Treatment for Drug-Exposed Women and Children: Advances in Research Methodology

Editor:

Elizabeth R. Rahdert, Ph.D.

NIDA Research Monograph 165
1996

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Institutes of Health

National Institute on Drug Abuse
Division of Clinical and Services Research
5600 Fishers Lane
Rockville, MD 20857

ACKNOWLEDGMENT

This monograph is based on papers prepared by principal investigators and other staff members from the 20 projects of the Perinatal-20 Treatment Research Demonstration Program, which was funded by the National Institute on Drug Abuse in 1989 and 1990. The purpose of this volume is to share experiences, products, procedures, and implementation strategies that have relevance for established investigators as well as clinicians involved in developing innovative services in sites not yet experienced in research.

COPYRIGHT STATUS

National Institute on Drug Abuse
NIH Publication No. 96-3632
Printed 1996

NIDA Research Monographs are indexed in the Index Medicus. They are selectively included in the coverage of American Statistics Index, BioSciences Information Service, Chemical Abstracts, Current Contents, Psychological Abstracts, and Psychopharmacology Abstracts

Contents

IV. DATA MANAGEMENT AND ANALYSIS

Introduction to the Perinatal-20 Treatment Research Demonstration Program

Elizabeth R. Rahdert

During the 1980s sufficient evidence accumulated to suggest that the use of cocaine and other illegal drugs by pregnant women presented a major public health problem (Frank et al. 1988; Neerhof et al. 1989). Prenatal alcohol and other drug abuse was occurring in all racial and ethnic groups and across all social strata. For example, the results of a study conducted in 1989 in Pinellas County, Florida, found that positive toxicology for illegal drugs varied little between women admitted to public health clinics for prenatal care and women seen in several private obstetric practices (Chasnoff et al. 1990).

In response, the National Institute on Drug Abuse (NIDA) undertook the support of research demonstration grant projects that focused on the treatment of drug-abusing pregnant and postpartum women and their drug-exposed offspring. The intent of this program was twofold: conduct treatment research and, at the same time, create many new treatment slots for the women and their children. Toward these aims, NIDA funded 10 projects in September 1989 and another 10 in September 1990, the sum of which became known as the "Perinatal-20 Treatment Research Demonstration Program" or "Perinatal-20" for short (see appendix for key Perinatal-20 personnel).

Each project was designed to scientifically evaluate either a comprehensive treatment program composed of an integrated system of services or a specific targeted therapeutic intervention embedded in a comprehensive continuum of care. Each study targeted either the drug-abusing woman of childbearing age (predominantly pregnant or postpartum) in treatment *with* her children or the woman in treatment *without* her children.

As slightly different research questions were addressed by one or more of the Perinatal-20 studies, each could be described as experimental, quasi-experimental, or correlative by design and be represented by one of the following:

Research Question 1:	What is the differential effectiveness of two types of treatment?

Example: Determine the differential effectiveness of a hospital-based *residential* treatment program (type 1) and a hospital-based *outpatient* treatment program (type 2).

Design Feature: Random assignment to each condition.

Research Question 2: What is the differential effectiveness of treatment-as-usual and treatment-as-usual plus enhancement?

Example: Determine the differential effectiveness of treatment for mothers in a therapeutic community with residence provided for the women but not for their children (*treatment-as-usual*) and treatment for mothers in a therapeutic community with residence provided for the women *and* their children (*treatment-as-usual plus enhancement*).

Design Feature: Random assignment to each condition.

Research Question 3: To what extent is type 1 treatment effective and to what extent is type 2 treatment effective when determining treatment outcome by use of a common set of intake and outcome measures?

Example A: Determine the effectiveness of an outpatient treatment program specializing in *women's services* (type 1) and the effectiveness of an outpatient treatment program based on a traditional model of *services for both men and women* (type 2).

Design Features: Nonrandomized, case-controlled, group comparisons.

Example B: Determine the effectiveness of an *outpatient* treatment program for mothers (type 1) and the effectiveness of a *residential* treatment program for mothers (type 2).

Design Features: Nonrandomized, case-controlled, treatment process study, with assignment of each woman to the outpatient program or the residential program based on the results of the intake assessment and clinical judgment.

Given that treatment research in this area was somewhat uncharted and certainly underdeveloped at the time the Perinatal-20 studies were first designed, many investigators said they might benefit from sharing

problems and solutions associated with their efforts to establish or expand a clinical site, integrate a research project within that site's daily operations, and conduct a study on a smooth and continuous long-term basis. Pooling knowledge and experience had the potential for a greater and more rapid yield than might otherwise occur if the 20 projects did not communicate until all study results were published.

Precedence was set for a cooperative examination of a variety of challenging issues asserted to be inherent in this area of research. In July 1990 a technical review was held on methodological issues in epidemiological, prevention, and treatment research on drug-exposed women and their children that resulted in the publication of a NIDA research monograph (Kilbey and Asghar 1992). Several Perinatal-20 researchers contributed to this earlier volume because they had already encountered some of the difficulties associated with conducting research in this area.

This volume presents a sample of what has been learned since then about the challenging areas of (1) services research implementation; (2) subject selection, recruitment, and retention; (3) clinical assessment and program evaluation; and (4) data management and statistical analyses. It anticipates that shared experiences, products, and procedures used in one or more of the Perinatal-20 studies might be of benefit to researchers, practitioners, and program administrators. Furthermore, it is hoped that the methodologies presented in this monograph can be of special value to those who wish to establish new services in combination with study protocols in clinical sites that heretofore have never participated in research.

REFERENCES

Chasnoff, I.; Landress, H.J.; and Barrett, M.E. The prevalence of illicit drug or alcohol use during pregnancy and discrepancies in mandatory reporting in Pinellas County, Florida. *N Engl J Med* 322:1202-1206, 1990.

Frank, D.A.; Zuckerman, B.S.; Amaro, H.; Aboagye, K.; Bauchner, H.; Cabral, H.; Fried, L.; Hingson, R.; Kayne, H.; Levenson, S.M.; Parker, S.; Reece, H.; and Vinci, R. Cocaine use during pregnancy: Prevalence and correlates. *Pediatrics* 82:888-895, 1988.

Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children.* National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. 399 pp.

Neerhof, M.G.; MacGregor, S.N.; Retzky, S.S.; and Sullivan, T.P. Cocaine abuse during pregnancy: Peripartum prevalence and perinatal outcome. *Am J Obstet Gynecol* 161:633-638, 1989.

AUTHOR

Elizabeth R. Rahdert, Ph.D.
Research Psychologist
Treatment Research Branch
Division of Clinical and Services Research
National Institute on Drug Abuse
Parklawn Building, Room 10A-10
5600 Fishers Lane
Rockville, MD 20857

APPENDIX. *Key personnel: Perinatal-20 Treatment Research Demonstration Program projects*

Farrokh Alemi, Ph.D. (PI)
Richard Stephens, Ph.D. (Co-PI)
Cleveland State University
Cleveland, OH

Elizabeth Brown, M.D. (PI)
Boston University School of Medicine
Boston City Hospital
Boston, MA

Ira J. Chasnoff, M.D. (PI)
National Association for Perinatal Addiction, Research, and Education
Chicago, IL

Shirley Coletti (PI)
Patrick Hughes, M.D. (Co-PI)
Operation PAR, Inc.
PAR Village Therapeutic Community
St. Petersburg, FL

Louis Cooper, M.D. (PI)
[Robert S. Neuwirth, M.D.]
St. Luke's/Roosevelt Hospital Center
New York, NY

Ernest Drucker, Ph.D. (PI)
Daniel R. Neuspiel, M.D., M.P.H. (Co-PI)
Montifiore Medical Center
Bronx, NY

Tiffany Field, Ph.D. (PI)
Emmalee Bandstra, M.D. (Co-PI)
University of Miami School of Medicine
Miami, FL

James Fine, M.D. (PI)
[Karla Damus, Ph.D.]
[Elaine Blechman, Ph.D.]
State University of New York at Brooklyn
Brooklyn, NY

Judy Howard, M.D. (PI)
Leila Beckwith, Ph.D. (Co-PI)
University of California at Los Angeles School of Medicine
Los Angeles, CA

Karol A. Kaltenbach, Ph.D. (PI)
[Loretta Finnegan, M.D.]
Thomas Jefferson University
Jefferson Medical College
Philadelphia, PA

Robert A. Lewis, Ph.D. (PI)
Purdue University
{Terros, Inc., Phoenix, AZ}
Lafayette, IN

Paul R. Marques, Ph.D. (PI)
National Public Services Research Institute
Landover, MD

Ellen Mason, M.D. (PI)
[Anne Seiden, M.D.]
Cook County Hospital
Chicago, IL

Mary E. McCaul, Ph.D. (PI)
Johns Hopkins University School of Medicine
Baltimore, MD

Lawrence A. Palinkas, Ph.D.
[James A. Hall, Ph.D.]
University of California at San Diego School of Medicine
San Diego, CA

Sidney H. Schnoll, M.D., Ph.D. (PI)
Virginia Commonwealth University
Medical College of Virginia
Richmond, VA

Richard S. Schottenfeld, M.D. (PI)
APT Foundation, Inc.
{Yale University School of Medicine}
New Haven, CT

Kenneth D. Stark (PI)
Washington State Division of Alcohol and Substance Abuse
{King County Hospital, Seattle, WA}
Olympia, WA

Sally Stevens, Ph.D. (PI)
[Peggy Glider, Ph.D.]
[Rod Mullen]
Amity, Inc.
Tucson, AZ

Irma H. Strantz, Dr.P.H. (PI)
Robert Nishimoto, D.S.W. (Co-PI)
University of Southern California School of Social Work
{Watts Health Foundation Uhuru Family Center}
Los Angeles, CA

KEY: PI=current principal investigator; Co-PI=coprincipal investigator;
[ ]=principal investigator submitting original proposal other than current PI;
{ }=performance site other than applicant organization

Building Bridges: Treatment Research Partnerships in the Community

Ira J. Chasnoff, Paul R. Marques, Irma H. Strantz, James Farrow, and Shoni Davis

INTRODUCTION

Substance abuse during pregnancy continues to affect the lives of families across the United States. Most communities are not organized to deal with the multiple problems that must be addressed when developing intervention and treatment programs for pregnant women. At the same time, many citizens and community organizations may be suspicious of outside investigators stepping in to design outreach and treatment programs within existing structures and organizations. These difficulties are multiplied if the community regards the research as an opportunity to take advantage of the community without providing any long-term benefit to the community members.

There is little debate about the need for a wide-ranging assortment of services for indigent women who are involved with alcohol or other drugs. Rather, the debate revolves around how best to provide and assess comprehensive drug treatment services. The predominant issues of cost, convenience, and effectiveness inevitably have an effect on research. Centralized programs conducted at comprehensive health care facilities, such as a university medical center, provide a congenial environment wherein to conduct treatment services research but provide relatively little access to the populations at highest risk. By contrast, community-based providers with direct access to a sufficient number of clients have little or no experience in conducting or participating in clinical research projects and are less apt to buy into the values of the research culture or appreciate the need for objectivity and experimental control.

Nevertheless, overcoming the many difficulties associated with conducting research in a community setting is important—even essential—because the majority of service providers and consumers are found in the community. It is here that empirical evidence of program effectiveness must be found. For this reason, there are strong incentives to identify and overcome the obstacles to conducting a research project in the community milieu.

This chapter identifies some potential barriers to implementation of a community-based treatment research program. The issues described are derived from the experiences of four treatment research programs located

in four different parts of the country (i.e., Chicago, Illinois, Los Angeles, California, Landover, Maryland, and Seattle, Washington). Each project was established to develop, implement, and evaluate the effectiveness of a substance abuse treatment program for pregnant and postpartum women within a community setting in conjunction with a variety of services provided through State and local agencies.

PROGRAM IMPLEMENTATION

The NIMBY (Not In My Back Yard) syndrome is not unique to any one social, ethnic, racial, or economic group. In spite of the devastation wreaked on some communities by substance abuse, many neighborhoods are reluctant to allow drug treatment programs into the community. Although residents often view the substance abuser as a threat to community safety and as a harbinger of the deterioration of local values, the promise of the availability of treatment facilities does not alleviate these fears, and many community-based programs face an uphill battle in trying to find an appropriate location for service delivery.

The Chicago program, operated by the National Association for Perinatal Addiction Research and Education (NAPARE), is a community partnership that includes treatment providers, health care agencies, and community hospitals in an inner-city neighborhood on the near west side of the city. This community has some of the highest rates of substance abuse, murder, infant and child mortality, and violence in the metropolitan area; drug deals are openly consummated on most street corners. However, many residents banded together to protest the opening of a treatment program in the neighborhood and worked to block changes in zoning restrictions that would allow the facility to operate. It took several months of attending community meetings and working with local leaders even to open the doors of the building.

Gangs also threatened the existence of the program. In a related program developed to support the original treatment research project, NAPARE opened a Head Start program within a Chicago Housing Authority site. Unknown to the staff at the time of selecting the site, the housing unit sat on the border of two gangs' territories. While children were in class, gangs attacked the program classroom with gunfire, shot out windows, and broke into the classroom and threatened the children. These repeated attacks resulted in three shutdowns within a week's time. Negotiations with the gangs to declare the area a "gang-free zone" took up more precious time before the program could be implemented fully. These delays caused the late startup of recruitment and service provision and affected the number of patients recruited into the study in the first year of funding.

BUDGET CONSTRAINTS

Startup problems are the bane of treatment research programs, which typically have a lifespan of 3 to 5 years. Therefore, the design and management of a budget for treatment research often demand more care and scrutiny than a budget for service provision only. Reallocating funds from staff positions to facility or operating costs can have a critical effect on the number of research participants who can be enrolled in the study and exposed to the intervention. If the overall number of participants in the final database is too small, apparent changes occurring in certain study variables may fail to achieve statistical significance. Also, if the number of participants involved concurrently with the program is too small, with artificially low counselor caseloads and peer group memberships, any benefits derived from the treatment milieu may be impossible to replicate in other treatment settings.

A treatment research program's budget can be affected by the seemingly simple task of ensuring the safety of staff members and program participants. A treatment research project is almost always a large investment in terms of carefully selected and trained staff members, and rapid turnover is an anathema. Therefore, to secure staff members for a long period, obtaining an attractive and safe facility is a major priority. In addition, research participants, particularly mothers with infants and toddlers, are not likely to participate in a treatment program where the setting is dangerous. In recent months, new concerns for the safety of research participants and staff members in inner-city research projects have been caused by escalating violence in the communities. The researcher must deal with this new reality and include safety precautions (e.g., bulletproof windows, iron fences and gates, security guards, alarm systems) in the budget.

In the original plan for the Los Angeles program, the community-based agency was to house both the experimental and control groups at two leased sites in south-central Los Angeles. However, after the grant was funded, it was necessary to modify the budget to provide for more costly leased space, especially for the experimental group, and an armed security guard at each site. Because of these budget changes, counselor positions were lost, which greatly reduced the ongoing treatment capacity at each site. However, it was concluded that research participants could be enrolled and tracked, in both the experimental and control groups, in sufficient numbers to provide for outcome comparisons that could reach statistical significance. Also, although the number of treatment slots at each site was reduced, the low counselor-to-client ratios and small therapy group sizes in the enhanced intervention program were preserved.

(Nevertheless, it was decided to seek supplementary funding to restore lost treatment capacity.)

Finally, as the demonstration treatment research grant nears its end, a transition for financial support for the treatment services must be made to State and local sources. This requires a realistic look at the budget with close attention to obtaining funding from a variety of agencies. The community will have come to expect the wide range of services made available by the large infusion of Federal funds that supported the treatment research, but few States or communities will be able to support that level of commitment over the long term. If the program ends at the conclusion of Federal funding, the community will have an even greater reluctance to welcoming the next treatment research program.

IMPLEMENTING THE RESEARCH PROTOCOL

When a study protocol is to be made part of a service delivery program at an existing community agency, its introduction can be facilitated if it is not viewed as a special case or as an activity of no lasting value to the organization's mission. For any new treatment program to be effective, the management and clinical staff members must understand and endorse its goals and philosophy and feel comfortable with their respective roles. Also, when a change in program content is being considered, there must be opportunities for staff members to provide input on the change to encourage a feeling of ownership and commitment. Therefore, all components of the program (e.g., relapse prevention, psychosocial and parenting education, mother-infant bonding activities), as well as the research instruments and schedule, should be developed or reviewed and accepted by a committee that includes members of the agency's drug treatment clinical and management staffs. By involving key drug treatment management and clinical staff members in each step of this process, some stereotypical, negative attitudes toward the research participants and the rich program of services and activities planned for them can be neutralized.

From the outset, the Los Angeles researchers felt that there must be an underlying consistency with the existing community treatment system so that staff members in the new program would not be placed under great stress in the work environment. The community agency involved in the research project had used three of the components of the experimental treatment program (i.e., Rawson's relapse prevention model [Rawson et al. 1990] for cocaine abusers, parent education, 12-step Narcotics Anonymous groups) at two other treatment sites. However, staff members with the most experience with and understanding of these approaches were

not available to help implement the new program because they had been transferred to other positions. In every human services system, experienced clinical staff members are routinely transferred or promoted in support of organization priorities. Therefore, it was determined that new personnel would be recruited and trained specifically for the day treatment program.

INFORMATION OWNERSHIP

Community-based treatment programs operate under a different set of priorities than those for university-based programs, especially with regard to the acquisition and use of information. From a research perspective, data are collected over a long period with only occasional delving into the database to evaluate progress and assess preliminary information. Treatment program personnel often want a more rapid turnaround of information, which may result in overinterpretation or premature use of data. In addition, if a research team uncovers information that could have implications for the treatment approach for a specific research participant, the question arises whether the researchers are responsible for reporting that information to the woman's therapist. This is an issue that frequently arose during the intensive individual interviews that were conducted as part of the research assessment of the women in the Chicago program. Clear guidelines for communication between the research and clinical teams were explained at the beginning of the project, and the working relationship between the two teams was under constant scrutiny throughout the term of the project.

TREATMENT SITE STAFFING

There are many benefits in collaborating with an established agency for the treatment intervention aspect of the study. In the Los Angeles study, as provided for by the grant, the University of Southern California contracted with a community-based agency to provide the specified drug treatment services to research participants. The agency, which was established soon after the 1965 Watts riot, provides a range of medical, public health, and substance abuse services to residents of south-central Los Angeles. At least three benefits resulted from the involvement of this agency in the collaborative treatment research effort.

1. Among key agency staff members, there existed a body of knowledge and experience in the delivery of drug treatment services, including those that target pregnant and parenting women; thus, program implementation was not a trial-and-error activity.

2. Through the community-based agency, a variety of medical, public health, and mental health services were made readily accessible to the research participants and their infants and families.

3. African-American community leaders and agency staff members were involved in treatment research in a way that was positive and brought new services to the community. Furthermore, the agency was known and accepted by the research participants, which avoided any potential biases related to client refusal to participate because of suspicion regarding the motives of the treatment provider.

However, established agencies have standardized job descriptions, salary schedules, hiring policies, and procedures, and various management reviews and approvals must be obtained before any personnel action can be taken. On the one hand, when many systems are already in place, it is easier to implement a new program in collaboration with an established agency. On the other hand, if the research program has personnel requirements that differ from the agency standard, much time can be lost in the review and approval process. For example, it often was found by all four research teams that although the counselor positions in the research protocol had duties and caseload standards different from those of the regular positions, it seemed easier to use existing personnel than hire new employees.

A frequent difficulty in staffing treatment research programs is retention of staff. It is well known that supervisory and line staff workers are seriously underpaid in most publicly funded drug treatment programs. Even if a research grant budget can support higher salaries for supervisory and line staff workers at research sites, the community agency's salary schedule for drug treatment personnel usually cannot. Therefore, new staff members, hired and trained specifically for the research treatment program, are soon receiving job offers from agencies that value the expertise they have gained from working at the research site and can afford to pay higher salaries.

Skilled counselors, regardless of ethnic or demographic mix, are sometimes difficult to find, especially those willing to work for the wages available in the communities served by the research treatment program. If this difficulty can be anticipated and there is adequate lead time to set up the program, it is wise to fully evaluate the counseling skills and backgrounds of direct services providers. Certification standards are not rigorous for drug treatment services providers, and even fully certified providers may not have experience with women on welfare and their particular problems. For example, without significant training and preparation, people with experience in working in a middle-class 12-step program cannot be transplanted usefully to a milieu of urban poverty.

COLLABORATION IN SUBJECT REFERRALS

The university-based researcher who engages in a study that involves one or more independent community organizations must understand and accept the fact that, even with general oversight responsibilities for the study, he or she is perceived as an "outsider" by each organization's staff. Therefore, the researcher needs to have "friends in high places," such as one or more key managers of the agency who support the research goals and will keep the researcher apprised of any organizational changes that might affect the research. The importance of this connection was demonstrated vividly in the Los Angeles study, where researchers depended on a collaborating agency for client referrals; they unexpectedly lost key manager support at the agency together with the promised flow of referrals.

