Guidelines for Studies on Substance Abuse Treatment (1981)

Pages
29

Size
8.5 x 10

ISBN
0309330688

Committee on Substance Abuse and Habitual Behavior; Assembly of Behavioral and Social Sciences; National Research Council

Guidelines for Studies on Substance Abuse Treatment

Dean R. Gerstein, *Editor*

Committee on Substance Abuse and Habitual Behavior

Assembly of Behavioral and Social Sciences

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1981

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Research Council was established by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. The Council operates in accordance with general policies determined by the Academy under the authority of its congressional charter of 1863, which establishes the Academy as a private, nonprofit, self-governing membership corporation. The Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in the conduct of their services to the government, the public, and the scientific and engineering communities. It is administered jointly by both Academies and the Institute of Medicine. The National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively, under the charter of the National Academy of Sciences.

COMMITTEE ON SUBSTANCE ABUSE AND HABITUAL BEHAVIOR

LOUIS LASAGNA (Chair), Department of Pharmacology and Toxicology, University of Rochester
HOWARD S. BECKER, Department of Sociology, Northwestern University
PETER DEWS, Laboratory of Psychiatry and Psychobiology, Harvard Medical School
JOHN L. FALK, Department of Psychology, Rutgers University
DANIEL X. FREEDMAN, Department of Psychiatry, University of Chicago
JEROME H. JAFFE, University of Connecticut School of Medicine and Veterans Administration Hospital, Newington, Connecticut
DENISE KANDEL, Department of Psychiatry, Columbia University, and New York State Psychiatric Institute, New York
JOHN KAPLAN, School of Law, Stanford University
GARDNER LINDZEY, Center for Advanced Study in the Behavioral Sciences, Stanford, California
GERALD McCLEARN, College of Human Development, Pennsylvania State University
CHARLES P. O'BRIEN, Drug Dependence Treatment Service, Veterans Administration Hospital, Philadelphia, Pennsylvania
JUDITH RODIN, Department of Psychology, Yale University
THOMAS C. SCHELLING, John F. Kennedy School of Government, Harvard University
RICHARD L. SOLOMON, Department of Psychology, University of Pennsylvania
FRANK STANTON, New York, New York
ALBERT STUNKARD, Department of Psychiatry, University of Pennsylvania
RICHARD F. THOMPSON, Department of Psychology, Stanford University

Staff

Peter K. Levison, Study Director
Dean R. Gerstein, Senior Research Associate
Deborah R. Maloff, Research Associate
Marie A. Clark, Administrative Secretary

CONTENTS

PREFACE

The National Institute on Drug Abuse (NIDA), along with its sister agencies in the Alcohol, Drug Abuse, and Mental Health Administration of the U.S. Department of Health and Human Services, has the dual mandate of promoting research on and treatment of substance abuse. In 1976, NIDA asked the National Research Council to form a Committee on Substance Abuse and Habitual Behavior to investigate and advise NIDA on important scientific issues related to NIDA's overall agenda, including research on the treatment of substance abuse.

During its initial two years, the Committee reviewed and organized a broad range of scientific data about a variety of drug-involved and nondrug-involved compulsive behaviors, preparatory to focusing on a narrower set of studies. In November 1978, five Committee work groups were organized, one of which was to prepare recommendations regarding NIDA's research program. This research recommendations work group included experienced past advisers to NIDA's research programs who were also familiar with exemplary local studies on substance abuse treatment. In its work, the group became convinced that the overall quality of treatment research proposals to NIDA, and NIDA's ability to promote good studies on treatment, suffered from the absence of a clear set of guidelines fully accessible to NIDA's managers, review groups, professional clientele, and overseers in the executive branch and in Congress. Too often, such activities as "clinical studies," "applied research," "treatment evaluation," "demonstration," "program evaluation," "policy analysis," and the like are confused or misidentified by officials and managers who have some role in the administration, oversight, or funding of NIDA's activities.

