attempts suicide but is prevented by physical restraint, a person who soils, a person who has no recreational activities or no work activities will meet the criteria; to be a chronic SBS case one or another of these

characteristics must be present every week for fifty-two consecutive weeks. There are combinations of less severe manifestations in each area of social functioning that will also qualify them. One remarkable finding was that only half of the chronic SBS cases were schizophrenic cases, the other half were scattered over a wide range of diagnoses. But the pattern of disability did not correspond to the diagnosis. The SBS syndrome describes a pattern of psychotic decompensation that can occur in any mental disorder and probably does; there is good reason to think that short episodes actually occur in the absence of any psychosis and if they were seen by psychiatrists they might be diagnosed as "transient situational reaction" or "no mental disorder."

These data on the declining incidence of chronic SBS provide the best evidence that community care of chronic mental patients is more than a fashion; it is actually a way of preventing long-term serious disability. The mechanism by which chronic SBS is being prevented is still not completely clear. Most entries into inpatient care follow the onset of an SBS episode, which tends to terminate very quickly after entering the hospital (one to five weeks). Most, but not all, episodes that occur in patients while living in the community result in a hospital admission. Very few of the chronic cases that start are in fact extreme examples of institutionalism or institutional neurosis. It is possible in a population that has been receiving a pattern of community care for more than twelve years to see the forms of chronic SBS

that develop in the absence of chronic hospitalization. This has not been studied in detail.

This is the first example in which a modification of the way in which a health-care delivery system is organized has produced evidence of improved health in the population being served, without introducing either a greater volume of service or a new technology of medical treatment (such as a new drug or operation). It raises fundamental questions about the pathogenesis of the deteriorating syndrome in the major mental disorders. It also requires a rethinking of the function of inpatient services as one element in the network of services used in the treatment of people with long-term serious mental disorders. The pattern of fluid movement between inpatient and outpatient care seems to depend upon a unified clinical team, that is, a clinical team that continues its treatment responsibilities for its own group of patients as these patients move in and out of hospital, family care, and outpatient status. The indications for hospitalization take on a new appearance.

These outcomes are emphasized to indicate that evaluation studies to determine the effectiveness of a new type of psychiatric intervention can sometimes do much more than answer the simple question, Was the effect being sought produced? Fundamental issues in clinical psychiatry can also become elucidated by coming into close contact with a rapidly changing situation. In fact, it is when great changes are in progress that certain types of

studies are most appropriate. Combining research with planned action not only provides information about the effectiveness of the action, but can throw light on previously unquestioned conventional wisdom regarding the nature of mental disorders.

Bibliography

American Public Health Association. *A Guide to Control Methods.* Program Area Committee on Mental Health. New York: Am. Public Health Assoc., 1962.

Barton, R. *Institutional Neurosis,* 2nd. ed. reprint. Bristol, England: John Wright, 1966.

Brandon, S. and E. M .Gruenberg. "Measurement of the Incidence of Chronic Severe Social Breakdown Syndrome—Has the Duchess County Service Been Associated with a Decline in Incidence?" in E. M. Gruenberg, ed., *Evaluating the Effectiveness of Mental Health Services. Milbank Mem. Fund* (1966)., 44, Part 2 (1966), 129-142.

Davis, A., S. Dinitz, and B. Pasamanick. "The Prevention of Hospitalization in Schizophrenia: Five Years after an Experimental Program," *Am. J. Orthopsychiatry,* 42 (1972), 375-388.

Drasgow, J. "A Criterion for Chronicity in Schizophrenia," *Psychiatr. Q.,* 31 (1957), 454-457.

Grad, J. and P. Sainsbury. "Evaluating the Community Psychiatric Service in Chichester: Results," in E. M. Gruenberg, ed., *Evaluating the Effectiveness of Mental Health Services. Milbank Mem. Fund Q.,* 44, Part 2 (1966), 246-287.

Greenblatt, M., M. Solomon, A. S. Evans et al. *Drug and Social Therapy in Chronic Schizophrenia.* Springfield, Ill.: Charles C. Thomas, 1965.

Gruenberg, E. M. "From Practice to Theory: Community Mental Health Services and the Nature of Psychoses," *Lancet,* 1 (1969), 721-724.

----. "Hospital Treatment in Schizophrenia: The Indications for and the Value of Hospital

Treatment," in R. Cancro, ed., *The Schizophrenic Reactions: A Critique of the Concept, Hospital Treatment and Current Research,* pp. 121-136. New York: Brunner/Mazel, 1970.

----. "Benefits of Short-term Hospitalization," in R. Cancro, N. Fox, and L. Shapiro, eds., *Strategic Intervention in Schizophrenia: Current Developments in Treatment.* New York: Behavioral Pubs., 1974.

Gruenberg, E. M., S. Brandon, and R. V. Kasius. "Identifying Cases of the Social Breakdown Syndrome," in E. M. Gruenberg, ed., *Evaluating the Effectiveness of Community Mental Health Services. Milbank Mem. Fund Q.,* 44, Part 2 (1966), 150-155.

Gruenberg, E. M. and J. Huxley. "Mental Health Services Can Be Organized to Prevent Chronic Disability," *Community Ment. Health J.,* 6 (1970), 431-436.

Gruenberg, E. M., R. V. Kasius, and M. Huxley. "Objective Appraisal of Deterioration in a Group of Long-stay Hospital Patients," *Milbank Mem. Fund Q.,* 40 (1962), 90-100.

Gruenberg, E. M., H. B. Snow, and C. L. Bennett. "Preventing the Social Breakdown Syndrome," in F. C. Redlich, ed., *Social Psychiatry,* pp. 179-195. Baltimore: Williams & Wilkins, 1969.

Hunt, R. C. "Ingredients of a Rehabilitation Program," in *An Approach to the Prevention of Disability from Chronic Psychoses,* pp. 9-27. New York: *Milbank Mem. Fund* 1958.

Hunt, R. C., E. M. Gruenberg, E. Hacken et al. "A Comprehensive Hospital-Community Service in a State Hospital," *Am. J. Psychiatry,* 117 (1961), 817-821.

Jarvis, E. "Influence of Distance from and Nearness to an Insane Hospital on Its Use by the People," *Am. J. Insanity,* 22 (1855-1856), 361-406.

Macmillan, D. "Hospital-Community Relationships," in *An Approach to the Prevention of Disability from Chronic Psychoses Milbank Mem. Fund,* pp. 27-50. New York: 1958.

Pasamanick, B., F. Scarpitti, and S. Dinitz. *Schizophrenics in the Community: An Experimental Study in the Prevention of Hospitalization.* New York: Appleton-Century-Crofts, 1967.

Roethlisberger, F. and W. Dickson. *Management and the Worker.* Cambridge, Mass.: Harvard University Press, 1939.

Sainsbury, P. and J. Grad. "Evaluating the Community Psychiatric Service in Chichester: Aims and Methods of Research," in E. M. Gruenberg, ed., Evaluating the Effectiveness of Mental Health Services. *Milbank Mem. Fund Q*., 44, Part 2 (1966), 231-242.