Unlike research in which investigators recruit research participants from a large pool of their service recipients, the women in the Los Angeles drug treatment study were drawn from the child protective services agency's caseloads. After being identified and reported to the agency at the time of delivery as probable drug abusers because of certain behavioral and physical indicators that they and their newborns exhibited, including a positive toxicology screen, the mothers were assessed further by agency workers. It was first determined whether there was endangerment to the infant if he or she was not placed with a relative or in foster care and then whether the mother would agree to complete a drug treatment program. If treatment was indicated, the worker referred the mother to this treatment research project or to another treatment program.

Soon after being notified about the funding of this research proposal, the investigators in Los Angeles learned that there had been certain organizational changes in the child protective services agency. The clinical policy and management staff members who had helped design the program had been transferred to other positions within the agency, and their firsthand knowledge of study objectives and special commitment to the project were lost. The agency also hired a new director who reorganized the agency and established new regions, with the result that many new teams were handling the south-central Los Angeles target area.

Although hospital reporting of drug-exposed infants remained constant during the period following the implementation of the Los Angeles project and a steady flow of referrals was expected from the agency, by the end of the sixth month of operation, the number of referrals was falling far short of expectations; only one-sixth the projected number of research participants were referred by child protective services workers. Thus, to increase referrals, there was an urgent need for researchers to establish new relationships with the management and line staffs at the agency.

A program of outreach and inservice orientation was developed and delivered by the researchers and program staff personnel to workers in the new regional offices.

When initiating, developing, and expanding a referral network from existing public services, it is not always clear which egos need to be stroked or whose indulgences need to be sought to gain full cooperation. In the Landover program, despite having acquired all the relevant letters of assurance from the hospital's chief executive officer and medical department chairpersons, researchers found that the hospital operations officer had the authority and inclination to prevent the smooth implementation of a referral system by initially allowing staff nurses to do no more than pass out brochures to prospective clients (mothers who tested postive for cocaine at delivery). Under this peculiar policy, the mother would have to initiate a telephone call to project personnel. This was an impossible way to form a sample, but it was an easy barrier to erect and to justify in the interest of patient confidentiality. This clash seemed related more to a turf problem or concerns about hospital image than to protection of confidentiality, but the motive never became completely clear. One neonatologist explained that any publicity that associated the hospital with area drug problems would have made it more difficult for operations personnel to implement their plans to build a new thoracic surgery center. On the other hand, the hospital had an ethical responsibility (and legal incentive) to protect patient confidentiality, and the operations department may have been genuinely motivated by such concerns. If this was the true motive, this particular type of barrier may have represented a case where assiduous risk control led to *greater* risk for mothers and their children (preventing access to remedial services) than would have occurred from a more measured degree of protection.

This barrier was eventually circumvented by repeated visits with direct services providers in the hospital. Caregivers (social workers, midwives, nurses), who have a more personal understanding of this population and are the key personnel who can make or break a referral process, must be convinced of the value of a research program. Administrative and operations personnel have to be involved in setting up the program, but their agendas should not be allowed to become barriers to the project goals. It is also important to have multiple referral sources so that difficulty with any one source does not have the potential to shut down a project.

RANDOM ASSIGNMENT AND COMPARISON GROUPS

Caregivers are often hesitant to make referrals to treatment research programs that plan to randomly assign some referrals to comparison groups,

so this topic should be approached with care. In the Landover project, it did not matter that the women in the comparison group were informed explicitly that they were free to take advantage of any public remediation services that would ordinarily be available to them independent of the research program. A great deal of "sales and marketing" by researchers was required to overcome resistance from referring providers. This involved chatting, informing, and yielding when necessary. Some tactics included hosting an open house for caregivers, soliciting their advice about gifts to be included in "goodies" bags provided to women (e.g., diapers, thermometers, coupons, etc.), providing inservice presentations, and sending periodic newsletters to keep the referral community informed.

In addition, any evidence of the investigator's humanity that can be shown while engaging interest in and attention to the research project can help. The investigator should make it clear that even the comparison group of women will be better off in some way for having been involved in the project. In Landover the investigators conceded the need to provide case management services (but not direct treatment) to the comparison group as an incentive for cooperation, and this concession helped overcome the reluctance of caregivers to make referrals.

Comparison groups should be studied carefully to log the extent of their involvement with nonproject-related supportive or treatment services and with outside resource people who play a role during the project period. This is always important in treatment research but may be particularly important with indigent mothers whose initial needs are often more basic than their need for drug treatment. Mothers often benefit as much from help with the practical aspects of living as from targeted drug treatment. In many cases, treatment may be premature because a sizable proportion of mothers will have little incentive to value treatment. That is, for many (or most) there are no expectations of a better life, no employer threats, no spousal threats, no legal threats, and surprisingly, only a rare possibility of action by child protective services (likely because of system overload [Ashery 1992, pp. 383-394]). As a consequence, relatively few women who go to "drug treatment" are ready for treatment, and therefore, comparison groups that receive case management services have a good chance to do almost as well as the treatment groups. Placing more attention on measuring nontreatment support activities helps identify multiple other factors that may account for outcome differences between groups.

Finally, alcohol- and other drug-free comparison groups should be screened carefully for substance use and abuse. In the Chicago program, 19 percent of the pregnant women who volunteered to serve as drug-free controls for the study subsequently had a positive urinalysis for an illegal drug and had to be removed from the comparison group.

PROVIDING TRANSPORTATION AND CLINIC-BASED CHILD CARE

No project aspiring to conduct outpatient treatment research on indigent drug-involved mothers can avoid contending with the double dilemma of providing transportation and clinic-based child care. A large institution with a car or van pool or a child care facility is apt to have less trouble than a community-based startup facility. However, for any research project—regardless of equipment or resources—getting mothers and children to leave home and contend with rigid schedules and urban transportation delays is difficult. However, because structure and punctuality are important parts of drug treatment, these barriers must be overcome. In Chicago and Landover, a portable cellular telephone used for making one or more calls in transit (to women who had telephones) improved the likelihood of participants being ready when the van arrived.

Most treatment programs develop a transportation plan by eliminating unacceptable and unworkable alternatives. At first, in the spirit of making use of existing resources, it frequently seems sensible to work with either public transportation or an existing business—such as a ride service or taxicab company with a fleet of cars and drivers—to deliver groups of mothers to treatment services. However, it often becomes clear that public transportation is too inflexible and is a clear disincentive for mothers whose ability to plan and follow through is not great. The cost of working through private businesses that provide ride services is usually unacceptably high. In addition, neither public transportation nor private ride services provide the sort of personal contact opportunity with the clients that a project employee or driver can deliver. A driver has the opportunity to be a surrogate friend or social worker and often can be called on to help a mother with minor chores. The Chicago and Landover research projects found that employing their own drivers was a better solution.

In terms of using a van, the most economical method was leasing a van for the duration of the project. For the Chicago and Landover projects, leasing costs for a passenger van appeared reasonable, but no lending institution would accept the risk of being the owner of record of the van unless it was protected by a $1- to $2-million insurance liability policy because, in the event of an accident, a passenger might choose to file a claim against the owner. Although understandable from a financial perspective, insurance-motivated self-protection can complicate the establishment of community programs. The least costly alternative for the programs was to purchase a used van. Insurance was still a problem, and innovative approaches to acquiring insurance were used. The Landover project was able to acquire an insurance policy only from the State-assigned risk pool (even though

there had been no claim against any vehicle insured by the program). Annual insurance costs from that pool have been approximately $4,000. This is approximately half the cost of insurance had the program leased a van but nearly five times greater than a regular policy without the unusual passenger liability requirements.

Providing child care to study participants presents a different type of problem. As seen across the Perinatal-20 projects, mothers had from one to eight children, with an average of three. Although many more programs have become available, only a few treatment programs are geared to accept children while a mother is receiving outpatient services. In addition to routine child care requirements, seasonal difficulties must be anticipated and planned for. One child care resource is generally adequate for pre-school-age children, but summertime brings new burdens for keeping groups intact and caring for older school-age children. There are no easy solutions, but communities may wish to consider coordinating child care services with local parks and recreation departments and libraries.

To develop onsite child care services, a program must consider the serious issue of licensing the child care facility. Most treatment programs do not have the space required, usually the first floor, for a licensed child care area. Sunlight, outdoor play area, fire regulations, and evacuation routes are all issues that arise in most State licensing plans. In some States, if the mother is in the same building, licensing is not required, but the program faces significant liability risks if the mother leaves the building.

DELIVERING HOME-BASED SERVICES

Clinical case managers and research personnel associated with the treatment program may be required to visit some of the most dangerous and violent neighborhoods in any urban area. Personal safety is thus an issue. In setting up the Landover project, the senior case manager and project coordinator solicited advice from social workers and local police who were familiar with the area. The advice and experience they received generally served them well and were useful for other high-risk communities.

The first home visit involves the most uncertainty, and case managers go out in pairs on the first visit. Ordinarily, the case manager has spoken only on the telephone with a woman prior to the visit. But for any home visit, the primary rule is always, “Don’t take risks, and don’t be too conspicuous.” From this rule, the case managers have derived secondary rules. Most rules have been developed from experience in home visits and adopted to minimize the sense of danger. Some of these rules are:

- The streets are safest before noon.
- Be more cautious when there is no evidence of care given to the outside home environment.
- Back off when a prospective client is hostile or uninterested.
- Back off when there are many other people in an apartment; schedule another time or meet in a neutral place.
- Be cautious in hallways where people are loitering or when urine smells fill the air.
- Be more cautious approaching apartments than single-family homes.
- Do not wear jewelry.
- Keep a purse out of sight.

The Landover program had three robberies in 2 years, all of which occurred when the case managers became comfortable with the environments they were working in and relaxed their usual caution. Two of these incidents involved theft of money from purses (one with a threat of violence), and one was a forced entry and theft of a cellular telephone absentmindedly left in plain view in a van.

PROGRAM EVALUATION AND SUBJECT ASSESSMENT

When engaging in complex intervention research, particularly when several independent agencies are involved, it is helpful to conduct a process or formative evaluation and share the findings and recommendations with the collaborators. There should be a response to each finding by the researchers. After the first 6 months of the Los Angeles study, the investigators conducted a formative evaluation to determine whether the treatment program and research protocol had been implemented as intended and what system problems were being encountered by the collaborating agency staff members.

Feedback from child protective services workers helped explain the low number of referrals of women into the program at that point; the reasons fell into four major categories.

1. There was a lack of knowledge among certain agency divisions regarding the services available through the project.

2. There was disappointment that onsite child day care for infants and siblings was not available.

3. Because a new departmental policy permitted each client a choice of agency for mandated drug treatment, potential research participants were selecting other programs that were less demanding of their time and effort (i.e., once-a-week counseling).

4. Some workers in the high-risk agency divisions disliked the research project's random assignment process.

Child protective services staff members in the last category felt that they should be the ones to make the assignment to the day treatment program (experimental group) or regular outpatient services (control group) on the basis of their clinical assessment of client needs.

In response to the formative evaluation findings, the researchers and the collaborating agencies took the following steps:

- Outreach efforts were expanded to all child protective services workers in the key district offices.

- The drug treatment agency's parent organization applied for and received State funding to establish a child day care program in the area, with priority access for children of research participants.

- Management staff members in the district offices took a strong proactive stance regarding the research project and its comprehensive treatment services, whether provided in the experimental or control group setting.

It is imperative that researchers maintain an effective communication network with each collaborating agency, using both formal and informal means. Through periodic meetings, progress reports can be shared and system problems addressed. Through informal contacts, small problems and complaints can be handled. Because it was learned, through informal channels, that three of the key Los Angeles child protective services offices were being used as training sites for new workers, the researchers accepted the fact that the regular provision of outreach and educational activities in these offices would be needed indefinitely.

Finally, a researcher may need to call on top leadership in the collaborating agency to focus sufficient attention and effort on a major problem, such as casefinding. This may be difficult when a large bureaucracy is involved and the treatment program effort is the first joint venture in community research; it is easier when there is a history of mutually beneficial

collaborative efforts. In the Los Angeles project, the resistance of certain child protective services workers to referring their clients to a project with random treatment assignments could not be handled solely through education and marketing. The agency's director was invited to visit the treatment research site and was impressed with the drug treatment programs provided by the community-based agency. Because of concern about the potential loss of both important research and services to the community, the director, with the assistance of the administrative staff, became personally involved in setting monthly referral goals and raising caseworker awareness of the benefits of the program to their clients.

CONCLUSIONS

A critical need in collaborative studies is strong centralized oversight of all aspects of the project. Although this type of management is time consuming, it is too important for the principal investigator to delegate the responsibility to someone else. When the researcher is an outsider, the importance of tact, persistence, and the ability to establish and maintain good interpersonal relations cannot be overstated. Close monitoring of the referral and intake systems and the treatment intervention process may not appear to be related to research supervision, but the integrity and continuity of the program as designed and the validity of the outcome data depend on this type of ongoing review.

The extra effort is worth it. Changing drug use behavior in the community where a woman lives is theoretically more effective for individuals and logically more defensible for communities. Studies in both animal models (Siegel 1983, pp. 207-246) and humans (Robins 1974) have shown that the environment where dependent drug use develops has strong eliciting power over drug-seeking behavior. Plucking a dependent user out of that environment and providing treatment does little to permanently break the associative links that control the craving for drugs. Building relapse prevention skills within the community context is important. In addition, if everyone needing drug treatment services were removed from the community, then the community would not benefit from a climate of change and renewal. Conceivably, a drug-involved community may reinforce drug-using behaviors for others in that milieu. Logically, as one wave of drug-dependent people moves out, the next wave would be ready to move in. Thus, despite the difficulties of making change in the context of the community, there are ample incentives to overcoming barriers to treatment research and establishing strong service links among providers.

REFERENCES

Ashery, R.S. Case management community advocacy for substance abuse clients. In: Ashery, R.S., ed. *Progress and Issues in Case Management.* National Institute on Drug Abuse Research Monograph 127. DHHS Pub. No. (ADM)92-1946. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

Rawson, R.A.; Obert, J.L.; McCann, M.J.; Smith, D.P.; and Ling, W. Neurobehavioral treatment for cocaine dependency. *J Psychoactive Drugs* 22(2):159-171, 1990.

Robins, L.N. *The Vietnam Drug User Returns: Final Report, September 1973.* Special Action Office for Drug Abuse Prevention. Contract No. HSM-42-72-75. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1974. 169 pp.

Siegel, S. Classical conditioning, drug tolerance, and drug dependence. In: Smart, R.G.; Glaser, F.B.; Israel, Y.; Kalant, H.; Popham, R.E.; and Schmidt, W., eds. *Research Advances in Alcohol and Drug Problems, Vol. 7.* New York: Plenum Press, 1983.

AUTHORS

Ira J. Chasnoff, M.D.
President and Medical Director
National Association for Perinatal Addiction, Research, and Education
Professor of Clinical Pediatrics
University of Illinois
200 North Michigan Avenue, Suite 300
Chicago, IL 60611
(312) 541-1272 (Tel)
(312) 541-1271 (Fax)
7115.1134@compuserv.com (Internet)

Paul R. Marques, Ph.D.
Director
Drug Research Programs
National Public Services Research Institute
8201 Corporate Drive, Suite 220
Landover, MD 20785
(301) 731-9891, ext. 102 (Tel)
(301) 731-6649 (Fax)
marques@pire.org (Internet)

Irma H. Strantz, Dr.P.H.
Associate Professor of Research
University of Southern California School of Social Work
University Park Campus
Los Angeles, CA 90089-0411
(213) 740-3469 (Tel)
(213) 740-9832 (Fax)
strantz@bcf.usc.edu (Internet)

James Farrow, M.D.
Associate Professor
Medicine and Pediatrics
Director
Division of Adolescent Medicine
University of Washington
CDMRC Building
1959 Northeast Pacific Avenue
Seattle, WA 98195
(206) 685-1273 (Tel)
(206) 543-5771 (Fax)
jaf@u.washington.edu (Internet)

Shoni Davis, D.N.Sc., R.N.
Executive Director
Perinatal Treatment Services
Affiliate Assistant Professor
Department of Psychosocial Nursing
University of Washington
CDMRC Building
1959 Northeast Pacific Avenue
Seattle, WA 98195
(206) 223-1300 (Tel)
(206) 223-1279 (Fax)

Conducting a Treatment Research Project in a Medical Center-Based Program for Chemically Dependent Pregnant Women

Ellen Mason

INTRODUCTION

In many respects a public medical center is an ideal location in which to develop and evaluate the efficacy of a comprehensive drug treatment program for pregnant women. Public hospitals have access to relatively large numbers of women whose medical and obstetrical problems may be a direct or partial result of the use or abuse of alcohol and other drugs. In addition, women residing in the community who are in need of chemical dependence treatment services are frequently identified and referred to a particular hospital by county board of health clinics and community-based health centers. Pregnant women who receive care at public hospitals often lack the financial resources and private medical insurance that are necessary to access drug abuse treatment in the private sector. Their only treatment option may be a program offered by a public hospital. Thus, low-income pregnant women served by a public hospital are a "captive" group for a newly launched, hospital-based addiction research demonstration project that needs to attract many such women over a short time.

In addition to possessing a large, accessible treatment (and therefore subject) population, a public medical center usually offers a broad spectrum of relevant primary through tertiary health care services, including high-risk prenatal care, neonatal intensive care, and specialized pediatric followup care. Often, the medical center is home to a variety of programs and professionals dedicated to clinical research in relevant areas, such as infant and child development, perinatology, psychiatry, and public health. These programs and people, although frequently understaffed and busy, will make themselves available to a treatment program through interdepartmental linkages, allowing for the formation of a comprehensive perinatal addiction treatment service without having to start entirely from scratch and without relying exclusively on external funding support for all program elements.

In spite of possessing this vital mix of ingredients—accessible subjects, skilled health care providers, and a full range of medical services—a medical center-based research demonstration program that serves chemically dependent pregnant women may encounter numerous

difficulties in implementing and running a treatment program and gathering data about the participants and the treatment. Overall, the problems that emerge are related to one of four domains:

1. The traditional structure of health service delivery in hospitals

2. Characteristics of the patient population and the intrinsic nature of addiction-related behaviors

3. Unforeseen difficulties with the research design and with mixing the research and clinical operations

4. Negative perceptions by hospital staff members and administrators of the problem of drug abuse and drug-abusing patients (subjects)

PROBLEMS ASSOCIATED WITH TRADITIONALLY STRUCTURED HOSPITAL SERVICE DELIVERY

Despite recent advances in the development of multidisciplinary care services for specific populations, such as oncology, human immunodeficiency virus (HIV)-positive, and high-risk prenatal patients, services in large hospitals are generally fragmented. Medical and psychosocial services are rarely delivered in a coordinated way. The traditional hospital care system has limited ability to deliver patient-centered care. Fragmentation of care within an institution is commonplace and constitutes a particular problem for an addiction treatment program that aims to provide a "one-stop shopping" service. For example, a common practical problem is the inability of a treatment program to obtain followup on what happens when a patient consults with a provider outside the treatment program. A patient may go to the hospital's oral surgery, orthopedic, or other subspecialty clinic, but no written or verbal report of the visit may be available for the chemical dependence treatment staff. Ironically, although the visit occurred in the same institution, it may be nearly impossible for the treatment program to obtain documentation of what medication was prescribed, what studies were ordered, and what followup care was arranged.

Hospital care is often fragmented geographically as well. For example, radiology, laboratory, outpatient, and medical records services are usually located far from each other and far from the treatment program. This separation works against the basic goal of comprehensive service delivery, which is to keep the services close to each other to facilitate patient compliance with treatment program elements and minimize patient frustration. The standard hospital and clinic layout requires frequent

waiting in lines, direction and redirection to various parts of the hospital, and a lot of walking. Negotiating the standard configuration of hospitals, especially large ones, is difficult even for able-bodied, motivated patients. In the case of the chemically dependent pregnant or postpartum woman who is new to treatment or in early recovery, the fatigue and frustration engendered by going from department to department for blood tests, ultrasound, appointment scheduling, and so forth may negatively affect her program participation and compliance with medical care, even when the health services appear to be easily accessible. Emphasis on case management and advocacy is needed to enable such a patient to successfully navigate "the system" both physically and mentally, although at times the additional help given her may foster an undesirable sense of dependence or intensify her internal feeling of chaos and nihilism.