The work group began with a brief account of problems in maintaining continuity in treatment research programs, including some specific recommendations regarding NIDA's current activities. This account was presented to the Committee's annual conference in June 1979, at which time detailed comments were received from other Committee members and from three invited discussants: Michael Agar, an anthropologist familiar with a broad range of substance abuse research settings; Carl Chambers, a sociologist with wide experience in the political and scientific aspects of substance abuse treatment research; and Elmer Struening, a psychologist with expertise in program evaluation methods. Based on

GUIDELINES FOR STUDIES ON SUBSTANCE ABUSE TREATMENT

INTRODUCTION

Research on and evaluation of treatment programs for substance related disorders--involving abuse of marijuana, heroin, alcohol, ethical drugs, and so forth--are topics charged with emotion and political controversy. Even people committed to the dispassionate, objective ideals of science have recognized and been compelled to respond to these emotional and political pressures. Partly because of those pressures, in the 10 years since federal funds have been heavily committed to research on and treatment of substance abuse, it has been difficult to develop the scientific knowledge necessary for rational decision making. Limited budgets and constituency concern have in the past often led to departures from scientifically optimal procedures, and studies appropriate for one stage of program development or implementation have been used for other stages while certain necessary studies have not--or have rarely--been performed.

This paper sets out some recommendations regarding research on treatments for drug problems. (In this report, the terms "drug" and "substance" are used interchangeably to mean psychoactive compounds of established abuse potential.) We first outline the different types of studies involved in developing and implementing new treatments for substance abuse. We then offer specific, detailed recommendations about two research areas: natural histories of drug use, which have been relatively neglected; and field replication of new treatments, which has been subject to great confusion. In the first case we hope to stimulate expansion; in the second case we hope to stimulate clearer thinking.

STUDIES ON NEW TREATMENTS FOR SUBSTANCE ABUSE

In this section we sketch here an ideal progression of studies that should occur as a new treatment for a drug problem is conceived, tested, modified, and discarded or implemented on the basis of its merits. We identify six different types of studies, each suited to a particular step in the development and subsequent implementation of a new treatment. The particular order and even the inclusion of each type of study are less important than that the distinctions between the different kinds of

studies are understood. Some of these studies are scientific research; others use some of the methods of science but are not what is generally meant by scientific research; still others are related more to accounting, law, and management than to scientific research. But all of them are important aspects of modern treatment for substance abuse, and they need to be properly identified and defined in order for our recommendations to be clearly understood.

Development Studies

Basic Research

Underlying all new developments in treatment research are advances in general scientific knowledge in the biological, behavioral, and social sciences, such as new understanding of brain metabolism, behavioral conditioning, and cultural control of patterns of habitual behavior. Research advances in the epidemiology of substance use and abuse, which identify the pattern of particular drug-using practices in the population at large and in various subgroups, are likely to have particular relevance. Examples include studies of the recent expansion in the use of cocaine in relatively affluent social strata and of phencyclidine (PCP) among juvenile populations and public response to the environmental health implications of tobacco smoking.

Yet it is impossible to predict exactly which basic research projects will pay off in specific applications to social problems or goals. From the point of view of the economics of science, basic research is a long-term investment in the production of knowledge, and attempts to decide allocations of funds to basic research according to any set of short-term productivity estimates can only distort that investment. A flow of scientific innovations that may lead to or be incorporated in substance abuse treatment is best assured by maintaining a broad base of basic scientific research, within which the interests and ideas that researchers pursue are subject to control principally by considerations of quality in their methods and logic.

Applied Clinical Research

New treatments usually emerge on an experimental basis in an established clinical center. (The obvious exceptions are found in the history of voluntary self-help treatments for alcohol and opiate addictions--e.g., Alcoholics Anonymous, Synanon.) Ordinarily, the treatment is initially developed on a case-by-case basis to establish its features well enough for a set of standard procedures to be assembled. During this development process, different methods are used in order to generate and examine hypotheses and to explore different combinations of treatment features. This development process can include: intensive staff discussions with and about the recipients of the new treatment; "open trials," in which the subjects are evaluated before, during, and after application of treatment; naturalistic observations of the course of both

the disorder and the treatment regimen; or small-scale controlled trials that may, for example, involve crossovers between one treatment and another. The same development process is involved whether the treatment includes some material commodity, such as a drug or device, or is essentially a series of interactional episodes, such as therapy groups or occupational assignments.

If the investigators are encouraged by the preliminary results--an encouragement that must be based on their previous practical, clinical, and research knowledge about how the subjects of the treatment would otherwise have fared--they generally initiate some form of clinical trial. A clinical trial involves a larger number of subjects--30-100 subjects, sometimes more--so that a sound statistical foundation for determining the probable efficacy of the treatment can be established. Such a trial may be performed on a residential basis, with a comprehensive structure of staff attention. During this period different parts of the treatment (duration, schedules of administration, ancillary services, differential staff allocations, etc.) are varied, and preliminary attempts may be made to partition specific components of treatment and determine their contributions to the outcomes.