Exacerbating this problem of scattered services is the fact that neither private nor public institutions are likely to assign conveniently located or adequate space to a new chemical dependence program. Given that most hospitals and clinics generally do not have much unused space to begin with and that chemical dependence treatment may be a low priority in a tertiary care hospital, a new treatment program may be located far from the main hospital and clinics, in space that especially mitigates against the integration of program and hospital services. Reallocation and renovation of the space for a treatment program is an extremely time-consuming process that may require extensive negotiations with hospital administrators, hospital labor union representatives, and municipal building inspectors. The process of locating and renovating space must be done with caution and tact. Serious animosity against the new program can arise if established providers of hospital services or other preexisting programs are displaced or inconvenienced by the new program.

Hospital departments and their service providers may be territorial about job descriptions and space. Institutional "turf" issues may arise when a chemical dependence treatment program appears to be encroaching on another department's area of expertise. Allied health service disciplines in the hospital, such as health education, occupational therapy, or social services, may refuse to allow treatment program staff members to be cross-trained or credentialed in functions that the other disciplines consider their domain. Moreover, these other disciplines may not be willing to accommodate the unique needs of a chemical dependence program. For example, HIV counseling generally is given prior to HIV serologic testing. In some institutions, this counseling may be available only through a specialized health education service. Despite the fact that each pregnant woman enrolled in chemical dependence treatment needs one or more HIV serologic screens during her pregnancy, scheduling these tests and obtaining the test results, along with the mandated pretest and posttest

counseling, from designated hospital personnel may be discouragingly labor intensive. At the same time, treatment program staff members, although already trained in this counseling, may not be allowed to deliver it.

Departmental directors may object to using personnel in ways that are new or innovative. For example, the use of certified nurse midwives in the addiction program's prenatal clinic may violate a nursing protocol that states that these midlevel providers are to care only for "low-risk" prenatal patients. Changing a nursing protocol, with the goal of providing superior services, may prove to be an almost insurmountable task, requiring many hours of negotiation.

An unwieldy hospital bureaucracy may negatively affect program implementation and functioning. Timely hiring of specialized clinical and research personnel, such as case managers and research associates, is difficult if the hospital personnel department is not flexible about position classification and job descriptions. Purchasing the materials needed for treatment or research purposes may prove to be a time-consuming and frustrating problem. If the items needed are not standard hospital supplies and if they cannot be acquired directly by the program but must be ordered through standard hospital purchasing procedures, the research timeline may be seriously delayed.

Confidentiality

Maintaining confidentiality in a medical center-based research demonstration program can be difficult. Federal law and ethical standards demand that patients in chemical dependence treatment not be openly labeled as chemically dependent. However, in the process of publicizing the program throughout the medical center to facilitate referrals and increase access to treatment for potential program participants, it becomes likely, from the time of the first patient contact, that the women entering the program will be labeled and will have their status known by the hospital and clinic staff members who work with or refer patients to treatment program providers. Further compromise of patient confidentiality occurs when the program refers its patients to other hospital services by using requisitions bearing the program name. Such requisitions are necessary for medical records purposes and for the program to retrieve test and consultation results through the hospital's internal communications system. For example, when a patient in a hospital-based addiction treatment program presents her stamped laboratory requisition to a registration clerk at the laboratory, the clerk who sees the program name may subsequently call out to the phlebotomist or laboratory technician that a drug abuse patient is waiting for service or behave in some

derogatory fashion prompted by negative feelings about drug use or drug addiction treatment.

Loss of confidentiality in terms of patient status in drug abuse treatment is especially common when inpatient hospital admission occurs. For pregnant women in treatment programs, confidentiality is frequently compromised when the patient is admitted for delivery. During the postpartum period, both treatment and research personnel need access to the patient and the neonate. These contacts are difficult to arrange without openly acknowledging the patient's status to medical and nursing staff members.

Hospital Accreditation Standards and Quality Assurance

The complex task of initiating and operating an innovative, comprehensive treatment program with an extensive research component is made more difficult by the need to ensure that the clinical program and the research activities, even as they are in the process of being developed, comply with all internal and external hospital standards. In hospital-based programs, patient assessments, clinician documentation, and staff credentialing and evaluation all must meet the standards for chemical dependence programs of the Joint Commission on Accreditation of Health Care Organizations (JCAHO). These standards are not formulated with allowances for "experimental" or new programming formats. Although JCAHO standards are designed to ensure comprehensive, state-of-the-art treatment service delivery, they are often rigid and inflexibly applied when the JCAHO evaluates drug abuse treatment programs. It is difficult to break new ground in the area of drug abuse treatment and not run afoul of a JCAHO standard.

Similarly, it is difficult for a new research demonstration project serving and studying a chemically dependent population to integrate successfully with internal hospital utilization review (UR) operations and quality assurance (QA) activities. In the early stages of program startup, this may be because the project treatment, research, and clinical staff members are focused on simply getting the program "off the ground." The intense effort involved in setting up a new program may leave treatment staff members and program leaders with little time or energy to devise and carry out routine QA activities. At the same time, designated hospital UR and QA personnel may be unfamiliar with the nature of the services offered in a drug abuse treatment program and therefore may have difficulty accomplishing the necessary independent reviews of treatment program activities. Evaluation of program performance may be impossible for QA staff members unless they obtain special training in the field of addiction services. When institution administrators are ambivalent about

starting chemical dependence services, the special UR and QA needs of the addiction treatment program may be yet more evidence to those institution administrators of the "troublesome" aspect of these types of programs.

PROBLEMS ASSOCIATED WITH THE PATIENT/SUBJECT POPULATION

The social and economic problems associated with poverty and drug use have been described in numerous publications. The overwhelming needs of chemically dependent women greatly affect routine treatment activities and data collection. Patients often lack basic resources such as food, clothing, or transportation.

Housing

Many patients in a research demonstration drug abuse treatment program may be homeless sporadically or continuously. Treatment staff members may spend an inordinate amount of time attempting to arrange temporary or permanent housing. Research associates, primarily master's-level members of the research team, often have difficulty locating patients for time-limited research measures. Patients may live in dangerous places where staff safety during outreach and data collection activities is problematic.

Transportation

Given patient needs and program requirements, assistance with transportation is a necessary component of a drug abuse treatment program for pregnant women. Achievement of regular program attendance is a problem when a woman in treatment must make long, tiring trips, often with small children in tow, using public transportation. When public transportation is the only available option, most women require a financial subsidy, especially if daily program attendance is expected. Hospital administrators may have an expressed or unspoken perception that providing a pickup and dropoff service or financial subsidy for chemical dependence treatment program patients is equivalent to "coddling." In part, this is because the women, even when pregnant, are seen as relatively able-bodied compared with other patient groups served by the institution. Hospital administrators may reason that if "sicker," that is, frailer, patients with physical disabilities do not receive transportation assistance, it is unfair for this group to receive it. There is also frequently a perception that the women and children coming to the addiction treatment program increase the need for insurance coverage carried by the hospital on its minivan transportation system.

Child Care

Child care is another essential service element in a program that cares for pregnant or parenting women with small children. Regular patient attendance cannot be expected unless child care is available to the women enrolled. Child care services may range from simple babysitting for a few hours each day to a specialized, daylong therapeutic nursery. Creation of child care services on hospital premises may prove surprisingly challenging, and many hospital administrators may be inexpert or unmotivated in handling issues that arise. Hospital facilities often will need substantial renovation to comply with the stringent codes that apply to areas where children are cared for. A day care license may have to be obtained. Quality-assured recordkeeping systems, daily screening for communicable diseases, proof-of-vaccination status, and an emergency pediatric care referral system need to be in place even for a "simple" onsite babysitting service. Although the mother is in treatment on the premises, many requirements and specific guidelines must be followed. An additional problem arises during the summer when the demand for child care services becomes heavier because school-age children enter day care. During summer vacation, the program may have to cope with large numbers of children, many of them teenagers, within the confines of hospital premises that are ill suited to accommodate adolescents.

Security Concerns

Even when a treatment program has a limited number of treatment slots or a low census, the regular presence on the hospital premises of a population of low-income, homeless, chemically dependent women and children, together with their significant others, relatives, and acquaintances, may be viewed as a security problem by hospital administrators. Thefts or other untoward incidents that occur within the medical center often may be attributed to the drug abuse treatment program, even when no substantiating evidence is available. Addressing both the legitimate and exaggerated security concerns associated with a drug abuse treatment program operating onsite in a medical center may prove challenging.

PROBLEMS ASSOCIATED WITH INTEGRATION OF RESEARCH REQUIREMENTS AND CLINICAL OPERATIONS

Whereas multidisciplinary programs function best when program staff members are flexible in assuming each other's roles and responsibilities, a potentially negative aspect involves the research staff taking on clinical duties. The research staff members are key figures in recruiting and introducing patients into a program, and they are responsible for following

the patients' progress in treatment. Because research staff members are so active in subject recruitment, the general hospital staff members are often more familiar with them than with the treatment staff members. As a result, the research team personnel often are inappropriately contacted by nonprogram hospital service providers when there is a problem or a concern with one of the program patients.

Limited space for program activities affects data collection. Physical proximity of the research and treatment staffs may make it impossible to maintain blinding with regard to patient assignment to the treatment or control condition. In addition, a research staff member who works next to a treatment staff member may be unwittingly kept abreast of patients' successes or failures in treatment. This knowledge has the potential to influence or bias the data gathered by the research staff from patients assigned to the treatment condition.

Case management often occurs in the course of patient-research staff contacts. As noted above, low-income chemically dependent women have extensive human service and emotional needs. The services the research staff members deliver to the patients, as well as the relationships that form between them as a result of the service delivery and frequent, nonjudgmental contacts for performance of research measures, are likely to alter outcomes and affect data collection. It is difficult to create and maintain a program with sufficient physical and organizational distance among the research team, the clinical team, and the patient cohort, particularly in a hospital setting where geographic separation may be impossible and where other hospital systems exert a "blending" pressure on the research and treatment teams.

PROBLEMS ASSOCIATED WITH INSTITUTIONAL PREJUDICE

Despite widespread public knowledge of the existence of the problem of maternal drug abuse and a stated societal commitment to addressing it effectively and humanely, deep-seated dislike, distrust, and distaste for chemically dependent pregnant women persist within medical centers. Some of this prejudice is the result of a belief that chemical dependence is not a medical condition like some others, that is, one that is worthy of medical treatment or in need of hospital-based services. Another important factor may be the uniquely negative, condemnatory feelings that a pregnant, drug-abusing woman often evokes. It is difficult for many people to acknowledge how powerfully they are affected by the concept of an "innocent baby" being "abused" in utero by an "unfeeling" or "monstrous" mother. These negative and frequently unconscious

feelings exist at every level of hospital personnel, from administrative personnel, to the clinical staff, to support personnel, to the janitorial and housekeeping staff. The prejudice is manifested in different ways, from the trivial to the serious: in prolonged timeframes for setting up program services caused by exacerbation of usual bureaucratic delays and as vague opposition by nurses or resident physicians to program needs and goals. Hospital administrators may require program patients to be sequestered physically as much as possible from other hospital patients. Janitorial and housekeeping service personnel may clean the program space less frequently. This issue is exceedingly difficult to deal with because it is rarely openly acknowledged by those who are responsible for it.

Negative feelings about chemically dependent pregnant and parenting women are exacerbated when the women in question are indigent or of low-income status. Many hospital personnel perceive these patients as dangerous, manipulative, and criminal. A woman's lack of education and inability to support herself and her family are seen by some hospital personnel as a result of her addiction rather than as a contributing factor. Fear and distrust of program patients may be openly or subtly conveyed to them by the hospital staff members they encounter and can be expected to have a negative effect on the patients' self-esteem and progress in treatment.

CONCLUSIONS

In spite of all the problems enumerated above, medical centers are desirable locations in which to conduct research in the area of maternal addiction. The following recommendations and suggestions for overcoming barriers to program implementation and service delivery may be useful for prospective investigators and funders.

- Substantial increases need to be made in the amount of time allowed for program startup. Additional time should be allocated for piloting the treatment program with a patient group whose data will be excluded from the final analysis.

- Extensive involvement of all branches of medical center administration and all hospital departments is mandatory in the planning of a proposal for treatment and study of chemically dependent women. Early, coordinated planning may help ameliorate or remove existing prejudices and promote institutional investment in the clinical research project's success. The planning should involve hospital QA and risk management personnel and relevant outside accreditation bodies. Space commitments and space renovation plans should be made as far in advance as possible.

- Widespread and ongoing education about drugs and addiction is needed for all hospital personnel. Inservice training provided by program staff members, as well as collaborative presentations by the mental health department and employee assistance program personnel, should be held throughout the medical center so that the treatment research project will become an accepted and vital part of medical center activities.

ACKNOWLEDGMENT

Ms. Quinshaunta Golden, program administrator, provided useful commentary and suggestions regarding the administrative problems encountered by the program.

AUTHOR

Ellen Mason, M.D.
Medical Director
New Start Program
Department of Obstetrics and Gynecology
Cook County Hospital
B Building, Room 241
1835 West Harrison Street
Chicago, IL 60605
(312) 633-3605 (Tel)
(312) 633-5495 (Fax)
emason%chi-cook@hub.fhcrc.org (Internet)

Two Therapeutic Communities for Substance-Abusing Women and Their Children

Peggy Glider, Patrick Hughes, Rod Mullen, Shirley Coletti, Lee Sechrest, Robert Neri, Bobbi Renner, and Donna Sicilian

INTRODUCTION

How does one design and implement a randomized clinical trial in a therapeutic community (TC)? This was the question confronting two Perinatal-20 Treatment Research Demonstration Program grantees, Amity, Inc., in Tucson, Arizona, and Operation PAR, Inc., in St. Petersburg, Florida. This chapter describes the issues encountered by these two agencies in developing clinical laboratories for research demonstration programs in two TCs for substance abusers. Some of the problems were readily worked through, but several remain.

Both research demonstration programs proposed randomized clinical trials to determine the efficacy of permitting substance-abusing women to bring one or more of their children to live with them in a TC. The idea had its origins in the 1960s when Synanon allowed some staff members and residents to bring their families to its California TC to live, realizing that some women would not enter the program if they could not bring their children. Even with the addition of children, Synanon was still primarily a male-oriented model and did not focus on the family unit. Although an etiological study of the children raised in the community was conducted (Missakian 1976, 1977),[1] no formal research was conducted on the effect of resident family members on treatment outcome for the women (Yablonski 1989).

Odyssey House in New York City appears to be the first TC to evaluate the effect of admitting children in a National Institute on Drug Abuse (NIDA)-funded demonstration program in the late 1970s. Cuskey and associates (1979) evaluated the program and found that substance-abusing women admitted to this TC—primarily women with children, pregnant

[1]Dr. Elizabeth Missakian received a National Institute of Mental Health grant to conduct an etiological study of children raised in a TC. Synanon designed an extensive kibbutz-like school at its Tamalas Bay, California, facility, which at one point had 160 children ranging in age from newborns to 17-year-olds. Outcomes showed that the children raised in the Synanon school scored as high as, and in many cases higher than, children raised in high socioeconomic status households in Marin County, California, on measures of emotional, psychological, and academic functioning.

women, or pregnant women with additional children—had better-than-expected postdischarge outcomes. Improved outcomes included decreased substance use and criminal behavior, increased employment, and improved self-concept, mother-child relationships, and parenting skills. In addition, these outcomes improved with longer stays in the residential program.

In 1982 Amity permitted women to bring their children into its TC residence and found that the women dramatically increased their length of stay (Stevens et al. 1989). Given the promising results of these initial programs, the authors felt further investigation was indicated.

The shared hypothesis of the two studies stated that admitting substance-abusing women into a TC with their children would increase the women's retention in treatment and would improve the long-term outcomes for them and their children. The authors felt a randomized clinical trial design would be optimal, but the literature contained few precedents for clinical trials within TCs. Bale and colleagues (1980) randomly assigned substance-abusing male veterans to varying modalities, including a TC. More recently, De Leon (1991, pp. 218-244) compared approaches to initial socialization into a TC by offering different approaches to admission cohorts during alternate months. However, no studies in the literature reported randomly assigning participants to different treatment conditions within a TC.

This chapter describes how Amity and Operation PAR developed laboratories in their TCs for randomized clinical trials to test the hypothesis explained above. Problems encountered and their solutions also are described. The chapter also discusses the implications of the authors' joint experiences for future treatment research in TCs.

IMPLEMENTATION OF THE DEMONSTRATION RESEARCH PROGRAMS

Study Design

A randomized, open treatment trial was the design selected by both Amity and Operation PAR. The studies hypothesized that women who were permitted to live in a TC with one or more of their children would stay in treatment longer and have better posttreatment outcomes than women who received the standard TC treatment (no children).[2] Improved

[2]In the standard TC, women must enter without their children. This means that a woman must place her child(ren) either with family members, often the very individuals who abused the woman as a child

outcomes mean less substance use and illegal activity and better attitudes toward parenting and child-rearing, employment, and other social activities. Women who agreed to participate and met eligibility criteria were randomly assigned to (1) the demonstration program with infant and child care available (experimental group) or (2) the standard TC without child care (control group). Eighteen months of residential care was projected for both treatment groups.

Eligibility criteria differed in the two research demonstration programs. Both programs required that women have custody of their children or legal permission to admit them to the TC and have no legal constraints that required incarceration following treatment. Operation PAR required that women be 18 years or older and meet Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised) (American Psychiatric Association 1987) criteria for cocaine abuse or dependence. Amity admitted women who were 18 years or older with a history of abusing any type of substance (including alcohol), although several adolescents (ages 15 to 17) also were included.

Both programs allowed each woman to bring up to two infants or children into residence. Amity set the child's upper age limit at 8 years; Operation PAR allowed children up to 10 years of age to live with their mothers. Pregnant women were not included in the Amity study.[3] Participants with serious or life-threatening medical problems—other than those who were asymptomatic HIV positive—were not eligible for either program. An exception was if a child or mother became ill while in the program. One participant at Amity gave birth to a severely medically compromised infant. The mother and child remained in the program, receiving all needed care and support.

Treatment Conditions

Amity operates two TCs in Tucson with a joint capacity of 220 beds (120 in the men's facility and 100 in the women's facility). The programs are located on two former guest ranches, separated only by a road. This proximity allows for joint programming and special community events. In addition, both the men's and women's programs have a similar

---

and/or who led her into substance abuse, or in foster care. Many women who have already lost their children to child protective services have no voice regarding what happens to their children when the women enter treatment.

[3]Pregnant women were not allowed into the Amity study because Amity staff members felt it would not be ethical to randomly assign a pregnant woman to the control group (no children) and thus have to remove the child from the mother after birth. To compensate for this, Amity wrote a proposal for and received a demonstration grant from the Center for Substance Abuse Treatment to provide residential treatment to pregnant and postpartum women and their children.

curriculum and programmatic structure, although sex-specific activities have been added for the women. Special buildings in the women's facility (with a capacity for housing 40 women and 80 children) were designated for use by study participants. Amity's developmental learning center at the women's facility provides specialized services throughout the day, after school, and on weekends for children in residence and for those who visit. Standardized developmental assessments are completed on each resident child on a quarterly basis. Staff members work to meet the needs of mother and child individually and as a family unit. Staff members for both the women's and children's components include academically trained professionals with degrees in social work, education, and nursing as well as experienced staff members who are certified substance abuse counselors and are credentialed by Therapeutic Communities of America, Inc. (TCA).

PAR Village is the Perinatal-20 treatment demonstration component of Operation PAR's 149 adult-bed TC in Largo, Florida. The Village permits women to live with their children while they participate in residential treatment. PAR Village is separated from the main TC grounds by a small pond and a 1-minute walk. It is composed of 14 houses with a capacity of 14 women and 28 children younger than 10 years. Operation PAR's child development center (CDC) occupies one building in the Village. It was licensed by the local government as a day care facility and had a capacity of 35 infants and preschool children. CDC activities were supervised by an experienced clinical social worker, and the staff included eight early childhood specialists and a master's-level clinical supervisor. Using standardized instruments, developmental specialists conducted individual assessments of each child. The child's developmental intervention plan was formulated with the participation of the mother. A developmental pediatrician visited once or twice a month to perform developmental evaluations and prescribe special interventions. A developmental psychologist visited weekly for additional testing and to focus on mother-child interactions and bonding.

Both Amity and Operation PAR had experience in developing special programming for substance-abusing women before implementing this study. Amity had made this a special emphasis beginning in 1981 and by 1982 had doubled the number of women participants. From 1982 to 1987 Amity permitted a few women to bring their children to live with them in the TC (Stevens et al. 1989); therefore, experienced staff members were available to begin the project in 1990. Operation PAR established its first program to meet the special needs of substance-abusing women and their children in 1987 so that experienced staff members were available to implement the demonstration program in 1989.

The general design of the program for women and children was similar in both the Amity and Operation PAR programs. Women in the experimental and control groups participated in the same habilitation activities and attended the same therapy groups and parent training classes. The major parameters of the 18-month TC programs adhered closely to the model described by De Leon and Rosenthal (1989, pp. 1379-1396).

The only differences for women in the experimental and control groups were the living quarters and the availability of onsite child care for the experimental group. Amity offered several housing configurations for women with children. In some, two women shared a bedroom, and their children shared adjoining bedrooms. In others, a woman and her children shared a bedroom as did other mother-child families in a house, with all families sharing the kitchen and family room. At Operation PAR, most experimental group women had one bedroom for themselves and an adjoining bedroom for their children, sharing a common kitchen and family room with another woman and her children. The pattern also varied, with some women sharing a room with their children; some houses were large enough for three or four women with their children.

Women in both programs were expected to take their children to the CDC in the morning, pick them up by 4 p.m., and care for them during evenings and weekends. At PAR Village, women were responsible for arranging pediatrician visits and babysitting, if they wished to attend a meeting during the evening or on the weekend. Babysitting was provided by community volunteers and TC participants who had received babysitting training.

The Amity control group women lived at the same facility as the experimental group women; however, they generally were housed in separate buildings. Women in the control group at Operation PAR slept in a 20-bed building on the adjoining TC campus, with an average of 6 to 8 women to a room. Consistent with the usual TC practice, control group women in both the Amity and Operation PAR programs did not live with their children. Their children resided with family or friends or in foster homes and were brought to the TC to spend time with their mothers during visiting hours. At Operation PAR these children were allowed to participate in outpatient groups held several times weekly for all children whose mothers lived at the TC. Transportation was provided for these groups.

IMPLEMENTING THE DEMONSTRATION PROGRAMS: PROBLEMS AND SOLUTIONS

Facilities

Although some remodeling was required at the Amity facility to accommodate children, the first cohort of women was admitted into the program and moved to the facility within 3 months of the grant award. These women assisted in readying the property for their children. All necessary licenses for occupancy and provision of services were already in place and covered the new program. Therefore, there were no facility-related delays in program implementation for Amity.

PAR Village also was able to implement its program quickly. Although full construction and renovation of the houses at the Village were required, work progressed quickly enough to begin admitting women and children 7 months after formal notice of the grant award. As a nonprofit community agency, Operation PAR was able to avoid the red tape that would have been inevitable if the facilities had been acquired by the local government or a university. Operation PAR was able to negotiate with a nearby city government to acquire 14 houses scheduled for demolition, obtain funds to move them to the TC campus, and then renovate them to meet programmatic, zoning, and licensing requirements.

Personnel

Both Amity and Operation PAR had the staff depth needed to expand the TC program to meet the special needs of women. For Amity, no new staff members were required to work with the women; experienced staff workers in other special programs for women were moved to this program. At PAR Village, women were served by existing TC counselors; about half were male.

Experience and expertise for infant and child care were limited at both programs. Amity had only one special education teacher, who became the program manager for the children's program. Operation PAR had only one child therapist. In addition, the new programs were to be broader in scope and more intensive than these staff members had previously experienced. Both Amity and Operation PAR had to recruit an entire team for their children's programs. The teams had to be trained not only in specific job responsibilities but also in the TC model and women's substance abuse issues.

Program for Infants and Children

Although Amity had previously provided the opportunity for some children to live with their mothers, a proposed therapeutic learning center for this program had not been established. This necessitated the development of a children's program in its entirety (i.e., facility, staff, curriculum).

Operation PAR already had established a child day care center for outpatient substance-abusing women in 1987 and had experience in acquiring age-specific furniture, educational materials, and toys. However, this program also included infants with perinatal cocaine exposure and provided an 8-hour-per-day learning environment. PAR Village staff members screened each child to identify developmental problems and establish individual treatment plans. The infants program was staff intensive: Local regulations required 1 trained staff member for every 3 infants younger than 12 months, for every 5 children between ages 13 and 24 months, for every 10 children ages 25 to 36 months, and for every 15 children older than 3 years. Because of the special needs of some children, it was often necessary to have a higher staff member-to-child ratio than regulations required.

Both programs provided for the physical safety of the children. Because the mothers and other participants were serious substance abusers who had been victims of sexual violence and other abuse and had neglected their families, safeguards were required to protect children from abuse. Procedures were introduced at Amity and Operation PAR to screen all residents for history of child abuse and potential for future abuse. Differences in the physical facilities at Amity and PAR Village led to different solutions. At Amity, the main TC campus was physically separate from the perinatal program, so it was possible to provide security by not permitting participants identified as potential child abusers to visit the perinatal program facility. At PAR Village the perinatal facilities and main TC campus were adjacent to one another, and the two programs were fully integrated. Therefore, it was necessary to transfer several PAR Village residents to other programs and to screen future admissions for child abuse potential. All participants (male and female) were educated regarding appropriate behaviors and language to use when in proximity to young children.

Program for Women

One of the major problems encountered by both programs was the issue of special privileges for women with children. Although TCs believe that all individuals are entitled to enter the TC if they are interviewed and found suitable by other participants and the staff, once individuals are in

the program, all their privileges must be earned. Everyone starts at the bottom and works up. Certainly one of the greatest privileges is for mothers to have their children live with them. The random assignment of women to the experimental or control group removed the community "clinical judgment" regarding the readiness of the women to receive this privilege. This goes against one of the most central tenets of the TC—that privileges must be earned and must be acknowledged by the community. An additional tenet of the TC dynamic is that residents be treated fairly and equally. That some women should receive superior or inferior treatment is an affront to these principles. Experimental group women required time for care of their children, so they had to be freed from some responsibilities assigned to control group women without children. They were given privileges such as additional snacks and rest periods. The living accommodations for women with children permitted greater privacy. Experimental group women were seen as receiving these rewards without working for them. Some women who did not have their children with them raised the issues of fairness and special treatment.

Clinical staff members were concerned about the negative consequences for children when their mothers had to be terminated from treatment for noncompliance or a major rule violation. For this reason, some Amity staff members were initially overprotective of the children by not permitting mothers to make mistakes and overlooking violations, which caused some experimental group women to believe and act as if they did not have to follow TC rules. At first, some of these women felt they should receive special treatment for participating with their children and providing research data.

To address the variety of issues raised by the demonstration programs, the following steps were taken at Amity (A) and/or PAR Village (P):

- Schedules were rearranged for *all* women in the program (experimental and control groups) so that women with children were neither missing out on essential curriculum components nor being dismissed from duties that all TC participants were expected to do (A, P).

- Parenting classes were provided for all participants, including those who were not in the demonstration program (P). All women (experimental and control groups) received parenting classes (A).

- Emphasis on staff training and TC resident curricula focused on helping the women take responsibility for their own decisions and behaviors as well as helping each other be accountable (A).

- Babysitting training classes were offered to all appropriate TC residents, and a roster of evening sitters was developed (P).

- Weekly family meetings consisting of the woman's primary counselor, a developmental specialist, the woman, and her child(ren) were held to assist with emerging problems. Meetings also were held with the control group women during or following visits with their children (A). Such meetings were held only when a mother or staff member felt it was indicated (P).

- A mother's group was organized for both experimental and control group women to teach basic skills such as housecleaning and organization, infant/child hygiene and feeding, meal preparation and nutrition, and time and money management (A, P).

The need for emergency and routine medical care for resident children and pregnant women at PAR Village was greater than expected. Much staff time was required to accompany ill children and mothers to the health care provider, which increased the workload of staff members who remained at the facility. Another significant expense at PAR Village was a threefold increase in the cost of prescription medications.

RESEARCH PROBLEMS AND SOLUTIONS

Randomization

When the programs were originally designed, administrative and clinical staff members in both programs expressed concerns about the randomization of participants to the two treatment conditions. They felt it was insensitive and arbitrary—principles that were not acceptable in a TC. However, it was recognized that randomization was the most scientifically sound procedure by which to study substance abuse in women and that any other procedure was less likely to be funded in a research demonstration proposal. In addition, programs for women were scarce, and the control side of the random assignment (treatment without children) was considered to be an acceptable alternative. Faced with the likely choice of *no* program for women and children or one with randomization, the clinical staff acquiesced to use of randomization.

The method for making the random assignment was considered to be extremely important. Amity staff members felt that the woman should be totally removed from the process, thus removing any sense of guilt for not choosing "correctly." It also was considered important to remove the clinical program staff from any connection with the decision so that staff

bias for or against a woman would not be an issue. The final decision was to use computer assignment, which would be given to each woman by the research assistant.

The Operation PAR team had similar concerns about randomization but developed a different procedure. The research team felt that the women might be suspicious if a computer assigned them to the less desirable treatment group. The procedure for assignment would have to be seen as fair and credible. The procedure chosen was to have the participant pick one of two cards that looked identical on one side but on the other side were printed with the words "Mom + baby in program" or "Mother only in program." The cards were periodically examined and replaced to avoid the possibility that participants could identify and choose the preferred card.

Clinical staff members and most women in both programs felt that the experimental treatment was superior and were highly sympathetic toward women who were randomized to the "wrong program." When randomization did not permit a woman to have her child(ren) with her, she would invariably be upset; some women became extremely distraught. To help resolve this problem, it was requested that a clinical staff member be present at the randomization to serve as witness to a proper and fair choice and provide support if the woman became upset.

The timing of when randomization occurred in the course of each woman's treatment differed in the two programs and within each program over time. The Amity program originally set the point of randomization at 3 months into treatment. This timeframe was chosen so that women would be fairly well integrated into the TC and less likely to drop out if they did not like the random assignment they received. After approximately 6 months of using this timeframe, staff members decided to shorten the randomization point to 1 month, which was done to increase the number of women in the program. (Too many women dropped out before the 3-month randomization point.)

An unanticipated complication of performing the randomization earlier in the woman's treatment was greater difficulty in clarifying the availability of the children for the program. At intake into Amity, several women said they would be able to admit their children to the program should they receive the random assignment. Unfortunately, the shortened period of 1 month prior to randomization did not allow enough time to negotiate custody details with child protection agencies that were frequently involved with children being considered for the study. This resulted in some women who were assigned to the experimental group being unable to admit their children to the Amity TC. Because of the nature of the research question, it was desirable for the child to enter the program within the first 6 months

of the mother's treatment. If this did not happen, the mother was removed from the study and placed in a comparison group consisting of all women at Amity's TC who were not part of the demonstration program.

In the pilot phase of the Operation PAR program, randomization was done prior to admitting the woman to the TC. This resulted in some control group women subsequently not reporting for admission. The protocol had to be changed to require TC residence of the mother at least several days before randomization. The majority of children of Operation PAR women were also in protective custody, which caused delays in their admission. There were several other potential Operation PAR participants who did not receive child protection agency permission to admit their children.

Randomization remained a chronic, unresolved issue. In both programs, researchers continued to prefer randomization for scientific reasons whereas clinicians and participants continued to be greatly distressed by it; this remains a significant, unresolved issue. It would be difficult to obtain TC clinical staff agreement to a similar randomized clinical trial in the future. The randomization in this study involved one of the most emotional issues that a woman could confront: whether to be with or without her child(ren) for an extended period. Most likely, randomization around other, less sensitive issues would be better accepted within the TC.

Early in the study clinical staff members of both programs and participants raised the question of whether a woman randomized to the control group could drop out of the study and reapply, thereby having a second, and possibly a third, chance to have her child(ren) with her in the TC. This question was raised because the women believed the experimental group condition was more desirable. To avoid questions of this nature, the clinical staff members and researchers of both programs developed a similar rule. For Operation PAR, this was called the "Two Worlds Rule": Women in one "world" could enter a TC and reside there with one or two children. There also was another world where women could enter a TC only without their children. Once assigned to either world, a woman would remain there until her 12-month posttreatment followup was or could have been completed. This rule was useful in explaining and helping to deal with questions about changing group assignment.

Maintaining Integrity of the Two Treatment Conditions

The study design called for the two groups of women to have as similar a treatment experience as possible. One difference was that experimental group women had overnight responsibility for their children, beginning at approximately 4 p.m. and ending at 8 a.m. the next day.

The issue was dealt with differently by the two programs. Given the tradition of equality and fairness for TC residents and the value placed on the mother-child relationship, clinical staff members wished to permit the control group women to have more frequent visits with their children and other family members. This increased visitation was written into the Amity grant so that the only significant difference between the two groups was the opportunity and responsibility of having children live with their mothers. All other opportunities to form mother-child attachments and to have hands-on parenting experiences with staff guidance and feedback were provided for both groups (although experimental group women had more time for these experiences). This design was considered sound by the research team because the issue of child residence was the main research question under study by the program. Should differences be found between the groups, these differences could be clearly attributed to the one factor that varied (mothers living with or without their children).

At PAR Village, the issue of equality between groups and the importance of the mother-child relationship also was raised. Treatment staff members and residents became increasingly enthusiastic about the value of mother-infant bonding and raised the issue of bringing the infants of control group women in for a weekend to facilitate bonding. In contrast, research staff members cautioned that the more the control and experimental programs resembled one another, the more likely no treatment group differences in outcome would be found at followup. Research staff members continued to "resist" efforts to make the control group experience more like that of the experimental group with respect to the major independent variable in the study. An exception was made for one control group mother who was pregnant when she entered the program and delivered in her second month of residence. She was permitted to have her infant with her in the TC for 3 weeks following delivery. The children of control group women attended weekly outpatient groups, participated in monthly outings, and occasionally slept over on weekend visits.

Ideally, the treatment in a clinical trial should remain the same from beginning to end, but the TC is a complex living, dynamic, and everchanging group process, with no possibility of freezing this process at any given time. However, it was possible to keep the two treatment conditions distinctive in their essential difference: child living with vs. not living with the mother in the TC.

Clinicians and Researchers Sharing Information

The research protocols for both Amity and Operation PAR were written to assure participants that they could report information to the researchers without fear of sanctions by the treatment staff or agencies to whom the

treatment staff reported (e.g., courts or child protective services) unless required by law. This policy did not pose problems at Amity, but it became an issue at Operation PAR regarding the sharing of urine drug test results.

In the initial planning meetings of the Operation PAR clinical and research teams, the principle of preserving an independent research database was accepted. It was agreed that the laboratory would send separate urine drug-testing reports to clinicians and to researchers so that the two groups need not share data. This seemed a reasonable arrangement, but problems developed. Research interviewers, seen as "TC outsiders," initially encountered difficulties in obtaining observed urines and having them sent to the laboratory. After a series of frustrating experiences, a sympathetic TC administrator facilitated the urine collection and shipment to the laboratory. Given the frustrations and delays, as well as recruitment of a new research interviewer, the "split urine" agreement was forgotten, and only the research team received reports. In the 18th month of the project, this developed into a major issue when a clinician heard about researchers collecting urine specimens and clinicians not receiving reports.

A series of tense meetings was held in which researchers acknowledged an honest mistake but were not able to turn over laboratory results or to tell clinicians whether there were positive urines in the TC. Clinicians were frustrated, and researchers were bewildered by the strong emotions and temporary breakdown in communications. The issue was extremely important to TC staff members, who felt frustrated with the researchers for not living up to an agreement, and to researchers, who wondered why urine drug-testing reports were suddenly so important when the researchers had such problems getting them done. For clinicians, not having results of the urine tests raised concern as to who among the participants might have been using drugs without staff knowledge. Not revealing information about a participant's drug use strikes at the heart of the TC dynamic.

The immediate crisis was resolved by holding a meeting with the study women residing in the TC and explaining the breakdown in procedure. Clinical staff members requested that the women voluntarily give permission for release of their urine test results, and all complied. One woman was informed she had a positive urine, which she disclosed to the clinical staff. It was handled as a treatment issue in keeping with usual practice at the TC.

Following this crisis, Operation PAR researchers surveyed other treatment research programs and noted a frequent arrangement in which drug screens were done at specific times as defined in these programs'

protocols. The dominant practice was to have both clinicians *and* researchers receive copies of the reports. For clinicians to receive a positive drug screen report was considered a benefit to the participant—not a danger—because treatment staff members then could work more intensively with the participant on the relapse issue. This would be a reasonable and economical arrangement in future studies. For Operation PAR, a change in procedure would require, among other things, the approval of the funding agency, review by the human subjects committee, a change in the consent form, and an approach to the women in the study with new consent forms. The original agreement to split urine screens between the clinicians and researchers was not changed; it was simply implemented.

Committed Clinicians vs. Dispassionate Researchers

During periods when TC clinical staff members were concerned about the restrictions of the study design, the research interviewers occasionally would receive negative comments during visits to the TC. On occasion, research staff members would be the recipients of highly emotional communications from both the clinicians and research participants. Although most of the researchers in both programs absorbed these communications with equanimity, some research staff members responded with frustration, bewilderment, and anger.

A related problem was the different value orientations of clinicians and researchers. The researchers were trained to be dispassionate, rational, and objective. The TC movement was a rebellion against the "professional" approach to treatment. A basic tenet of the TC is the removal of the "we/they" dichotomy. The staff members at Amity and Operation PAR, many of whom had experienced the recovery process, felt that professional distance is harmful and represents the complete antithesis of an effective TC. In comparison, research staff members in both projects saw this professional distance as crucial to appropriately performing their jobs. At times, clinicians would become uncomfortable when researchers spoke matter-of-factly about relapse of subjects, "interesting" relationships in the research data, and potential participants who did not meet study admission criteria even though they were in great distress and required the type of help available through the study.

Both Amity and Operation PAR resolved these differences in much the same way. These incidents and issues were discussed informally and at the monthly formal team meetings. Differences in the communication and emotive styles of clinicians, researchers, and participants were accepted, and substantive issues were resolved. For example, Amity and Operation PAR each held two slots in the mother-with-child condition as "compassionate beds." These slots were filled by the clinical staff on the

basis of need with women who did not fit the study design and were not considered part of the study.

Subject Recruitment

This has been one of the most difficult problems to resolve for both programs. When Amity wrote the grant proposal for this program, there were no other services available in the Tucson area for substance-abusing women and their children. By the time the grant was awarded, three local programs for women had been established. Because the other programs did not require random assignment, eligible women applied there first. This reduced the pool of eligible participants for Amity. In addition, most major referral sources expressed concern over not knowing which treatment condition their participants would be receiving and preferred to refer participants to programs where the women would be with their children. Through intense efforts, this situation improved but continued to be a recruitment problem.

Initially, recruitment into Operation PAR was slow because the original study design required each child to have been exposed to cocaine during pregnancy and to be younger than 6 months of age. To speed up recruitment, the perinatal cocaine exposure criterion for children was dropped, and the age criterion was expanded to 10 years or younger. Recruitment again became a problem in the second year when competing programs for female substance abusers were established in the local community. Although Operation PAR was able to keep the program beds full, competing programs continued to cause recruitment problems.

DISCUSSION

Positive Accomplishments

Researchers often refer to the TC as the "black box"—a complicated and everchanging treatment system that cannot be operationally defined for systematic clinical trials. Recent efforts by TCA have attempted to gain consensus on the essential components of the TC, and a generic TC program has been described (De Leon and Rosenthal 1989, pp. 1379-1396). This chapter describes two examples of TCs that collaborated with university researchers to make the TC black box more friendly to rigorous research.

An important aspect of the programs is that they demonstrated the feasibility of conducting randomized clinical trials in TCs. It was possible to develop clinical laboratories for the research by establishing facilities

and programs for children and their mothers within existing TCs. Additional time will be required to test the study hypotheses.

Shared Problems

People in both Amity and Operation PAR had serious problems with randomization, primarily because of the mismatch between random assignment and the basic tenets of the TC. In addition, an issue raised by Amity concerned the probability that the random assignment in effect changed the program and, therefore, did not measure the TC model as generally implemented. People in both programs have strong reservations about participating in a future randomized trial to compare treatment conditions where staff members and participants believed one condition was clearly superior.

The types of services described herein were not available at the time the study was designed. Moreover, publicity surrounding the programs helped stimulate expansion of treatment services for women and their children. In addition, most of the women in this study did not have legal custody of their children at the time of admission. Randomization to the control group continued the noncustody relationship with their children. Randomization indeed facilitated the ability of experimental group women to live with their children during long-term treatment—a highly unlikely scenario if the study had not been conducted.

Both programs observed tensions between clinicians and researchers, although such tensions and differences in orientation are not unique to these two programs. Some differences may become exaggerated when researchers work with TC clinicians, who expect total respect for the TC treatment environment. In these programs, clinicians did not decide which women would be allowed to bring their children into treatment, but in the traditional TC, this decision would be made by a concerned community of staff members and senior residents—not by random selection or computer assignment.