The process just described, from emergence of the treatment through its careful exploration, can be called applied clinical research. It is important to keep clearly in mind that this process is not basic research, but that, without a basic research foundation, the flow of clinical innovations and applied research to understand them will slow down and ultimately cease.

Field Replication Studies

If a treatment seems promising after applied clinical research--if there are solid indications of clinical effectiveness and no indication of high toxicity, low appeal, uncontrollable costs, or other negative features--the procedures are detailed to lay or professional peers and review groups. This initiates a period during which a number of different sites use the treatment so that experience of it is broadened to several hundred or several thousand cases. Although several treatment centers may be involved, it is not likely that these centers will constitute a representative mix of all potential treatment sites. The sites involved in such testing are usually university-affiliated, major urban, or research-oriented clinics or programs.

The wider trials generally involve evolutionary changes in the treatment plan, due in part to the diversification of investigators and in part to the larger number of subjects. This is desirable, but it is important that the design be kept constant for each particular trial. This is the step during which tight experimental controls are most important. Randomly assigned comparison or control groups may receive no treatment, established treatment methods ("current practice"), or placebo-like methods.

If the treatment fulfills its promise in this multiple-site replication and proves in some way superior to other available treatments, the result will be a set of closely related treatment

descriptions or protocols, including guidelines or criteria for differential diagnosis and application. Only if the treatment has been carefully defined and controlled during this step can outcome standards be developed, against which subsequent efforts to deliver the treatment more widely can be measured. If this step has been bypassed for some reason, or not done properly, powerful obstacles will have been created to making the kinds of assessments necessary in the next phase. Because the main intent here is to learn whether a wider group of patients and investigators can replicate "in the field" the methods and results of the original group, we call these investigations field replication studies.

Implementation Studies

If a treatment has proven both effective and deliverable through field replication studies, it is ready for acceptance as a standard therapeutic practice. At this point the new treatment can be adopted into the regular curricula of relevant professional schools, training institutions, or on-the-job instruction; written into professional or administrative codes; and possibly underwritten by government or private insurance. At this stage, delivery of the treatment is diffused to hundreds, perhaps thousands of sites.

It should be noted that even when a treatment has reached the stage of widespread implementation, it is not necessarily fixed in a timeless mold. It is always subject to "after-market" innovations, which must be tested through the same steps of clinical and replication research specified above. In fact, the "new" treatment with which this discussion is concerned is usually a mutation or modification of an established procedure rather than a radically different approach.

Once implementation has begun, three distinct types of studies are appropriate. These are identified, approximately as used here, in the recent report of the President's Commission on Mental Health (1978:1802).

Program Information

In any unit so large that its ongoing operations cannot be followed or monitored in detail by one person, there is a need for internal data collection and analysis for managerial purposes. This activity includes auditing, monitoring, time accounting--processes that make the critical aspects of their staff's work visible to program managers, so that the managers can make effective and realistic plans and decisions. We do not call this activity "research," although it is often included under the rubric "evaluation research." Rather, this is a program information function, which entails internal analysis of the distribution and effectiveness of services provided. The burden of data collection and analysis for this purpose falls most directly on program staff or consultants who create and maintain management information systems.

Quality Assurance

There is always a need for assuring the quality of treatment; indeed, outside interest in specific programs is usually concerned primarily with quality. The question here is whether a service delivery unit is following accepted procedures to guarantee that the best known treatments are being properly used. Such assurance is generally provided through an external audit in the form of an "accreditation review," occurring periodically in relation to granting licenses, subsidies, or special treatment contracts.

Public Policy Analysis

The questions that arise in connection with expenditure of public funds are: How much better off on the whole is the public with this treatment than without it? Does (or will) the treatment as implemented yield the expected results--not only for the persons treated, but also in expected effects on problems identified as public concerns? Most important, given the enormous competition for public funds, can providers of this treatment continue to justify their qualification for government support in terms of either long-term or short-term government commitments? These are very different questions, involving distinct modes of inquiry. The study of such questions is generally called public policy analysis.

Summary

We have identified six kinds of studies related to drug abuse treatment, each with a distinct form and purpose:

- Basic research, to respond to the empirical and theoretical interests of respective scientific disciplines.
- Applied clinical research, to improve and develop innovative forms of treatment for drug-related disorders.
- Field replication studies, to test the efficacy of treatment innovations beyond the original test site, with a variety of personnel, procedures, and subjects.
- Program information, to collect and analyze the internal data about program operations that are essential to managers for good decision making.
- Quality assurance, to compare the practices of a service delivery unit with professional and legal standards.
- Public policy analysis, to compare the outcomes of a treatment with those that might be produced by alternative allocations of government authority and resources.