TCs frequently encourage expression of intense emotions and often have group norms that differ from those of a research staff. Researchers who visit TCs often are trained to be unemotional and dispassionate in their work and frequently are unaware of how these styles diverge from TC norms. In this program, initial training of research staff members attempted to sensitize them to participant issues by asking researchers to review clinical records and attend treatment team meetings. Despite such efforts, tensions and misunderstandings can be anticipated, and clinicial and research teams should meet frequently to resolve issues. It is also

important to maintain strong support for the unique requirements of the research at the highest administrative levels of the TC.

Implications for Therapeutic Community Research

Prior to these two studies, researchers had implemented clinical trials in which participants were randomly assigned to different treatment modalities, including TCs (Bale et al. 1980). The two studies described in this chapter demonstrate the feasibility of clinical trials that randomly assign participants to different treatment conditions within TCs, albeit not without complication. Clinicians, researchers, and participants were blind neither to the psychosocial treatment conditions being compared nor to which participants received which treatment. In both the Amity and Operation PAR studies, one treatment was viewed by clinicians and participants as more desirable, which caused considerable tension and had the potential for biasing results.

Yet the overall experience suggests that the TC modality is likely to benefit from further experience with the clinical trial methodology, especially in examining the efficacy of treatment innovations. The Center for Therapeutic Community Research, headed by George De Leon, Ph.D., is the recipient of a NIDA research center grant to conduct a systematic study of TCs.[4] The center will focus on the issues encountered by the two programs described herein. This argues well for increasingly systematic clinical research in a collaborative network of TCs.

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised).* Washington, DC: American Psychiatric Press, 1987. 567 pp.

Bale, R.N.; Van Stone, W.W.; Kuldau, J.M.; Engelsing, T.M.; Elashoff, R.M.; and Zarcone, V.P., Jr. Therapeutic communities vs. methadone maintenance. A prospective controlled study of narcotic addiction treatment: Design and one-year follow-up. *Arch Gen Psychiatry* 37(2):179-193, 1980.

Cuskey, W.R.; Richardson, A.H.; and Berger, L.H. *Specialized Therapeutic Community Program for Female Addicts.* National Institute on Drug Abuse Services Research Report. DHEW Pub. No. (ADM)79-880. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1979. 31 pp.

[4]De Leon, G. Principal Investigator, National Institute on Drug Abuse center grant # 1P50-DA-007700-27, Center for Therapeutic Community Research, N.D.R.I., 11 Beach Street, New York, New York 10013, 1993.

De Leon, G. Retention in drug-free therapeutic communities. In: Pickens, R.W.; Leukefeld, C.G.; and Schuster, C.R., eds. *Improving Drug Abuse Treatment*. National Institute on Drug Abuse Research Monograph 106. DHHS Pub. No. (ADM)91-1754. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

De Leon, G., and Rosenthal, M.S. Treatment in residential therapeutic communities. In: Kleber, H., ed. *Treatment of Psychiatric Disorders: A Task Force Report of the American Psychiatric Association (Volume 2)*. Washington, DC: American Psychiatric Press, 1989.

Missakian, E. "Social Play Behavior in Communally-Reared Young Children." Paper presented at a meeting of the Western Psychological Association, April 1976.

Missakian, E. "Social Behavior of Young Children: An Ethnological Study." Paper presented at a meeting of the International Ethnological Association, August 1977.

Stevens, S.; Arbiter, N.; and Glider, P. Women residents: Expanding their role to increase treatment effectiveness in substance abuse programs. *Int J Addict* 24(5):425-434, 1989.

Yablonski, L. *The Therapeutic Community: A Successful Approach for Treating Substance Abusers*. New York: Gardner Press, 1989. 220 pp.

ACKNOWLEDGMENTS

The Amity program was supported by NIDA grant R18-DA-06918. The Operation PAR program was supported by NIDA grant R18-DA-06369 and by the Center for Substance Abuse Treatment demonstration grant H86-SP05699.

AUTHORS

Peggy Glider, Ph.D.
Executive Director
Quail Enterprises
3965 West Orangewood Drive
Tucson, AZ 85741
(520) 744-2153 (Tel)
(520) 744-6401 (Fax)

Patrick Hughes, M.D.
Professor
Institute for Research in Psychiatry
University of South Florida
Psychiatrist
J.A. Haley Veterans' Hospital
3515 East Fletcher Avenue, Third Floor
Tampa, FL 33613
(813) 972-2000, ext. 6647 (Tel)
(813) 978-5808 (Fax)

Rod Mullen
Executive Director
Amity, Inc.
P.O. Box 32200
Tucson, AZ 85754
(520) 749-7144 (Tel)
(520) 749-5569 (Fax)

Shirley Coletti
President
Operation PAR, Inc.
Suite 1000
10901-C Roosevelt Boulevard
St. Petersburg, FL 33716
(813) 570-5080 (Tel)
(813) 570-5083 (Fax)

Lee Sechrest, Ph.D.
Professor
Department of Psychology
University of Arizona
Tucson, AZ 85721
(520) 621-9182 (Tel)
(520) 621-6320 (Fax)
sechrest@u.arizona.edu (Internet)

Robert Neri, M.A.
Vice President
Residential Services
Operation PAR, Inc.
Suite 1000
10901-C Roosevelt Boulevard
St. Petersburg, FL 33716
(813) 570-5080 (Tel)
(813) 570-5083 (Fax)

Bobbi Renner
Department of Psychology
University of Arizona
Tucson, AZ 85721
(520) 326-1465 (Tel)

Donna Sicilian, M.S.W.
Director
Children's Services
Operation PAR, Inc.
Suite 1000
10901-C Roosevelt Boulevard
St. Petersburg, FL 33716
(813) 327-7900 (Tel)
(813) 726-8106 (Fax)

Case Management: A Method of Addressing Subject Selection and Recruitment Issues

Mary Ann LaFazia, Jeanne Kleyn, Jean Lanz, Teri Hall, Kris Nyrop, Kenneth D. Stark, Christopher Hansen, and D. Heather Watts

INTRODUCTION

Enrolling pregnant substance-using women into chemical dependence treatment continues to present a major challenge for social services and health care professionals. Recruitment problems are exacerbated further when enrollment includes participation in a clinical trial. In 1989 Washington State embarked on a 6-year research and demonstration project funded through the National Institute on Drug Abuse (NIDA) to study optional treatment modalities and the recruitment of pregnant substance-using women. The project was implemented in Washington State as the King County Perinatal Treatment Project (MOM's Project). Project researchers hypothesized that pregnant substance-using women and their infants would have improved pregnancy and child development outcomes if mothers participated in chemical dependence treatment during pregnancy. Project eligibility criteria required women to be in their first or second trimester of pregnancy (prior to 28 weeks gestation), be older than age 16, and have a diagnosis of chemical dependence.

This chapter describes the difficulties and successful strategies associated with recruiting pregnant women into chemical dependence treatment under the auspices of a clinical trial. Some of the recruitment problems were associated with the refusal of the women to participate in a clinical trial. Other problems were related to the nature of the disease of chemical dependence. Still other problems stemmed from the reluctance of social services and health care providers (traditional gatekeepers) working with this population to refer clients for treatment. The success of the recruitment efforts resulted from use of community maternity case managers as primary recruitment agents. Recruitment strategies incorporated a project philosophy of patient recovery and used recruitment scripts that featured brief motivational interviewing techniques.

A major ameliorator of the negative consequences of alcohol and other drug use during pregnancy is participation in chemical dependence treatment. Preliminary findings of the MOM's Project suggest a positive

trend in pregnancy outcomes for women who remain in treatment 90 consecutive days or longer. For example, these women have demonstrated a reduced likelihood of delivering preterm babies (less than 37 weeks gestation), experiencing intrauterine demise or neonatal death, delivering babies who produce a positive toxicology screen, and delivering babies who require an extended hospital stay.

RECRUITMENT CHALLENGES

Recruitment can be considered a two-stage process: *finding* pregnant substance users and *engaging* the women into treatment. Agencies in contact with pregnant substance-using women may fail to make referrals for treatment, although a majority of these women eventually become known to community social services and health care professionals. These gatekeeper clinicians are often reluctant to confront women about their chemical use for fear the women will stop using prenatal and other services. Clinicians also may be unaware of the availability of chemical dependence treatment in their community, lack the skill to recognize addictive diseases, or have negative impressions of the efficacy of treatment.

There may be few referrals even when other chemical dependence treatment agencies are involved. The chemical dependence treatment providers outside the MOM's Project resented the use of non-chemical-dependence personnel as recruiters. They expressed concerns that project case managers "enabled" pregnant women by providing them an array of support services that reduced the consequences of drug-using behavior.

If clinical trials are involved, recruitment challenges increase. Some initial problems in recruitment of eligible women into the MOM's Project centered on two project criteria: the early gestation criterion (prior to 28 weeks of pregnancy) and the randomization process. Potential referral sources, not understanding the nature of a clinical trial, were resentful that their referrals could be rejected solely on the basis of gestational age. They viewed the project as a resource that should be open to all pregnant women in need. This initially caused considerable animosity in the community, leading to refusal to refer women. Providers especially resented the randomization process that failed to consider the women's diagnostic assessments.

After pregnant substance-using women are *found*, the challenge is to *engage* them into treatment. Women who are referred for an assessment often do not appear for appointments; women who are assessed may not enter treatment. In developing engagement strategies, it is important to

recognize the concerns of these women that may interfere with following through with assessment and treatment.

On their own, the women often are not able to recognize the multiple defense mechanisms of the disease of chemical dependence, so they continue to abuse alcohol and other drugs during pregnancy (Davis 1994). Some women have little or no contact with social services and health care agencies and are able to conceal their pregnancies until delivery. Other women fear the loss of their children to child protective services agencies (Streissguth and Giunta 1992, pp. 137-154). Many pregnant women fear prosecution if their drug use should come to the attention of any authority, especially the criminal justice system.

RECRUITMENT IN THE MOM'S PROJECT

By using results from focus groups of women in treatment, it was determined that reluctance to enter treatment had to do with denial of chemical dependence rather than concerns about research or randomization. Problems associated with participation in a clinical trial appeared to be related to community clinicians' attitudes and not the opinions of the pregnant women being recruited. Only 7 percent of recruited women refused their randomized assignments at enrollment.

Recruitment Strategies

To address recruitment challenges, public health nurses and social workers who were maternity case managers were selected as primary recruitment agents for two reasons: (1) to access referral from the parent agencies of these traditional social and health care services and (2) to capitalize on the history of these two disciplines in coordinating and linking service delivery, skills vital for working with the target population (Ridgely and Willenbring 1992, pp. 12-33).

Approximately 50 percent of case management time was spent on direct recruitment. Remaining time was spent on providing ongoing case management services to women enrolled in project treatment. Project case managers functioned in a manner similar to that of field-based caseworkers, a method that is the predecessor of many case management models (Schilling et al. 1988). The workers helped each woman identify her needs and the needs of her family and worked with her to ensure prenatal and postpartum medical care, maternity support services, and community social services such as safe housing, food, transportation, and child care. The provision of case management services was an additional incentive to a woman's ongoing participation in the project.

For example, one commodity in constant demand was safe, alcohol- and other drug-free affordable housing, which case managers helped locate for women participating in treatment.

Recruitment Sources

Early in the project, maternity case managers spent considerable time and energy identifying and "courting" personnel from traditional agencies for potential referrals. The list of agency recruitment sources for the project included 30 traditional agencies (e.g., jails, public health clinics, public assistance offices, child welfare agencies). The time was spent at agencies providing staff training, making presentations, observing characteristics of the client population, and maintaining personal contact with key personnel on a systematic basis. (However, referrals from some of these agencies failed to materialize regardless of the efforts of the project staff.)

The largest single referral source was women who self-referred (approximately 25 percent of all women who participated in the project). Many self-referred women most likely received initial information about the project from an established agency. Local public assistance offices, a constituent agency of the case manager social workers, referred consistently (about 19 percent). A most successful outreach effort was made to incarcerated women in the county jail facility. A social worker and public health nurse case management team visited the jail weekly to recruit new clients, which resulted in 16 percent of the approximately 19 percent referrals from the criminal justice system. Figure 1 shows a breakdown of recruitment sources.

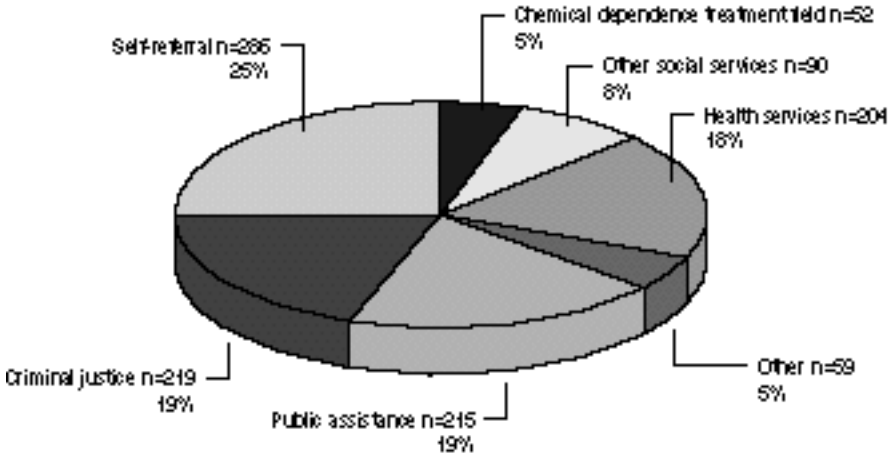


FIGURE 1. *Women recruited, by source (n=1,125)*

Client Characteristics

Women of color accounted for more than half of all referrals to the project (about 51 percent). More than a third of the referrals (approximately 35 percent) were African-American, about 9 percent were Native American, and slightly more than 3 percent were Hispanic. Caucasian women accounted for the other project referrals (approximately 49 percent). Figure 2 presents the ethnicities of the women recruited.

Case manager recruitment staff members quickly recognized that women who were eligible for the project presented problems as difficult as those of any clients with whom they had worked. Many of the women either lived on the streets or moved among friends on a regular basis. The women were "tuned in" to their addiction and used every resource available to satisfy it. Information from women who were eventually enrolled into the project (n=366) indicated that 89 percent were severely chemically dependent according to the *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised)* (American Psychiatric Association 1987). Seventy percent had at least one treatment experience before entering the MOM's Project, with almost one-third (32 percent) having participated in three or more prior treatment programs. A substantial majority (79 percent) reported having been incarcerated at least once, and 43 percent reported involvement with police before age 18. More than half the women with children (58 percent) had a history of involvement with child protective services.

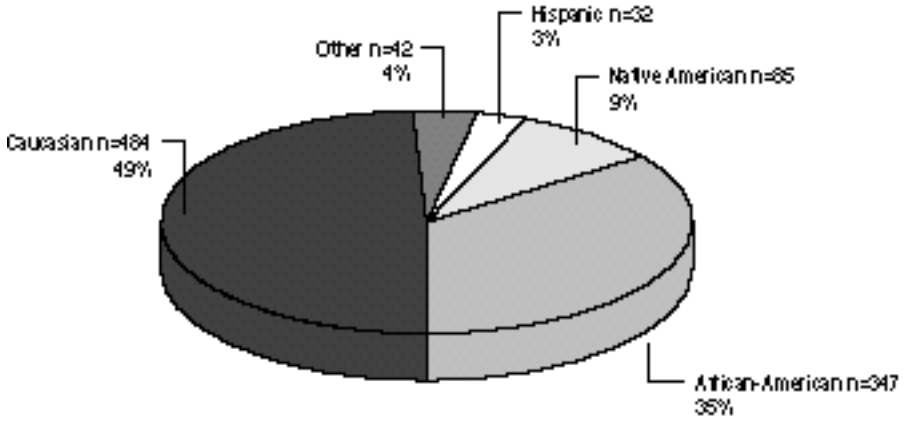


FIGURE 2. Women recruited, by ethnicity (n=990)

PROJECT RESPONSE TO RECRUITMENT COMPLICATIONS

To integrate the diverse agendas of recruitment agencies and other participating organizations, the project adopted a philosophy of recovery and emphasized the research context. MOM's Project treatment and community staff members adhered to a philosophy-of-recovery framework for working with the women in a multidisciplinary environment. In practice, this prevents the various treatment and support staff members from unintentionally interfering with the recovery process. This philosophy is resilient and focuses on solutions. Each woman's long-term goal is to reach her greatest potential as both an individual and a parent. The recovery process helps the woman evaluate herself by identifying her own strengths.

Under the recovery philosophy, the entire multidisciplinary staff practices giving positive feedback and role modeling. This allows clients to restructure their dysfunctional behavior patterns and practice positive interactions in a safe and empathetic environment. This philosophy led to the termination of individual project staff members who were unable to provide the demonstration services in a research context. Other project staff members were unable to prioritize recovery above individual agency agendas. The emphasis became *recruitment to chemical dependence treatment* as opposed to recruitment to other services being offered by the project, such as child care or prenatal care. This approach called for each project component to merge its agency and professional goals into the MOM's Project philosophy, with the focus on the client's long-term commitment to recovery. It became incumbent on the case managers, in their initial recruitment contacts with the women, to set the tone for this recovery milieu. The case managers focused on assisting clients to identify defense mechanisms that maintain denial, learn and reduce signs and symptoms of relapse, and reduce or eliminate destructive behavior patterns (Davis 1994).

To further minimize problems of diverse agendas, the project established standing committees to work on research- and demonstration-related issues and develop protocols and procedures for implementation. Each protocol and procedure was committed in writing in the *MOM's Project Implementation Manual* and the *MOM's Project Research Matrix* (LaFazia 1993*a*, 1993*b*). The implementation manual describes details for each demonstration component. The research matrix includes all research measures, timelines, and responsibilities for administering measures and tracking clients. These products were used to resolve conflicts, clarify roles, and serve as the final authority for each participating organization.

Midway through the project, a panel of outside experts was convened to assess recruitment efforts and provide recommendations for improvement. The panel included three researchers with at least 5 years of experience working with street-based drug users and two outreach workers who had worked with low-income pregnant women.

The panel members met with project staff members over several weeks and cooperatively identified the following barriers to increased enrollment:

- The parent agencies of the project's public health nurses and social worker case managers did not provide the anticipated number of referrals.

- Adolescents ages 16 and older were intended to be a major segment of the treatment population; however, few adolescents entered treatment.

- Recruitment as a priority appeared to be lost in a complicated maze of issues facing case managers and other project staff members.

- The potential for street-based and other more direct methods of recruitment had not been examined.

- Case managers made decisions concerning the focus of recruitment efforts without adequate information about the number of pregnant substance-using women an agency served.

- Case managers had limited information about agencies' attitudes toward treatment in general and the MOM's Project specifically.

- Case managers had varying and, in some cases, low levels of skill and knowledge about techniques for engaging substance-using women into treatment.

- Case managers appeared to be spending extensive time with some clients without engaging them into treatment.

Recruitment Enhancement

The project implemented the following five recruitment-specific activities: (1) A recruitment committee was formed to meet monthly to review recruitment activities and progress; (2) the recruitment protocol was enhanced to incorporate existing knowledge about treatment engagement techniques and limit the amount of time case managers spent on recruiting individual women; (3) a recruitment coordinator was hired to survey social services and health care providers, especially providers of services to

adolescents, assess the usefulness of direct recruitment, and otherwise strengthen the recruitment process; (4) other project staff members, especially those from the chemical dependence treatment program, were encouraged to become involved in recruitment efforts through providing ongoing presentations to health care and social services clinics and agencies; and (5) project staff members encouraged representatives of potential referral sources to visit the treatment facility.

Social Services and Health Care Providers Survey

A survey was conducted of the agencies on which the case managers were expending their recruitment efforts. The primary focus was on traditional health care and social services agencies. The survey focused on the time spent by case managers on those agencies and their referral rates. Discussions were held with agency personnel to assess the level of negative attitudes toward the project and to determine whether continued involvement would result in increased referrals. Based on the survey, it was determined that many of the traditional agencies targeted did not provide sufficient referrals to the project to warrant the time expended. Therefore, the outreach plan was revised, and the focus moved to nontraditional agencies (e.g., missions, needle-exchange locations, food banks). (Nontraditional agencies provide services to clients who do not access traditional social services and health care systems, such as very poor people and homeless substance-abusing people who live on the street. More than 25 nontraditional agencies were added to the agency recruitment list.)

The initial study design of the MOM's Project proposed that up to 40 percent of clients in the study would be pregnant adolescents. From the onset of the project, this population was difficult to recruit, engage, and enroll into the project. Although extensive effort was spent on adolescent services agencies, schools, and the juvenile justice system, adolescent referrals did not materialize. Youth care providers were surveyed to determine why adolescent referrals were not forthcoming and revealed the following: (1) Pregnant teens were hard to reach, and even providers who worked exclusively with teens had a hard time gaining access to them; (2) certain agencies receive supplemental funding to work with pregnant teens and thus were reluctant to refer teens to other programs; (3) teens, still early in their substance abuse, were better able to cease their alcohol and other drug use (at least temporarily) when they discovered that they were pregnant; and (4) agencies and teens did not identify substance abuse treatment as necessary.