THE ROLE OF THE NATIONAL INSTITUTE ON DRUG ABUSE

Responsibility for supporting and guiding these six kinds of studies can not be carried by a single agency. No one funding source can support all the basic research in a scientific discipline, nor can any central agency be responsible for the internal management information needs of all treatment programs. It is our intention here to draw attention to those areas in which one agency, the National Institute on Drug Abuse, should take major responsibility.

Development Studies

Basic Research

The basic research programs at NIDA have begun in recent years to organize, rationalize, and provide reliable support for biobehavioral and sociocultural research that may advance understanding of substance abuse. But development of epidemiological studies has been limited. While NIDA has provided support for a few exemplary studies of drug-using practices (such as activities related to consumption), there has not been a continuing program of research on the changing supply of substances and the activities and transactions between drug-using populations, treatment providers, and other institutions.

Certainly, NIDA has limited resources to support basic research, and we cannot attempt here to review NIDA's entire basic research program, either to affirm or dispute specific program decisions. However, from the point of view of treatment for drug abuse, we are convinced that NIDA should place greater emphasis than it has in the past on studies of natural histories of drug-taking (see below).

Applied Clinical Research

The development of innovative treatments cannot be a routinized process; new treatment procedures usually result from a fortuitous but happy combination of factors. At best, one can try to keep several key factors--for example, knowledge of basic research advances, diverse clinical populations, concerned investigators, supportive institutions--in reasonable proximity to each other and to avoid choking off innovation unintentionally as a consequence of bureaucratic routine. In general, we believe that the multidisciplinary clinical research center is an important seedbed for treatment innovation and that NIDA should support such centers.

Field Replication Studies

Sharp distinctions must be maintained--and have not been--between applied clinical research, field replication, and program information. As we

described above, applied clinical research should be boldly exploratory and small in scale; it is not a necessary part of any clinical unit except specially supported research clinics or wards. And program information should be a routine activity as large in scale as the organization in which it occurs; it is essential to good management in any clinical unit.

Field replication is different from both of these. It is an attempt to test a still-new treatment on a medium or large scale, often spanning a number of organizations; it requires procedures of program control, design, and data collection that are different and much more elaborate than routine program management. In particular, a field replication study requires that a service delivery unit be organized as an experimental station (although without sacrificing the standards of clinical care to which any such unit should always be subject). We set out below specific recommendations about how such studies should be carried out.

Implementation Studies

Program Information

While we recognize the need for good program information, we are concerned about attempting to centralize this function. We suspect that centralization tends to reduce the ability and responsibility of program managers to make timely use of such information. Centralization may also tempt a central unit to substitute manipulation and analysis of such information in place of field replication research, quality assurance, or public policy analysis. Hence, we recommend that NIDA not consider program information a major responsibility.

Quality Assurance

Assurance of quality is an extremely sensitive topic, and we cannot here provide the comprehensive analysis that it deserves. Such matters as the use of "consumer satisfaction" surveys and global clinical rating scores to assess program effectiveness, and the accreditation and professionalization of drug abuse treatment workers, deserve careful study and informed action. We only insist that this activity be clearly distinguished from the other kinds of activities in the process of development and implementation of new treatments.

Public Policy Analysis

The most untenable situation in the development of a treatment innovation occurs when its support becomes a question of public policy too soon: before there is good data about the likely magnitude of a problem, about the field-tested effects of the new treatment, and about the overall quality of the systems available to deliver the treatment. These are not

all the data that policy analyses require, but if such data are not available, one can have little confidence in the outcomes of the analyses. Of course, public needs do not parallel scientific progress, so it is difficult to keep promising work on the back burners simply because it is not yet fully tested. Nevertheless, NIDA should be responsible for carrying out the appropriate studies of innovative treatments even (indeed especially) when they are prematurely adopted in response to political pressures.

RECOMMENDATIONS FOR RESEARCH

Of all the areas in which it might be possible to spell out research criteria or recommendations, we have chosen two for special emphasis. The first is a part of basic research; the second concerns field replication research.

First, we recommend that NIDA commit itself to the support of long-term studies on the natural histories of substance-using practices, norms, knowledge, associated activities, and consequences. These studies, which should be based on ethnographic, longitudinal, demographic, survey, and life-history interview methods, should be established in a number of communities to maximize the chances for increasing basic comparative research knowledge about drug use and abuse. This type of research has been immensely significant both for clinical and public policy developments, yet NIDA has not made it an important part of its overall research program.

Second, we recommend that NIDA develop and apply a set of explicit scientific standards governing any proposal for field replication research (as defined above). These standards should emphasize the basic scientific principles required for drawing sound conclusions from such studies. We are convinced that this stage is the weakest methodological link in the chain of treatment development and suffers primarily from a lack of clarification and commitment to scientific standards.

Natural Histories of Substance Use

When not based on broad and well-constructed foundations of scientific knowledge, understanding of substance use and treatment is easily misguided. Knowledge of substance use and abuse has derived from limited data bases: from populations, often studied for short terms, of the relatively small proportion of people for whom substance use has already led to serious clinical or criminal identification. Much federal effort has been expended to increase this sort of data base, such as the Client Oriented Data Acquisition Program (CODAP), the Drug Abuse Warning Network (DAWN), and the National Drug-Alcohol Collaborative Project (NDACP). There has not been a comparable effort to develop data outside such institutional contexts. Moreover, one cannot get good information about the behavior of interest from annual cross-sectional probability samples of school or household populations.

When we refer to natural histories, we mean the detailed study of patterns of substance use and related consequences across individuals' lives, with systematic attention to variations in the social structure and cultural significance of drug use. This concept of natural history has been of immense importance in behavioral and social science research, particularly in application to the use and abuse of substances. The basis for this approach has been nicely defined by the sociologist Lee Robins (1980:7-8):

> Unlike schizophrenia, which is a rare disorder but one which is recognizable in every culture and in every historical period, drug abuse has emerged as a series of "epidemics" of abuse of different drugs affecting different age, sex, and socioeconomic groups at different historical times and in different countries. As the groups affected vary, the natural history may vary, just as the natural history of measles differs in adults and children, and in children who are chronically undernourished as compared with those who are well fed. The particular drug or drugs abused may each have its own natural history of abuse, as well. To take an analogy from the infectious diseases, to attempt to talk about a natural history of drug abuse may be equivalent to trying to describe the natural history of "infection" rather than the natural history of particular infectious diseases. As both agent and host vary over time and place, our description may be accurate only for a particular moment in time and a particular location. Thus while we can describe the natural history of schizophrenia with some confidence as a rare disorder having its onset in young adulthood, and having a chronic course if untreated, there is no such simple description of the natural history of drug abuse.

In order to cope with this complex variation, researchers have used three general methods: historical-comparative analysis of the broad institutional and societal variations that affect drug use; longitudinal assessment of individual "careers," to see how drug use and abuse vary across a life span; and ethnographic study, which includes relatively unobtrusive participant observation by trained analysts in the actual social environment of substance use and abuse, with analysis of the setting and world as seen and acted in by users. These forms of archival scholarship, life-span study of individuals, and ethnographic assessment of the subjective meanings and dynamics of drug practices are a necessary counterpart to the more familiar laboratory and survey methods generally identified with the behavioral and social sciences.

Successful natural history studies are sufficiently numerous to qualify it as an established research approach. It has been applied in studies of returned Vietnam veterans (Robins 1974); of former Lexington patients from rural Kentucky (O'Donnell 1969) and New York City (Vaillant

1973); and of drug-using communities in Baltimore (Nurco et al. 1975), New Haven (Gould et al. 1974), New York (Feldman 1968; Preble and Casey 1969), Chicago (Hughes 1977), Oakland (Sutter 1966; Blumer 1967), and San Diego (Gerstein 1976).

However, in contrast to the national and regional youth drug surveys to which NIDA has made stable, relatively long-term commitments (see Kandel (1980) for a comprehensive review), there has not been a systematic, sustained investment in natural history research. Specific fads such as PCP have received attention (Feldman et al. 1979), but this sporadic interest needs to be converted to a well-considered program of regular support for natural history studies. Such studies should be encouraged in a number of communities selected for geographic and socioeconomic diversity and for the availability of high-quality research establishments. Such projects will become scientific data resources for the next several generations of research on substance use, especially for research leading to new treatments.

As we discuss below, a missing component in most treatment-related studies has been good empirical information about the effects of past and projected intervention on the vast majority of people who are not in treatment but whose behavior is nonetheless of concern to policy makers. The studies we are recommending would provide invaluable information for evaluating long-term as well as immediate effects. For example, natural history information even on so widespread a practice as tobacco smoking is remarkably scarce and coarse. We wish to emphasize that, to be maximally useful, natural histories cannot be focused on single substances but must consider a broad range of substances.