Direct client recruitment often involves "working the streets." Street-based recruitment is more time consuming than agency recruitment.

Like most urban areas, Seattle-King County comprises a variety of geographic zones with differing population characteristics and drugs of choice. The characteristics of these zones can change over the course of a day (e.g., a downtown park frequented by homeless, chronic alcoholics during the day can become dominated by crack users at night), as well as over longer durations, and thus need to be monitored on a regular basis.

The recruitment coordinator spent time working the streets in each of the city's primary drug zones and in two major prostitution areas. The initial focus of these efforts was exclusively on pregnant women. Over time, the effort shifted to the goal of spreading the word about the project to men and nonpregnant women as well, in hopes that peer referrals would occur. These efforts succeeded in increasing project visibility and familiarity to street-based substance users. Time also was spent fielding requests for treatment by women ineligible for the project. Flyers routinely were placed in laundromats, beauty parlors, and neighborhood stores and on utility poles in neighborhoods with large concentrations of substance users.

The effort on direct client recruitment showed that such activity, although valuable for targeting women who are often invisible, was too costly in recruiting actual numbers. It was determined that time would be more effectively spent working with staff members from nontraditional agencies.

PROJECT RECRUITMENT PROTOCOL

By design, the project's designated recruitment teams of maternity case managers were not specialists in the substance abuse field. However, they received extensive training on substance abuse issues and shared client responsibility with the chemical dependence treatment staff. The enhanced recruitment protocol incorporated basic principles from field-based casework practice as well as from motivational interviewing for non-substance-abuse specialists (Rollnick and Bell 1991, pp. 203-213). A straightforward script was developed, giving prospective clients information and advice about the need to change their behavior. The case manager's role was to prepare women for a change in their drug lifestyle while emphasizing chemical dependence treatment as a necessary and available resource.

Although the project could not modify the recruitment criteria of early gestation and randomization, all the positive services available to the women were incorporated into the recruitment script. The benefits list was extensive and included referral to women-specific treatment for up to 12 months postpartum regardless of delivery status and custody (stillbirth, miscarriage, abortion, adoption, or foster care); opportunity for children

to be in residence with mothers; child care services while parents were in treatment; participation in research-related services that provided health care, social services, and child development assessments; gift certificates to a local department store for participation in research; assistance with financial and medical applications; and initiation of prenatal care through uninterrupted medicaid coverage (through 50 days after delivery for women at or below 185 percent of the Federal poverty level at the time of entry into care).

The recruitment script, including the benefits available, was presented in the context of a recovery approach. Rapport was established through gentle but firm confrontation with a woman regarding her drug-related issues, that is, other problems in her life that were connected to her substance use. The script included assuring the woman that chemical dependence is not *bad* or *wrong* but that it is a disease that progressively worsens and will continue to affect her life. She was reminded that regardless of her situation, she needed to start working on her recovery so that she could effectively handle the stresses in her life.

The enhanced recruitment script also established some time limits in working with each woman. The elusive behavior of some pregnant substance-using women often required an enormous amount of time to establish contact and build meaningful relationships. Initially, case managers spent considerable time working with a small number of resistant clients. Setting time and contact limits proved to be in the best interest of the chemically dependent pregnant women as well as beneficial to the case managers. When a woman is still enmeshed in denial, supportive persons tend to assume the role of enabler, which is counterproductive to recovery. The parameters for case management engagement activity included a limit of 1 month. If after a month the woman had been offered all the basic program services and still had not committed to a chemical dependence assessment, she was referred to the community for ongoing services, including referrals to community chemical dependence treatment agencies.

DISCUSSION AND CONCLUSIONS

The use of the public health nurse and social worker team as the project's primary recruiters proved to be effective in getting women into project or community treatment. The maternity case managers incorporated their expertise in working with difficult populations and their knowledge of community resources into a specialized advocacy role, working with women toward the primary goal of referral to chemical dependence treatment. Over the length of the project, more than 1,000 women were recruited, and 366 were enrolled in the project. Using maternity case

managers as recruiting agents was fiscally advantageous because, in Washington State, maternity outreach and case management is a billable medicaid expense. Programmatic advantages for replication stem from the availability of public health and public assistance social workers in every community throughout the State.

The largest age group recruited (approximately 34 percent) was 20 to 25 years of age; the second largest age group included women 26 to 30 years old (27 percent) (figure 3).

The majority of women recruited to the project were identified in either the first or second trimester of pregnancy (about 80 percent), with a large number recruited in the first trimester (approximately 38 percent) (figure 4).

In summer 1996 the MOM's Project will complete the evaluation of recruitment and admission to chemical dependence treatment (Lanz et al. 1995). Through the evolution of the project, it has become evident that case management may be an important link enabling women to successfully engage in treatment services. Further research is needed to explore the effect of case management services on the retention of pregnant substance-using women in treatment *and* on their transition back into the community.

The MOM's Project also is involved in a 3-year cost-effectiveness study funded by NIDA through the Boston-based Health Center for Economics Research. The study will examine the cost of various services received during pregnancy and after delivery for the project's treatment program. The cost-effectiveness of recruitment will be examined as part of the analysis.

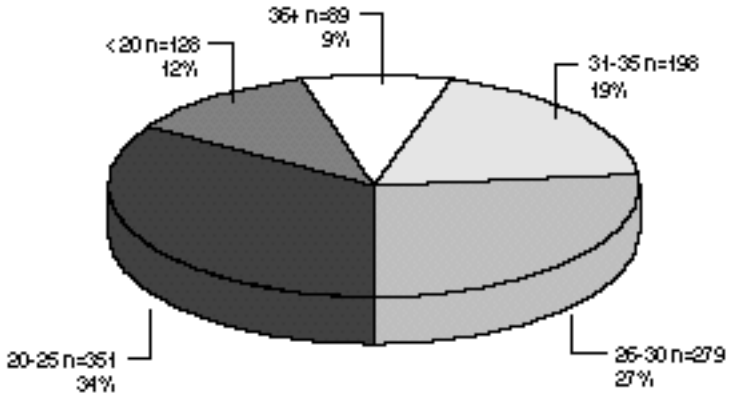


FIGURE 3. Women recruited, by age (n=1,045)

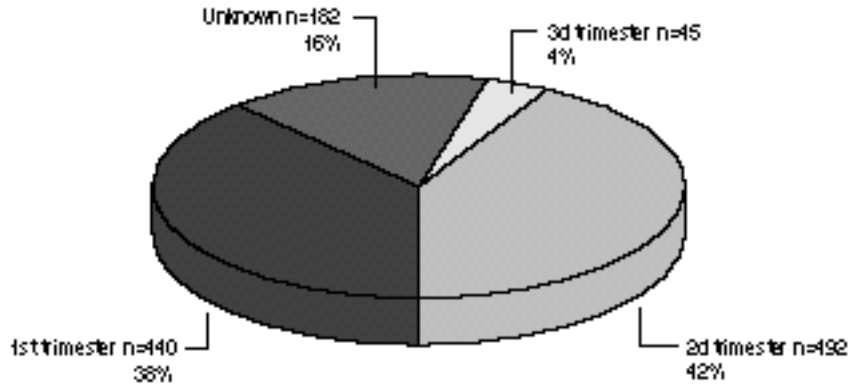


FIGURE 4. Women recruited, by trimester of pregnancy (n=1,159)

OUTREACH/RECRUITMENT/CASE MANAGEMENT VIGNETTES

The following are two descriptions of typical legal, social, and medical support advocacy efforts provided to substance-abusing pregnant women.

County Jail Outreach

Client Description. The client is a 22-year-old African-American woman, 5 months pregnant, and in jail at time of referral.

Client Needs. The client had no prenatal care prior to arrest, has no financial support, has no medical coverage, has been denied entry into the medicaid program because of her criminal justice status, and has a court hearing on criminal charges pending.

Outreach/Recruitment/Engagement. The referral came from jail health services. The case management team interviewed the client at the jail and facilitated resolution for medicaid coverage (through negotiations with the State office) and presented an argument to the court that the pregnant client and her 4-year-old child would best be served through treatment rather than further incarceration; the woman was enrolled for prenatal care.

Outcome. The client received an amended sentence, was randomized to a community substance abuse treatment program, participated in long-term residential treatment, delivered a healthy baby, returned to the community with her children, and remains in recovery.

Community Outreach

Client Description. The client is a 27-year-old pregnant Caucasian woman with two children, ages 5 and 6, who recently terminated a long-term violent relationship (including several hospitalizations). She is addicted to alcohol and cigarettes with occasional marijuana use, was raised in an alcoholic environment, has no high school diploma, has a low-wage employment history, and recently began a relationship with a nonabusive partner.

Client Needs. The client needed public assistance because of a breakup with an abusive partner, was served an eviction notice from her apartment (which cost $525 a month), was 5 months pregnant with no prenatal care and no extended family support, received threats from her ex-partner, and had no transportation to treatment and no child care for her children while in treatment.

Outreach/Recruitment/Engagement. The client was referred by a public assistance worker; she admitted chemical dependence problems but was overwhelmed with financial and housing crises and other problems that interfered with her getting into the treatment program. The focus was placed on activities to help her resolve her immediate housing and financial concerns so she could enter treatment. Case managers interceded with the landlord to help resolve the rental dispute, initiated referral to prenatal care, located child care so that the client could participate in treatment, provided counseling about domestic violence, and established a supportive role to assist her with other problems.

Outcome. The client entered MOM's Project outpatient treatment within 1 week of referral.

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised)*. Washington, DC: American Psychiatric Press, 1987. 567 pp.

Davis, S. Drug treatment decisions of chemically-dependent women. *Int J Addict* 29(10):1287-1304, 1994.

LaFazia, M., ed. *MOM's Project Implementation Manual.* Olympia, WA: Division of Alcohol and Substance Abuse, Department of Social and Health Services, 1993*a*. 82 pp.

LaFazia, M., ed. *MOM's Project Research Matrix.* (Revised). Olympia, WA: Division of Alcohol and Substance Abuse, Department of Social and Health Services, 1993*b*. 11 pp.

Lanz, J.; LaFazia, M.A.; and Hall, T. "Evaluation of Recruitment and Admission to Chemical Dependence Treatment." Unpublished paper, Department of Social and Health Sciences, University of Washington, 1995.

Ridgely, M.S., and Willenbring, M.L. Application of case management to drug abuse treatment: Overview of models and research issues. In: Ashery, R.S., ed. *Progress and Issues in Case Management.* National Institute on Drug Abuse Research Monograph 127. DHHS Pub. No. (ADM)92-1946. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

Rollnick, S., and Bell, A. Brief motivational interviewing for use by the nonspecialist. In: Miller, W.R., and Rollnick, S. *Motivational Interviewing: Preparing People To Change Addictive Behavior*. New York: Guilford Press, 1991.

Schilling, R.; Schinke, S.; and Weatherly, R. Service trends in a conservative era: Social workers rediscover the past. *Soc Work* 43(1):5-9, 1988.

Streissguth, A.P., and Giunta, C.T. Subject recruitment and retention for longitudinal research: Practical considerations for a nonintervention model. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children.* National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

ACKNOWLEDGMENTS

The services and research reported in this chapter were supported by NIDA grant R18-DA-06361 and by the State of Washington through the 1989 Omnibus Drug Act. A special acknowledgment is made to the pregnant women who overcame difficult obstacles to participate in the project.

AUTHORS

Mary Ann LaFazia, M.S.W.
Director
MOM's Project
Division of Alcohol and Substance Abuse
(360) 438-8058 (Tel)
(360) 438-8057 (Fax)
lafazm@dshs.wa.gov (Internet)

Kenneth D. Stark, M.Ed., M.B.A.
Director
Division of Alcohol and Substance Abuse
(360) 438-8200 (Tel)
(360) 438-8078 (Fax)
stark@dshs.wa.gov (Internet)

Department of Social and Health Services
P.O. Box 45330
Olympia, WA 98504-5330

Christopher Hansen
Supervisor
Family and Prevention Services
Division of Alcohol and Substance Abuse
Hansen and Associates
3840 85th Avenue, N.W.
Olympia, WA 98502
(360) 866-0721 (Tel)
(360) 866-0735 (Fax)
chansen@tssnet1.tss.net (Internet)

Jeanne Kleyn, Ph.D.
Recruitment Coordinator
MOM's Project
2216 Minor Avenue East
Seattle, WA 98102
(206) 285-4139 (Tel)
(206) 285-3532 (Fax)

Kris Nyrop, M.A.
Recruitment Coordinator
MOM's Project
Alcohol and Drug Abuse Institute
University of Washington
3937 15th Avenue, N.E.
Seattle, WA 98195

Jean Lanz, Ph.D.
Research Coordinator
MOM's Project
(206) 543-5557 (Tel)
(206) 543-3915 (Fax)
lanzjb@u.washington.edu (Internet)

D. Heather Watts, M.D.
Coinvestigator
MOM's Project
Professor
(206) 543-9867 (Tel)
(206) 543-3915 (Fax)
hwatts@u.washington.edu (Internet)

Department of Obstetrics and Gynecology
University of Washington, RH-20
Seattle, WA 98195

Teri Hall, M.S.W.
Case Manager Supervisor
MOM's Project
Economics and Medical Field Services
Department of Social and Health Services
3600 South Graham, Mail Stop N41
Seattle, WA 98118
(206) 721-2750 (Tel)
(206) 721-4729 (Fax)

Issues in Subject Recruitment and Retention With Pregnant and Parenting Substance-Abusing Women

Judy Howard and Leila Beckwith

INTRODUCTION

A substance abuse treatment program probably has its greatest effect on individuals who participate in all treatment sessions and who continue with the course of therapy until it is completed. Thus, efforts to promote compliance and retention are crucial to an effective study of the effect of substance abuse treatment (Mrazek and Haggerty 1994). Such efforts require knowledge about the characteristics of the target population's lifestyle and needs. To help researchers and clinicians develop strategies for decreasing noncompliance and attrition, this chapter provides examples of how the lifestyle and needs of addicted women with children affect recruitment and retention in treatment research. Pregnant women and women of childbearing age present an array of real-life circumstances that challenge traditional models of substance abuse treatment as well as traditional means of recruiting and retaining subjects in research programs.

This chapter reviews a range of subject recruitment and retention issues specific to pregnant women and women of childbearing age who use illicit substances. Recruiting and retaining study participants overall, in the drug treatment field (Gilchrist and Gillmore 1992, pp. 1-17; Hansen et al. 1990; Stark 1992) and in many other areas of health research, pose difficult issues. For example, in a recent review of treatment research to reduce mental disorders, methodological problems recurred in a variety of studies (Mrazek and Haggerty 1994). The most frequently identified problems related to the difficulty of adhering to a strict randomized trial design and to high attrition among study participants. When subjects are substance-abusing women with children, these same problems commonly arise, but they often are made more complex because they occur in combination with a variety of other unique issues that relate specifically to this population.

SPECIAL ISSUES AFFECTING RESEARCH ON PREGNANT AND PARENTING SUBSTANCE-ABUSING WOMEN

Treating addiction is the core purpose of the Perinatal-20 Treatment Research Demonstration Program. Addiction arises out of multiple adverse circumstances and in turn sets the stage for further difficulties. Within the

populations studied by the majority of the Perinatal-20 projects, a variety of clinical issues were identified as complicating factors in research with this population, many of which stemmed from subjects' family histories and life circumstances. The majority of the addicted women recruited and enrolled in these projects had experienced harsh, abusive childhoods and had been reared by parents who were alcoholics or substance abusers and were unable to provide consistent care for their children. Moreover, the addicted women had experienced problems related to school achievement and employment. They also reported difficulty in establishing stable, supportive relationships with friends and significant others. During the time they participated in the research projects, most subjects were single parents living in poverty with several children, and they were frequently involved with the legal system. Their limited economic and psychological resources produced a broad spectrum of needs in addition to those associated with addiction.

Within this subject population, investigators delineated seven clinical factors that directly contributed to the extraordinary staff efforts needed to recruit and retain subjects in treatment research: (1) addiction severity level, (2) involvement with the legal system, (3) housing problems, (4) difficulties with interpersonal relationships, (5) parenting responsibilities, (6) employment-related issues, and (7) the need for many comprehensive services. Although these variables often cannot be controlled, they have a profound effect on subject retention and attrition and thus on the success of research evaluations. Accordingly, it is important to identify and code these factors to enable researchers to determine those that differentiate subjects who remain in treatment from those who do not (Reed and Grant 1990, pp. 10-56). Such comprehensive information also can help inform future research efforts in defining specific study samples and effective treatment components.

Addiction Severity Level

Addiction is a chronic, relapsing disease that in its most blatant form suppresses other personal goals and becomes the driving force that determines all activities in an addicted individual's life. Within this context, remaining in treatment and becoming abstinent occurs for only a minority of enrolled subjects during a defined time. Because of the way various research investigators view addiction, they approach subject retention in different ways. Some define it by both attendance in the program and abstinence, whereas others define it on the basis of attendance alone. Furthermore, some researchers consider abstinence as not using illegal drugs, whereas others feel that the use of alcohol constitutes a relapse. In still other cases, retention is defined as a subject's continuing willingness to participate in at least some components of the treatment program.

Strict retention guidelines that require both regular participation in the treatment program and abstinence or sobriety may result in the loss, dropout, or attrition of a majority of enrolled subjects. However, such strict guidelines also may better demonstrate the effectiveness of the treatment program by showing significant differences in outcome measures among study group subjects who are retained, subjects in the control group, and study group subjects who drop out.

On the other hand, the application of less rigorous retention guidelines also benefits research by adding new knowledge to the field, thus contributing to the development of appropriate services to treat the majority of women who form the attrition group and allowing investigators to trace patterns of relapse and recovery that can inform new treatment paradigms using different outcome measures. For example, there is a growing recognition in the field that addicted clients may need pretreatment services to help them work through denial and facilitate their commitment to abstinence. Development of pretreatment services and evaluation of their effectiveness may result from less restrictive retention guidelines.

Involvement With the Legal System

Because the use of illegal drugs involves illegal activities, many subjects are involved with the legal system, within either the criminal or civil courts. For example, among subjects who are involved with the criminal justice system, extended periods of incarceration often interfere with retention in a research study. In the authors' experience, a subject who was highly motivated to stop using cocaine attended treatment sessions regularly and remained abstinent during the last 4 months of her pregnancy; however, she relapsed just prior to delivery, with subsequent positive urine toxicology screens for herself and her infant at delivery. In spite of the positive toxicology screens, she was awarded custody of her infant because of her involvement in the treatment program. However, when the subject was discharged from the hospital, the maternal grandmother refused to allow her daughter and grandchild to live with her. Following this event, despite staff members' efforts to obtain housing, the subject resisted treatment services, eventually returning to the streets and addiction. She was arrested on drug-related charges, lost custody of her infant, and was incarcerated for 4 months. Thus, although intensive efforts on the part of clinical staff members had resulted in 4 months of abstinence for this subject, she ultimately was dropped from the study.

On the other hand, brief periods of incarceration do not necessarily end a subject's participation in a research program. However, such legal involvement can place extraordinary demands on clinical staff members who are attempting to decrease subject attrition. In another example from

the authors' project, an enrolled subject who experienced several brief periods of incarceration repeatedly called staff members from jail to express her desire to continue with substance abuse treatment and receive additional assistance from program personnel. The clinical staff members responded to her telephone calls from jail by making preparations for housing and treatment on her release. Invariably, when released from jail, the subject not only resumed program attendance but also returned to her familiar neighborhood and continued with her illegal activities, which resulted in repeated incarcerations. Thus, despite her stated willingness and attempts at attendance, after several months it became clear that this subject could not participate in treatment over the long term, and she was dropped from the study.

Another issue related to subject involvement with the legal system is child custody and mandated treatment. Not infrequently, addicted pregnant women or women who have lost custody of their children enroll in drug treatment programs to demonstrate to the court that they are trying to become abstinent. Furthermore, the court may mandate such participation for these women to maintain or obtain custody of their children. Sometimes, once the court makes its ruling regarding child custody, the subject drops out of the study because of lack of motivation, thus increasing attrition. For example, loss of child custody may precipitate a mother's decision to give up and not return to treatment. On the other hand, a court decision to award custody also may remove the motivation for undergoing treatment.

Housing Problems

It is critical for a subject to have a stable residence so that staff members can locate and contact her. Yet the housing problems commonly experienced by low-income substance-abusing mothers can place considerable demands on clinical staff members who are trying to maintain subjects in a research program. For example, in one study conducted by the authors, a subject was unable to continue making rent payments and was evicted from her apartment. She felt intimidated and was unable to secure a place by herself; she asked staff members to help her locate housing, make telephone calls to get information about rentals, and transport her and her infant to view prospective residences. In addition, she needed help in filling out rental application forms, acquiring furniture, and setting up the household. Similarly, clinical staff members assisted another subject in obtaining housing by reviewing advertisements with her and taking her to various neighborhoods to help her decide where she could live. In both these instances, staff members' efforts were rewarded, and the subjects were retained.