As a first step toward getting a natural history program under way, it would be appropriate for NIDA to convene a group of scholars to discuss technical issues involved in such studies. It is important that a variety of disciplines and methods be represented, including anthropologists, ethnographers, other field-study specialists, epidemiologists, survey research experts, and perhaps others. A desirable result of such a meeting would be a set of model research strategies that could be replicated with appropriate tailoring in a number of communities. Finally, but of great importance, the natural history data so collected should, after reasonable amounts of time and within the limits imposed by protection of participants, be made available as public-use files to qualified scientific investigators.

Field Replication Studies

We have argued above that high-quality comparative data on natural histories of substance use comprise a basic scientific resource for understanding substance abuse and its treatment. Field replications, in contrast, are not a basic scientific resource. But they are a practical necessity if treatment innovations are to be chosen rationally for implementation, and subsequently taught to practitioners and evaluated by program managers, quality controllers, and policy planners. The relevant problem to date has not been a paucity of studies but rather a widespread failure of field replications to include appropriate scientific criteria

in their designs. In many cases there are good reasons why such studies have not lived up to strict scientific criteria (just as there are good reasons to obtain program information even though its scientific value is generally nil). But we believe that it is appropriate for NIDA, as the lead scientific agency, to explicitly publicize and apply rigorous scientific standards as conditions for its support of field replication research.

The main question to keep in mind is: Why "replicate" treatment at all? In the history of treatment for drug abuse, mandates to provide public treatment have often been handed down in the absence of clearly formulated goals or of the priorities of various goals. Ordinarily, a "drug problem" has been identified, a rough guess has been made of how many "treatment slots" need to be funded, and some program using the latest technology or some mix of techniques has been undertaken. But what is to be achieved? "Fill treatment slots?" "Reduce the community's drug problem?" "Give drug-affiliated people a chance to better themselves?" Presumably, all of these, and others as well, are reasonable goals. But there are considerable differences in how best to determine treatment effectiveness, depending on which goal is emphasized. Replication study must clearly indicate which goals it is trying to investigate and must designate measures appropriate to each goal.

If "filling the slots"--providing opportunities for care--is at issue, one would require only a management information system capable of accurately monitoring how many slots are actually filled how much of the time. (Such information is generally called a "process measure" in the evaluation research literature.) If, on the other hand, the treatment is expected to affect the prevalence of a problem in a community, such as to reduce its rates of crime, public intoxication, or lung disease, there is little point in trying to monitor the delivery of specialized services as such. How the program delivers treatment, or the group of people to whom the program is delivering, is not exclusively at issue. Rather, one wants to know what has happened to the rate of the problems specified in the community of concern as measured by aggregate problem rates or indicators.

For example, suppose that an alcoholism treatment program was instituted to reduce the number of daytime drunks on the streets of a downtown business district. Assume that enough treatment openings were created to equal the number of drunks on the streets, that all these slots were in fact continuously filled, and even that all the treated people stopped drinking not only during treatment but after "graduating." Is the program a success? One cannot know unless the number of drunks on the downtown streets before treatment is compared with the number on the streets after treatment. Despite all the assumptions above, there could be no change in the crucial index. If the program fills with people who were not downtown drunks, if the movement of downtown drunks into treatment makes room for uptown drunks to move downtown, or if the availability of treatment makes it more attractive to become a downtown drunk, the program, successful by any number of process measures, could fail to solve the specific problem for which it was intended. And even if virtually all the downtown drunks disappeared, one

could still not know, in a scientific sense, that the program was responsible. Without appropriate (untreated) controls, one could not rule out other factors that might have accounted for the disappearance of the drunks, such as mass police raids, a liquor strike, a new free nearby amusement center, etc.

The third type of goal, that a program is meant to "give people a chance to better themselves," is one that is generally addressed by replications and the one with respect to which deployment of control procedures requires the greatest care. At issue is how the people receiving an innovative treatment would have fared without that treatment. In other words, one needs a comparison by which the effect of treatment, using outcome measures, can be shown. Most of what ought to have been controlled field replication research, as reported in the literature, has been done without a null hypothesis and has simply reported the status of people in treatment after varying lengths of time. Such replications may have been politically vital, and the only practical method under some circumstances, but they do not fulfill the need for scientifically valid field testing.

The typical approach in research that has used a null hypothesis procedure is to use the people who are to receive treatment as their own controls. One measures the people on some scale before or upon entry to treatment and then on the same scale at some point after entry, and one ascribes any difference to the intervention of treatment. The null hypothesis in this approach is that, however the people started, so they would have remained without treatment. This approach is only a small improvement over uncontrolled designs because this null hypothesis is quite dubious. People generally come to treatment when things are going badly for them, and the little natural history that has been accumulated shows that drug users undergoing bad times very often improve, sooner or later, with or without treatment. There are also changes in psychosocial functioning that accompany maturation, with or without treatment. Therefore, one cannot conclude that people would have remained the same without treatment. Consequently we do not advocate using people as their own controls to demonstrate the effects of treatment.

The use of control groups has been developed to correct for the inadequacies of the approach noted above. The difficulty in the use of control groups lies in how to select or create an appropriate group as a control for the group receiving the sample drug treatment. One way is somehow to divide the treatment group itself by refusing or interrupting treatment for some members--in other words, to create early dropouts. Another way is to provide a placebo treatment condition for the control group, such as putting them on a waiting list or providing an ineffectual "holding" therapy. In general, however, in order to study a treatment innovation that is not yet proven to be either superior or inferior, the most logical choice for the control procedure is current therapeutic practice.

Whatever method of creating the control group is used, it must be done by random assignment, or something as close to this as possible; otherwise, one is measuring not the effect of treatment, but the treatment admission procedure, a complex unanalyzed set of judgments about who should go into which treatment. There are often practical and

ethical problems in the use of randomizing control procedures, and these must be resolved with the help of people experienced in the clinical process. Despite the problems, however, such a strong scientific case exists for random assignment that NIDA should ordinarily insist on it as a condition for supporting field replication research.

Two other basic features that NIDA ought to insist upon in replication studies are thorough initial data collection and follow-up of the total sample of applicants for treatment. A good standardized set of initial data helps in deciding whether specific subgroups of the study population account for differences in outcome measures between the treatments studied. Even in the absence of net differences, one may find that different treatments succeed or fail with different parts of the population. Since the point of all field research is to find ways of more effectively prescribing or distributing treatments, it is of greatest value to base studies on information that can be collected and acted on at initial contact or application for treatment.

We cannot stress too much the importance of insisting that all applicants for treatment, or at least a reasonable representation of all applicants, be included in the measurements of success or outcomes of treatment. This is not because it is important to compare early dropouts with nondropouts, but because, in comparing treatment groups, staying in or dropping out is part of the response to treatment.

It is important to include a strong ethnographic component in replication research. Ethnographic inquiry helps in understanding the unexpected and often difficult-to-measure qualitative features of human variation in treatment programs. It particularly enables one to gauge the ways in which different participants respond to treatment, ways that may escape the usual boundaries of age/sex/race/job classifications. While the standard treatment approach focuses on the relations between direct care-giver and recipient, it is clear that many complex relationships of treatment (between primary care staff and management) and of users (between patients, their families and friends, and the local community) have to be considered in determining the effectiveness of a treatment program. Researchers from the relevant branches of anthropology, sociology, social psychiatry, and social and community psychology can carry out such studies.

Although we have stressed the need for NIDA to enforce strict scientific criteria for field replication studies, it is essential that the studies be devised with specific reference to the nature of local communities and individual motivations and histories. While NIDA should leave the responsibility for specifying appropriate procedures to the researchers, it should insist that the procedures show reviewers that the issues defined here have been considered and that the conclusions that the replication studies are trying to reach can follow from the design features proposed.

A final consideration in field research is whether conclusions drawn from replication programs can be assumed to be applicable to people who are not in treatment in such programs. This question entails two issues. One is recruitment, which is widely ignored in discussions of treatment. Ordinarily, a program does recruit clients, rather than simply opening its doors and waiting for people to wander in. Some form

of advertising, outreach, or referral arrangement accompanies virtually every type of substance abuse treatment program. Recruitment tactics are very important because the history of drug treatment shows that expansion of availability of a new treatment is inevitably accompanied by alteration in the mode of recruitment. A program that begins with a few slots and a long waiting list makes admission decisions quite differently than a program (even the same program at a later time) with a large number of slots and not enough people to fill them. Unless there are specific data on the differences in treatment effects among populations recruited under different conditions, one cannot accurately predict the effect of program expansion. Our central recommendation here is simply that treatment replications must include a well-defined descriptive component regarding recruitment, i.e., how the sample population is brought into contact with the program.

The second issue involves quality control. As we have indicated, quality assurance is really a separate activity from replication research. However, the standards to which therapeutic practitioners are held can be no better than the quality of the descriptions of therapy features in the reports of replication projects. Criteria for diagnosis, prescription, care delivery, and other features tested in replication projects are the basis on which quality assurance standards are established. The more accurately that replication studies can define the nature of the care and the type of recipients for whom different treatment applications are appropriate, the greater will be the likelihood that implementation studies will identify high-quality programs. Policy planners will then have more realistic expectations about how and how many people may receive quality treatment and what effects should be expected.