Difficulties With Interpersonal Relationships

Subjects' relationships with their significant others are an important part of their lives and therefore also affect their participation in research studies. Within the population reported here, these relationships often are marked by instability, discord, and in some cases, violence. For example, one subject was involved in a violent episode with her significant other during which, under the influence of cocaine, he broke down her front door and demanded money. Fearing for her own and her children's lives as he battered her, she stabbed him. The police were summoned, and the man was incarcerated. Following this incident, the subject required extensive support and one-on-one counseling from clinical staff members to recover from the emotional trauma, develop strategies to prevent its recurrence, and continue her participation in the study.

In another situation, staff members invested months of effort attempting to address the relationship problems between a subject and the father of her children. To retain her in the study, clinical staff members worked extensively with the couple, making numerous home visits and conducting couples counseling at the program site. Although the mother was the identified subject, her significant other used the program hotline extensively during all hours of the day and night because the subject frequently left him with the children when she was bingeing. Despite these efforts, maintaining the subject in this relationship failed, and staff members secured a placement in a homeless shelter for her and her children. Her attendance in the program increased after this placement, but because of the lack of social and family supports, she eventually decided to move to another area to be with her mother and was subsequently lost to the study.

Clinical services for this population often must be directed not only toward the subject and her significant other but also toward the subject's family of origin. Not infrequently, grandmothers provide housing and money to enrolled subjects and their children, and such enmeshment in this extended relationship can make unusual demands on the clinical staff. For example, after one subject was enrolled in the study, staff members became aware that she was violent and abusive toward her own mother and siblings. After staff members had worked closely with her and her family to obtain a psychiatric assessment as well as placement in an inpatient mental health facility, the subject disappeared. Although she was dropped from the study, her mother continued to seek assistance and counsel from the clinical staff in caring for her grandchild.

Parenting Responsibilities

Parenting responsibilities also present obstacles to attendance and compliance with drug treatment. In one instance, a subject was making good progress in recovery after participating for 6 months in a drug treatment program with her infant when, despite her wishes, the legal system returned her four older children to her custody. She had never had custody of all her children before. In this case, family reunification resulted in the subject's dropping out of the center-based day-treatment program and increasing demands on the clinical staff members, who helped this mother locate a larger apartment, apply for Aid to Families With Dependent Children, identify local schools for her children, and organize afterschool activities. Staff members also had to help her develop daily routines to shop, cook, and do laundry for her newly enlarged family, as well as provide inhome counseling to help her deal with the older children's behavioral issues. Because of these increased parenting responsibilities, the mother was no longer able to participate in the center-based program and had to be dropped from the study.

Thus, family reunification, which for many clients is a marker of success in a treatment program, also can result in subject attrition. Many subjects in the authors' attrition group were unable to coordinate their attendance at the day-treatment program with their responsibilities in caring for their children after kindergarten and elementary school. Although the program provided transportation and child care for one pre-school-age sibling beyond the focus (target in the research program) child, there were not enough resources to provide comprehensive child care (afterschool care) for the enrolled subjects' offspring. (These mothers had an average of three children of varying ages.)

However, for other subjects, clinical staff members' support of parenting responsibilities, when family reunification occurred, resulted in the mothers' recognition of the help that the program could provide, not only in maintaining abstinence but also in providing for the day-to-day needs of their children. For example, one subject disappeared from the authors' treatment program when she became overwhelmed with the responsibilities of caring for her new infant and 9-year-old son, who was reunited with her after having been in out-of-home placement for several years. Following family reunification, the 9-year-old exhibited a variety of behavioral problems at home, was absent from school for weeks on end, and acted out when he was in class. On the basis of recommendations made by program staff members after they lost contact with this subject, child protective services took custody of both children until the mother resumed her participation in the treatment program and demonstrated efforts at maintaining abstinence. Once this occurred, clinical staff members helped

this subject reach out to her own mother for support; the grandmother in turn assisted the mother as she planned for the safety and care of her children. This added support resulted in the subject's being retained in the study.

In summary, knowledge to date does not enable investigators in this field to determine which subjects will be able to work on recovery and parenting responsibilities simultaneously. Yet this is a critical issue in retaining parenting women in substance abuse treatment.

Employment-Related Issues

During the course of their participation in a research program, some parenting women who are receiving substance abuse treatment become legally employed. Generally, this is viewed as a measure of the subject's recovery, and treatment programs can count the subject's employment as a measure of success. However, within a research design that involves set times and days for treatment, employment schedules often conflict with the treatment protocol and thus push subjects into the attrition group. Nonetheless, some women are able to combine both treatment and employment. For example, one subject in the authors' project became employed as a home health care aide; her sufficiently flexible work schedule allowed her to perform her job duties (bathing and caring for elderly people in their homes) as well as continue her attendance in the research program.

Need for Many Comprehensive Services

Research demonstration projects that investigate treatment approaches for substance-abusing mothers require an extraordinary range of service components to promote subject enrollment and retention. When treatment efforts are directed toward pregnant women or women of childbearing age, basic parenting-related needs (e.g., child care; physical space for infants' and young children's sleep, eating, and play activities; transportation to day treatment centers) must be met if substance-abusing mothers are to be able to address their addiction through participation in appropriate treatment. Furthermore, in addition to having drug treatment counselors on staff, programs need to employ staff members who can provide child care while mothers are participating in substance abuse treatment sessions. Although it is physically possible for mothers to care for their own children during drug treatment sessions, it has been the authors' experience that this added responsibility increases mothers' stress and interferes with their ability to focus on critical issues of addiction and recovery. In addition, a child care component should encompass food, formula, child car seats, diaper bags,

clothing, toys, kitchen equipment, and appropriate furniture (including floormats, cribs, child-size tables and chairs, and potty chairs). Moreover, if parenting and health education are incorporated into the treatment design, additional space must be provided and staff members must be hired to provide these service components as well. Finally, the various activities of these interdisciplinary staff members need to be coordinated in an organized, integrated, and meaningful way to help subjects benefit more fully from program participation.

These types of service components are expensive and thus necessitate a small number of subjects within a study sample. However, the costs of providing such treatment for pregnant women may be offset by the reduced costs to society related to decreased perinatal complications for newborns during and after delivery and fewer inhospital days. For example, Lee and Svikis (1995, p. 482) compared health costs for the delivery of infants born to cocaine-abusing women who received drug treatment with those of cocaine-abusing women who did not receive drug treatment during pregnancy. Neonatal intensive care unit (NICU) use rates decreased by nearly 50 percent for infants whose mothers had been enrolled in drug treatment. The average length of stay in the NICU also decreased—6.5 days ($9,750) for infants of mothers in the treatment group vs. 41 days ($61,500) for infants of mothers who did not receive treatment. The average cost for drug treatment during pregnancy was $6,700 per mother.

Despite this evident cost-effectiveness, one strategy for reducing the number of research dollars required for research demonstration projects is to fund discrete studies collocated within ongoing treatment programs. This approach can involve hazards if the continuing treatment program and the research team have not established mutual goals, mutual trust, and a similar philosophy of treatment. On the other hand, if there is close collaboration between the research and clinical teams, this solution for decreasing costs per study may be effective, thus enabling more research to be conducted.

SPECIFIC ISSUES AFFECTING RECRUITMENT

Samples of Convenience

As in any field of study, it is useful in addiction research for investigators to examine recruited samples that span the full range of the disorder (Mrazek and Haggerty 1994). If, for various reasons, samples are biased—as they were among the majority of the Perinatal-20 projects, which

included only low-income, chronic, heavy users in their midtwenties to late twenties—then study results cannot be generalized across the broader population of addicted women of childbearing age (Gorelick 1992).

Because almost all Perinatal-20 study samples were recruited on the basis of referrals from public health or social services agencies, they are biased but convenient samples. This method of recruitment allowed investigators to recruit quickly from populations that have a visibly high rate of addiction. Problems related to addiction were already fairly certain among referred subjects because public health and social services agencies have been making efforts to identify addiction among pregnant women or women with children throughout the past decade.

Despite these convenient samples, after referral the research teams had to follow their own protocols for inclusion criteria, which incorporated a toxicology screen (e.g., urinalysis, analysis of meconium, hair analysis, self-report). Although the decision to use one or more of these specific criteria was simple, implementation of screening procedures was more or less difficult, depending on the referred sample. Recruiting from large county hospitals or child protective services agencies—rather than from many small, private sector providers, for example—yielded a higher rate of return for staff time invested and resulted in a lower subject recruitment cost. Conversely, recruiting from the private sector forced investigators to solicit across a larger number of agencies and to educate staff members, monitor testing, and in some cases, institute toxicology screens and interviews within the offices of a larger number of referral sources—all with a potentially low rate of return.

Although diversity of samples within a field of study is valuable, within an individual research demonstration project, homogeneity of client characteristics is a necessary requirement to control for intrinsic variability, which would swamp the observed effects of treatment. For instance, current research contains a paucity of information about the effectiveness of treatment for mothers who are just beginning to use drugs. Furthermore, little is known about the extent and severity of addiction among adolescents and young women or about effective treatment for these groups. Moreover, treatment studies of women who have greater economic and psychological resources are lacking as well. To remedy limitations related to recruitment of convenient samples and in recognition of the importance of studying a broader range of substance abusers while still constituting homogeneous study groups, future requests for proposals should encourage and support researchers in identifying and recruiting subjects from more diverse segments of the population. Unless this effort is made, future investigators are likely to continue to study only the limited, convenient sample that has been most widely researched to date.

Maintaining a Strict Randomized Design

A strict randomized trial design that includes a formal retention protocol and a designated duration of treatment also is necessary to ensure study validity. It is also necessary to describe how many participants were recruited, how many were screened, how many passed and how many failed the screening, how many consented to participate in the study, and why refusals occurred. After enrollment, it is important to assess how many subjects entered their assigned randomized groups and why others did not. After group assignment, it is critical to determine how many subjects completed the program and why others did not. Finally, researchers also should examine the baseline factors that were linked to dropout and determine whether they were the same for the experimental and comparison groups (Mrazek and Haggerty 1994).

Such stringent requirements may challenge the morale and cooperation of researchers and clinicians who are working together in treatment research (Howard et al. 1990, pp. 66-79). Thus, to ensure a successful project, both groups must be informed about the research plan and objectives and be enthusiastic about and committed to study goals. They also must recognize the necessity of adhering to the research design and the consequences of deviating from it (Mrazek and Haggerty 1994). However, even when such initial agreement is present, there commonly exists a basic tension between research requirements and clinical services and needs (Sacks 1983). Although it is difficult to lessen these types of tensions, it is imperative that investigators consider these problems and make efforts to prevent or contain them to maintain a study design that will enable an evaluation of treatment effectiveness. Even when research investigators are not able to control external forces, these factors at least should be recorded to inform future studies. There are at least three levels on which such tensions may occur.

Clinical Staff. Substance abuse treatment research is no different from research in other clinical fields, insofar as conflicts often emerge between the clinical staff and the research design. These conflicts may arise even before data collection is initiated, when an investigator is trying to select an appropriate individual to conduct recruitment—a task that is crucial to the success of any study. On the one hand, a researcher-recruiter understands the need to fill both the treatment and control groups equally but may not possess sufficient clinical acumen to interest subjects in participating. On the other hand, a clinician-recruiter may be able to increase the participation rate but may resist enrolling subjects in the control group. In the authors' Perinatal-20 project, for example, the decision was made to employ a drug treatment clinician-recruiter.

However, because of ambivalence about the research design, this individual impeded recruitment for both treatment and control subjects. Her growing frustration about her lack of control over random client assignment to the experimental or community comparison group and her emerging belief that experimental subjects received more effective and appropriate services eventually resulted in her hesitancy about recruiting subjects for fear that they might be assigned to the control group.

Another conflict between the research design and clinical staff may arise when clinical staff members' perceptions about good clinical care run counter to the research treatment protocol. For example, one pregnant, cocaine-addicted woman who continually expressed interest in participating in a study involving day treatment attended sporadically and did not keep her prenatal obstetric appointments. When she delivered her infant, who tested positive for cocaine metabolites at birth, clinical staff members felt that this subject could no longer benefit from the program's services and referred her and her newborn to a residential drug treatment program in the community. After this woman had remained abstinent for 6 months, one of her three older children also joined her and her infant in this residential treatment setting. However, despite this positive outcome, which occurred as a result of the efforts of the study's clinical staff members, for research purposes this subject had to be counted as a member of the attrition group. In another case, also based on the perceptions about appropriate clinical care of this same clinical team, another client, who had a similar pattern of noncompliance in the day treatment program, was referred to another community-based residential program. After 6 weeks, this subject left the residence without permission and was terminated from the residential service. She recontacted the day treatment staff, admitted to living in a crack house and using crack cocaine daily, and stated that she was not interested in a second referral for residential treatment. She told the clinical staff, "I do not need a drug program. What I need is a place to live with my kids, and then I'll be able to stop using drugs on my own." Although the therapeutic outcome for this subject was not successful, once again research rules required that she be included as a member of the study attrition group. In most cases substance abuse treatment professionals are not yet able to determine precisely which subjects will benefit from specific treatment modalities (e.g., day treatment, residential treatment, transitional housing, therapeutic community, etc.). Thus, clinical staff members must be aware of what is known and what is not yet known in the substance abuse treatment field so that the research design is not unnecessarily compromised (Price and D'Aunno 1992, pp. 37-60).

Statistical considerations within the research protocol also may cause dissension between the research design and clinical staff members.

For example, in the authors' project that had a designated treatment period of 18 months, some subjects delivered two children within that interval. Although, for purposes of research, only the first child was considered the target child, clinical staff members felt an ethical responsibility to focus as much effort on the new babies as on the designated target children. Clinical staff members and mothers alike pressured research staff members to conduct research assessments of these younger siblings, which restricted the research staff members because of the extra time and money required to perform these additional evaluations. Furthermore, the data related to these second children could not be included in statistical analyses because of the need to exclude correlated data.

Referring Community Professionals. When the research design includes a control or comparison group, tensions may develop between the program and community professionals who refer subjects to the study. Because community professionals want to secure the best possible treatment for people in need, they may resist referring clients to a research program where there is random assignment to a control or comparison group. In addition, referring parties may try to pressure clinical staff members or negotiate services for those subjects who are assigned to the control or comparison group. These circumstances can undermine clinical staff members' confidence in the research program and can place stress on staff morale if it calls into doubt the community's regard for the value of the program with which staff members are associated.

Study Subjects. Some study subjects also experience tensions associated with a research design that involves random assignment. For example, experimental group subjects who value the intervention services they are receiving commonly want to refer friends to the program but only if they can be assigned to the intervention group. If random assignment places a friend in the control or comparison group, the referring subject may pressure clinical staff members to provide additional services. For example, in the authors' experience, where groups comprise women randomly assigned to residential treatment programs with and without their children, mothers assigned to the group that did not include children expressed guilt and concern about how the study may have interfered with their maternal responsibilities. Likewise, within a program where residential treatment was compared with outpatient or day treatment, tensions developed when women did not like their assignments and requested to transfer to the other treatment option.

Resentment or ambivalence about prescribed treatment in a population that over the past decade has become more knowledgeable about treatment options can contribute to attrition. Attrition represents an additional potential bias because subjects who remain in a study may

differ generically from those who do not. In one Perinatal-20 project, the majority of subjects who stayed in treatment had reported more psychological distress and less independence at the onset of their participation in the study than the women who rejected treatment (D. Haller and S. Schnoll, personal communication, June 16, 1994), whereas other studies have found different factors to be associated with retention (Gainey et al. 1993; Kleinman et al. 1992; Stark and Campbell 1988; Williams and Roberts 1991). These findings emphasize the importance of including sufficient and appropriate baseline measures—such as measures of intelligence, personality, and social supports—to enable researchers to determine potential biases in samples related to retention and attrition (Mrazek and Haggerty 1994).

ETHICAL ISSUES

Ethical concerns, which are inherent in conducting treatment research with substance-abusing populations, arise out of many of the issues described above. First, research to date has not clarified specific types of treatment approaches that are effective in meeting specific client needs. Investigators can be caught between this lack of empirical knowledge and clinical staff members' perceptions about subjects' treatment needs, particularly when staff members feel that a subject requires interventions that lie beyond the research parameters. For instance, clinical staff members who provide day treatment may perceive that a subject's addiction is so out of control that, in their clinical judgment, she requires residential treatment to separate her from a high-risk environment. Such anxiety on the part of clinical staff members then poses an ethical issue for the principal investigator, who must weigh the need to retain subjects and ensure the integrity of the project vs. the risk of providing insufficient or inappropriate treatment, when current knowledge provides no clear guidelines regarding this treatment option.

A second ethical question relates to the termination of pregnant women from treatment studies because of relapse. Even during periods of noncompliance, a woman's enrollment in substance abuse treatment may have a mitigating effect on her level of drug use. Because of concerns about a possible association between higher levels of drug use and premature labor and delivery, staff members may have strong concerns about the consequences of a subject's potential escalated drug use once she is terminated from a study and finds herself in an unsupervised situation and without the medical and supportive services that the program provides. From this point of view, although the study design may consider only the woman to be the subject of the treatment research, in reality she and her fetus cannot be differentiated as a treatment unit. The resulting ethical

dilemma relates to maintaining the integrity of the research retention protocol (i.e., adhering to termination guidelines) vs. potential damage to the fetus, who is at high risk for preterm delivery.

A third ethical issue relates to the role of the investigator when a subject is involved with the legal system because of child abuse and neglect. For many such subjects, the civil court and child protective services agencies determine whether the mothers retain custody of their children. Frequently, the court may mandate participation in a drug treatment program as a condition for either obtaining or maintaining custody. However, obeying the court order to obtain custody may conflict with a mother's participation in a research demonstration treatment study. Two situations highlight this point. First, if the court stipulates drug treatment as one of the conditions for child custody, random assignment to a control group that does not receive treatment in cases where drug abuse treatment is part of the study design would violate this mandate. Second, if a research project offers residential treatment with random assignment to groups that do and do not include children, and the judge awards child custody to a mother on the condition that she enroll in drug treatment, the subject faces a dilemma if she is randomly assigned to the treatment group that does not include children. How does she comply with the court order to enter treatment and maintain custody when random assignment may place her in a situation where she can receive treatment but is not allowed to live with her children while obtaining treatment?

A related but separate situation may arise when a mother does not have custody of her children on entrance into a research demonstration study, but later, when she begins to recover from her addiction, the court is pleased with her progress and returns the children to her custody. In such cases, a subject not only may feel overwhelmed by her struggle to maintain abstinence (although she is successful to date) but also may be reluctant to tell the court that she is not yet ready to assume the added responsibility of day-to-day parenting for fear of permanently losing custody of her children. What priority is placed on the mother's efforts to become well? At what point during a subject's recovery is she ready to successfully take on the responsibilities of parenthood? Furthermore, should research investigators attempt to negotiate with the courts to foster subject retention and support staff members' efforts to provide effective substance abuse treatment?

These are some ethical concerns faced by investigators involved with this population. Future studies may be able to circumvent some of these difficulties by carefully considering them in advance, deciding on practical solutions, and addressing these issues in the consent form signed by each research subject at the time of enrollment. For example, such statements might include the following:

- Termination from the project may occur if the treatment staff members perceive that another form of treatment will be more helpful to you. If this occurs, staff members will recommend programs that offer the preferred method of treatment.

- Relapse may be a condition for termination from the study. If you are pregnant at the time of relapse, staff members will refer you to a health care facility that has agreed in advance to provide necessary medical services to promote the health of you and your baby. Staff members also will refer you to other available drug treatment programs.

- Your participation in this substance abuse treatment program is separate from any involvement that you may have with the legal system regarding your addiction. Program staff members will not communicate with the court about your drug use or your recovery from addiction. If the court requires reports about your drug use, we will refer you to other drug treatment programs. However, if you agree to participate in our program, we *will* communicate with the legal system about any incidence of suspected child abuse or neglect, as required by State law.

  Or

- Besides reporting any incidences of suspected child abuse or neglect, program staff members will comply with court requests for information about your substance abuse and your efforts at recovery. As required, staff members also may report such information at custody hearings. However, the court will base its decisions on all the information it has regarding your case, and not on our reports alone.

SUMMARY AND CONCLUSIONS

To advance knowledge about the treatment of addiction among pregnant women and other women of childbearing age, investigators must adhere to the requirements of a strict experimental research design while concurrently providing clinical services. This means that researchers must address a variety of difficult questions, including the following:

- Was the sample large enough?