Our recommendations regarding field replication studies can be briefly summarized as follows:

- Studies must designate measures appropriate to the goals of the treatment program in question. Process measures, aggregate indicators, and outcome measures are appropriate for different goals, and the logical relations between objectives and instruments must be made clear by the investigators.

- Null hypotheses in the use of outcome measures must be specified and appropriate control procedures used. Randomized control group procedures are ideal; where justifications for departing from them cannot be dismissed, appropriate changes in null hypotheses must be identified and justified.

- Thorough initial data collection must be required.

- Follow-up must be based on all those initially contracted or admitted to treatment, not just on those who completed or remained for long periods.

- All replication projects should include ethnographic components using suitably trained researchers.

- Replications must make provisions for clear, detailed, accurate descriptions of recruitment process, diagnostic criteria, and other features of therapy so that quality assurance standards and good profiles of subject populations can be constructed from them.

SUMMARY

In this brief report we have tried to explain the ways in which innovations in the treatment of drug-related problems come into being, to make clear the differences between activities that are often blurred, and to stimulate certain needed changes in the ways that the National Institute on Drug Abuse has been involved in research and development of drug abuse treatments.

We have specifically recommended strengthening of NIDA's basic research program on the natural histories of substance-using practices. This program must explicitly avoid compartmentalization of interest to one specific method or substance, since this has slowed down research advances in the past.

We have also recommended strengthening of NIDA's methodological standards in the specific area of field replication studies. The development of strong and uniform scientific criteria should never be made into a fetish in research: disorderly methods and broken rules precede major scientific advances often enough to caution against an overly formal or bureaucratic view of all research. However, in testing an innovation that is meant for clinical use by a variety of delivery units, a clear stage of rigorously controlled field trials is an absolute necessity.

REFERENCES

Blumer, H. (1967) The World of Youthful Drug Use. Berkeley, Calif.: School of Criminology, University of California.

Feldman, H.W. (1968) Ideological supports to becoming and remaining a herion addict. Journal of Health and Social Behavior 9:131-139.

Feldman, H.W., Agar, M.H., and Beschner, G.M. (1979) Angel Dust: An Ethnographic Study of PCP Users. Lexington, Mass.: D.C. Heath & Co.

Gerstein, D.R. (1976) The structure of heroin communities (in relation to methadone maintenance). American Journal of Drug and Alcohol Abuse 3(4):571-587.

Gould, L.W., Walker, A.L., Crane, L.E., and Lidz, C.W. (1974) Connections: Notes from the Heroin World. New Haven: Yale University Press.

Hughes, P.L. (1977) Behind the Wall of Respect: Community Experiments in Heroin Addiction Control. Chicago: University of Chicago Press.

Kandel, D.B. (1980) Drug and drinking behavior among youth. Pp. 235-285 in A. Inkeles, N.J. Smelser, and R.A. Turner, eds., Annual Review of Sociology, Volume 6. Palo Alto, Calif.: Annual Reviews.

Nurco, D.N., Lerner, M., Bonito, A.J., and Balter, M.B. (1975) An approach to the classification of the lifestyles of narcotics abusers. Pp. 77-94 in D.J. Lettieri, ed., Predicting Adolescent Drug Abuse: A Review of Issues, Methods and Correlates. Washington, D.C.: National Institute of Drug Abuse.

O'Donnell, J.A. (1969) Narcotics Addiction in Kentucky. Washington, D.C.: U.S. Government Printing Office.

Preble, E., and Casey, J.J. (1969) Taking care of business: the heroin user's life in the street. International Journal of the Addictions 4(1):1-24.

President's Commission on Mental Health (1978) Task Panel Reports. Volume IV Appendix. Washington, D.C.: U.S. Government Printing Office.

Robins, L.N. (1980) The natural history of drug abuse. Acta Psychiatrica Scandinavica 62, Suppl. 284:7-20.

Sutter, A. (1966) The Vietnam Drug User Returns. Final Report, Special Action Office Monograph, Series A., No. 2. Washington, D.C.: U.S. Government Printing Office.

Robins, L.N. (1980) The world of the righteous dope fiend. Issues in Criminology 2(2):177-222.

Vaillant, G. (1973) A 20-year follow-up of New York City narcotic addicts. Archives of General Psychiatry 29:237-241.