- Were the criteria for subject inclusion and exclusion well defined?

- Did the process of recruitment result in a sample that could be generalized to a larger population, or was the sample biased in some way?

- Was assignment to groups clearly random?
- What was the attrition rate?
- Was attrition the same in both experimental and comparison groups?
- Did baseline measures collect enough information to permit a description of the factors that were associated with attrition in each group?
- Was the attrition rate so high that the retained sample had special characteristics? If so, what were these features?

This chapter highlights several problems related to these questions, describes the difficulties that investigators have faced in meeting clinical and research challenges to date, and suggests strategies for overcoming some obstacles.

In establishing the Perinatal-20 project, the National Institute on Drug Abuse took an informed first step in organizing a substantial research effort to investigate treatment modalities that incorporate services specific to the needs of substance-abusing women who have children. This initial effort has resulted in a beginning knowledge base that can be used to refine and expand future treatment efforts. Even the issue of the "study unit" for this population is evolving. Today's researchers are attempting to determine whether the mother alone or the mother along with her dependent children constitutes the study unit. This question also has led professionals in the field to examine a range of specific outcome priorities, and investigators just now are beginning to determine exactly what needs to be evaluated in gauging the effectiveness of treatment. Is success measured on the basis of the woman's progress with abstinence alone, or does it also include her role with her children? Is it determined on the basis of her relationship with her children or the children's growth and development? Compared with providing services for and studying single adult subjects, developing treatment for women and their children presents researchers with a more complex task and requires expanded clinical services (Gallagher 1990, pp. 540-559).

As in most fields of study, initial research data in substance abuse treatment for pregnant and parenting women are derived from samples of convenience, as described above. To put this information in perspective, future research will require a wider and more representative spectrum of the population. Furthermore, tensions between clinical needs and research requirements must be considered in advance, and methods for relaxing these tensions will be critical to the success of future efforts. For example,

members of both the research and clinical staff teams must be absolutely clear about the study design and the requirements of reliable research. Where possible, potential ambiguities about group assignment, project services, subjects' responsibilities, and so forth must be incorporated into subject consent forms so that the subjects also are apprised of potential problems and their solutions. A final caution to future investigators is to be aware of the economic, physical, and personnel limitations of the range of treatment services that can be provided in a research demonstration study involving this population. Because of these limitations and the extensive range of services the subjects of the studies require, treatment components must be discrete and carefully defined to prevent programs from becoming impractically diverse and unclear. Research goals must be attainable and measurable. Finally, researchers must not underestimate the contribution that a small but well-designed and well-described study can make to this developing field.

REFERENCES

Gainey, R.R.; Wells, E.A.; Hawkins, J.D.; and Catalano, R.F. Predicting treatment retention among cocaine users. *Int J Addict* 28(6):487-505, 1993.

Gallagher, J.J. The family as a focus for intervention. In: Meisels, S.J., and Shonkoff, J.P., eds. *Handbook of Early Childhood Intervention*. New York: Cambridge University Press, 1990. 760 pp.

Gilchrist, L.D., and Gillmore, M.R. Methodological issues in prevention research on drug use and pregnancy. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children*. National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

Gorelick, D.A. Sociodemographic factors in drug abuse treatment. *J Health Care Poor Underserved* 3(1):49-58, 1992.

Hansen, W.B.; Tobler, N.S.; and Graham, J.W. Attrition in substance abuse prevention research: A meta-analysis of 85 longitudinally followed cohorts. *Eval Rev* 14(6):677-685, 1990.

Howard, K.I.; Cox, W.M.; and Saunders, S.M. Attrition in substance abuse comparative treatment research: The illusion of randomization. In: Onken, L.S., and Blaine, J.D., eds. *Psychotherapy and Counseling in the Treatment of Drug Abuse.* National Institute on Drug Abuse Research Monograph 104. DHHS Pub. No. (ADM)91-1722. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Kleinman, P.H.; Kang, S.Y.; Lipton, D.S.; Woody, G.E.; Kemp, J.; and Millman, R.B. Retention of cocaine abusers in outpatient psychotherapy. *Am J Drug Alcohol Abuse* 18(1):29-43, 1992.

Lee, J.H., and Svikis, D. Cost-effectiveness of comprehensive care for drug-abusing pregnant women. In: Harris, L.S., ed. *Problems of Drug Dependence 1994: Proceedings of the 56th Annual Scientific Meeting, The College on Problems of Drug Dependence, Inc., Volume II.* National Insitute on Drug Abuse Research Monograph 153. NIH Pub. No. 95-3883. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1995.

Mrazek, P.J., and Haggerty, R.J., eds. *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Institute of Medicine. Washington, DC: National Academy Press, 1994. 605 pp.

Price, R.H., and D'Aunno, T.A. The organization and impact of outpatient drug abuse treatment services. In: Watson, R.R., ed. *Drug Abuse Treatment. Vol. 4. Drug and Alcohol Abuse Reviews.* Totowa, NJ: Humana Press, 1992.

Reed, R.J., and Grant, I. The long-term neurobehavioral consequences of substance abuse: Conceptual and methodological challenges for future research. In: Spencer, J.W., and Boren, J.J., eds. *Residual Effects of Abused Drugs on Behavior*. National Institute on Drug Abuse Research Monograph 101. DHHS Pub. No. (ADM)91-1719. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Sacks, M.H. The interaction between clinical research and treatment. *Hillside J Clin Psychiatry* 5(2):169-176, 1983.

Stark, M.J. Dropping out of substance abuse treatment: A clinically oriented review. *Clin Psychol Rev* 12(1):93-116, 1992.

Stark, M.J., and Campbell, B.K. Personality, drug use, and early attrition from substance abuse treatment. *Am J Drug Alcohol Abuse* 14(4):475-485, 1988.

Williams, M.T., and Roberts, C.S. Predicting length of stay in long-term treatment for chemically dependent females. *Int J Addict* 26(5):605-613, 1991.

AUTHORS

Judy Howard, M.D.
Professor of Pediatrics
(310) 794-1456 (Tel)
(310) 206-4215 (Fax)

Leila Beckwith, Ph.D.
Professor Emeritus
(310) 794-1456 (Tel)
(310) 206-4215 (Fax)
lbeckwith@ucla.edu (Internet)

Child Development Studies
Department of Pediatrics
University of California, Los Angeles School of Medicine
300 UCLA Medical Plaza, Suite 3300
Los Angeles, CA 90095-7033

Recruitment and Retention of Adolescent Women in Drug Treatment Research

Lawrence A. Palinkas, Catherine J. Atkins, Pamela Noel, and Christopher Miller

INTRODUCTION

Although randomized clinical trials (RCTs) are universally regarded as the standard for establishing treatment efficacy (Shapiro and Louis 1983), relatively few RCTs have been attempted in the field of outpatient drug abuse treatment. Of those that have been attempted, most have reported subject recruitment and retention to be among the most pervasive problems encountered. Considerable attention has been devoted to the issue of subject retention (Carroll et al. 1991; Howard et al. 1990, pp. 66-79; Gainey et al. 1993; Kleinman et al. 1990, pp. 24-38). However, relatively little is known of the specific problems associated with recruiting and retaining adolescents in drug treatment studies. Given that most drug abuse problems begin in adolescence (Kandel and Logan 1984; O'Malley et al. 1984), a better understanding is required of the needs of these potential clients and the methodological issues involved in recruiting and keeping them in drug treatment studies.

This chapter examines the issues of recruitment and retention of adolescents in drug treatment studies, with a particular emphasis on the experience of one of the Perinatal-20 studies that deals exclusively with adolescent women between ages 14 and 19.

RECRUITMENT

Subject Recruitment in Drug Treatment Clinical Trials

Subject recruitment depends on natural constraints on the pool of available subjects and on decisions made by both the investigator and prospective clients. Only a minority of persons with drug abuse problems participate in drug treatment programs, experimental or otherwise, and those who do so are not representative of the entire population of drug abusers (Carroll and Rounsaville 1992; Howard et al. 1990, pp. 66-79). Moreover, failure to recruit a sufficient number of subjects within the specified timeframe has several consequences, including disruption of established timetables; reallocation of scarce personnel and other resources to increase recruitment

efforts; a weakening of the intervention's therapeutic power, especially in interventions requiring group participation, if the groups are chronically undersized; a reduction in the study's sample size and power to detect statistically significant differences; and a reduction in staff morale (Ashery and McAuliffe 1992).

Three factors have been identified as contributing to problems in subject recruitment in drug treatment studies. The first pertains to the motivation and willingness of individuals with drug abuse problems to participate in such studies. Howard and colleagues (1990, pp. 66-79) note that prospective clients with drug abuse problems usually make their own assessments of the probable costs and benefits arising from study participation and agree or decline to participate for their own reasons. "For example, they may not think that the kind of treatment offered is relevant to their condition or consistent with their values. Further, their motivation to volunteer for research may be related to their motivation to recover (Cox and Klinger 1988) and hence interacts with the treatment condition to which they are assigned" (Howard et al. 1990, p. 69). Clients may be reluctant to volunteer for treatments that are unfamiliar to them or to engage in the role of a "guinea pig" by virtue of participating in an experimental form of treatment (Ashery and McAuliffe 1992). This is especially true with behavioral interventions. "Relatively unmotivated clients seem to find chemotherapies (such as methadone detoxification and maintenance) and short-term residential treatments more attractive than outpatient drug-free, psychosocial treatments" (Ashery and McAuliffe 1992, pp. 312-313).

A second factor contributing to difficulties in recruiting subjects for drug treatment RCTs lies in the selection (inclusion and exclusion) criteria established by the investigators. Subjects often are selected on the basis of age, sex, diagnosis, history and severity of condition, presence or absence of psychiatric comorbidity, or suitability (either for reasons of relevance to the problem under study, ethical reasons, or the sake of convenience) (Ashery and McAuliffe 1992; Howard et al. 1990, pp. 66-79; Carroll and Rounsaville 1990, pp. 91-104). Selection criteria often have the practical consequence that more individuals are excluded than included, making it difficult to obtain large samples for sufficient statistical power to detect significant differences generalizable across a larger population.

A third factor is that drug treatment studies are frequently hampered by institutional constraints in selecting and recruiting potential participants. These constraints include fewer abusers than anticipated in an area, difficulties reaching abusers, slowness of developing referral networks, caution by referring counselors in accepting a new treatment form, and

problems in collaborating with treatment agencies (Ashery and McAuliffe 1992).

Drug treatment studies for adolescents face additional recruitment problems. Adolescents are more likely to be gateway drug abusers (alcohol, marijuana, tobacco) than regular users of opiates, cocaine, hallucinogens, and other "heavy" drugs (Dusenbury et al. 1992, pp. 832-842). Consequently, the problem of finding sufficient numbers of "abusers," as reported in other RCTs, becomes magnified in recruitment of adolescents.

The pattern of drug use among adolescents also has implications for willingness to participate in drug treatment programs. A study by McAuliffe and colleagues (1991) found that only 4 percent of adults who met clinical criteria for marijuana abuse in a survey of drug abuse in Rhode Island had ever sought treatment. In contrast, 65 percent of opiate abusers had sought treatment at some point. Because most adolescents are gateway drug abusers, they may be reluctant to seek treatment for what they do not perceive to be a serious drug problem, a perception reinforced by the fact that a large percentage of their peers are also gateway drug abusers.

Another problem faced by RCTs of drug treatment interventions is that unless legally emancipated, most adolescents are considered "minors" and require parental approval before participating in an RCT. When drug-using adolescents are dealt with, this approval can be difficult to secure for several reasons, including the absence of a parent or legal guardian, reluctance of parents to admit that their children are abusing drugs, lack of information, inability to communicate with parents and guardians regarding the objectives of the study, unclear requirements for participation, and anticipated risks and benefits.

Finally, as has been the case with treatment programs for women in general, most drug treatment programs have difficulty recruiting adolescents because these programs are not structured to meet the specific needs of this client population. With respect to teenagers, these needs include continuing education, child care services, job training, social skills training, and so on.

On the other hand, family ties and social networks can facilitate recruitment of adolescent subjects. Parents can be a powerful influence in motivating an adolescent to participate in a drug treatment study. Social networks also can help in this regard, especially when adolescents who have already participated in a study communicate their experience to peers, encouraging them to participate also.

Recruitment of Adolescents in Project PALS

The purpose of Positive Adolescent Life Skills (PALS) is to evaluate the effectiveness of social skills development and social network restructuring and case management for two groups of adolescents: (1) pregnant adolescent females who are using drugs or who are at risk for using drugs and (2) nonpregnant adolescent females who are either using drugs or at risk for using drugs and who are at risk for pregnancy.

After screening, teens are randomly assigned to one of two skills training conditions: (1) PALS skills training or (2) no skills training. These groups meet for 90 minutes once each week for 16 weeks. Teens in each of these skills training conditions also are randomly assigned to one of two case management conditions: (1) casework (case management by master's-level social workers) or (2) no case management. All teens participate in an educational course called the "Facts of Life," which covers the consequences of drug use, child and adolescent development, and sexual responsibility. Thus, teens assigned to the PALS skills training condition also attend the 90-minute Facts of Life class once each week for 16 weeks. Each session is taught by a nurse practitioner or clinical social worker.

Based on a model developed by Catalano and Hawkins (1985, pp. 157-181), the PALS skills training program is a combination of cognitive and behavioral training to improve social skills and restructure the teen's social network. Social skills include (1) saying no to alcohol, other drugs, and unsafe sex; (2) handling criticism; (3) providing positive support to family and friends; (4) being assertive; and (5) making healthy decisions based on problemsolving procedures and good information. Network restructuring includes skills to improve the quantity and quality of the teen's social network through (1) network definition, (2) positive and negative support identification, and (3) planning for network restructuring.

The case management model, traditionally called casework, has been operationally defined and is "delivered" by a master's-level social worker. Case management consists of 16 weekly sessions with a teen. Sessions are normally 1 hour long, although the case manager has the freedom to transport a teen to other services (e.g., immigration lawyer, food bank, free baby clothes provider). During the first 4 weeks of case management, an extensive needs assessment and a psychosocial assessment are completed. Based on this information, the teen participant and the case manager develop a treatment plan, begin counseling, and make appropriate referrals.

PALS recruits teens from a variety of sources, including the Adolescent Medicine and Teen Obstetric Clinics of the University of California, San Diego (UCSD) Medical Center; UCSD Family Practice Clinics;

the Options for Recovery Program of San Diego County Drug Abuse Services; San Diego Adolescent Pregnant and Parenting Project (SANDAPP) and alternative school programs of San Diego public schools; and San Diego County's juvenile probation program. In the existing PALS program, 38 percent of referrals have come from UCSD Adolescent Medicine and Teen Obstetric Clinics, 11 percent from SANDAPP and alternative school programs of the San Diego public schools (pregnancy and parenting programs, school nurses, counselors), 7 percent from staff outreach efforts (advertisements, nonagency referrals), 6 percent from other professionals (probation officers, community health clinic staff members, public health nurses, therapists), and 38 percent from parents, PALS graduates, self-referrals, and other sources.

Basic guidelines for referrals to PALS are as follows:

- Participants must be females between ages 14 years and 18 years, 11 months.
- Participants must reside within or near the borders of the city of San Diego.
- Participants must not reside in juvenile hall, a residential treatment facility, or a group home.
- Teens must not be involved with a similar research project or agency during the course of the program.
- Teens who have developmental or mental health issues may be prohibited from fully participating in the treatment group sessions.
- Participation must be strictly voluntary (i.e., not a condition of probation).
- A legal minor must have a signed parental consent.
- Participants must be at risk for drug use based on study screening criteria.

A clinical screening is conducted on entry to assess a teen's risk factors for drug use as well as background and demographic information. Referrals for standard medical care (including prenatal care for pregnant teens) are made to UCSD Family Practice Clinics or the UCSD Adolescent Medicine and Teen Obstetric Clinics. Screening also includes the use of two assessment tools. The Adolescent History Interview Schedule (Hawkins et al. 1987) provides descriptive data on the teen and her family

in the following areas: demographic characteristics, including age, ethnicity, education, and language spoken with friends and family members; sexual history; her and her family's alcohol and other drug history and use; her school involvement and attitudes; delinquency; and social network characteristics (e.g., the extent, domain, and duration of acquaintances; whether members use alcohol and other drugs; and whether the teen's drug use has caused the network members any problems). The Problem Oriented Screening Instrument for Teens (POSIT) (Rahdert 1991) is used to identify problematic functioning in 10 critical areas of adolescent development: substance use/abuse, physical health, mental health, family relations, peer relations, educational status, vocational status, social skills, leisure and recreation, and aggressive behavior and delinquency. Endorsement of a "red flag" item or a critical number of items within a functional area item-set indicates that further indepth evaluation is required.

To date, 1,097 adolescent women have been evaluated for participation in PALS. Of this number, 802 (73.1 percent) did not participate, and 295 (26.9 percent) were found to be eligible and were willing and able to participate. Participants were recruited in 7 separate waves of approximately 40 to 50 students per wave.

A comparison of the social and demographic characteristics of these two groups of potential and actual participants is provided in table 1 below. The two groups were similar with respect to number of years of schooling, marital status, and pregnancy status (pregnant). PALS participants were significantly younger (t=3.89, p<0.001), more likely to be Latina and less likely to be white or "other" (Asians and Pacific Islanders, Native Americans, others) (x^2=28.3, degrees of freedom *[df]*=3, p<0.0001) than nonparticipants. Nonparticipants were more likely to live in residential treatment facilities, group homes, and foster homes or on the street than participants (x^2=58.3, df=1, p<0.0001). Nonparticipants were also more likely to drop out of school than participants (x^2=7.1, df=1, p<0.01).

Table 2 provides a comparison of the psychosocial characteristics of PALS participants and nonparticipants. PALS participants scored higher in the mental health, peer relations, educational status, vocational status, and aggressive behavior or delinquency domains than nonparticipants.

Referral sources of participants and nonparticipants are indicated in table 3. Participants were more likely to be referred by noninstitutional sources (current or former PALS participants, self-referral, parent or guardian) than nonparticipants (x^2=29.8, df=1, p<0.0001). However, it was only after the project was well under way that noninstitutional referrals became an increasingly greater percentage of total referrals,

TABLE 1. *Social and demographic characteristics of potential adolescent women drug treatment clients, by eligibility and participation status*

Characteristic	PALS Participants (N=295)	Nonparticipants (N=802)
Mean (SD) age	16.6 (1.3)	17.0 (1.9)*
Mean (SD) highest grade level enrolled	10.2 (1.8)	10.1 (2.0)
School dropouts (percent)	9.7	16.7†
Ethnicity (percent)		
African-American	38.0	34.2*
Latina	45.4	33.2
White	10.8	19.7
Other	5.8	12.9
Marital status (percent)		
Single	92.2	92.2
Married or living with boyfriend	7.8	7.8
Pregnancy status (percent)		
Pregnant	18.0	16.8
Parenting	10.8	17.4
Neither	71.2	65.8
Residence (percent)		
With family‡	98.2	77.3*
Without family§	1.8	22.7

*$p<0.001$
†$p<0.01$
‡In single-family dwelling, apartment, duplex, other
§In residential treatment, group home, or foster home or on street
KEY: PALS=Positive Adolescent Life Skills; SD=standard deviation

a point illustrated by comparison of the percentage of referrals from former or current PALS participants with the percentage of referrals from the UCSD Adolescent Medicine and Teen Obstetric Clinics by wave (table 4).

Reason for nonparticipation was available for 640 (79.8 percent) of the 802 nonparticipating teens referred to PALS. Almost two-thirds (63.6 percent) were disqualified because of inability to meet program eligibility requirements; slightly more than one-third (36.3 percent) did

TABLE 2. *Psychosocial characteristics of potential adolescent women drug treatment clients, by eligibility and participation status*

POSIT Domain	PALS Participants (N=295) Mean (SD)	Nonparticipants (N=485) Mean (SD)	Significance
Substance use/abuse	1.18 (2.29)	1.12 (2.18)	NS
Physical health	2.93 (1.87)	3.00 (1.99)	NS
Mental health	8.03 (4.78)	6.82 (4.72)	$p<0.001$
Family relations	5.56 (3.33)	5.39 (3.43)	NS
Peer relations	3.72 (2.35)	2.83 (2.29)	$p<0.001$
Educational status	9.31 (3.97)	8.18 (4.29)	$p<0.001$
Vocational status	6.46 (2.92)	5.67 (3.22)	$p<0.001$
Social skills	3.26 (1.89)	3.19 (2.07)	NS
Leisure and recreation	5.02 (2.25)	4.73 (2.53)	NS
Aggressive behavior or delinquency	5.36 (3.67)	4.61 (3.45)	$p<0.01$

KEY: POSIT=Problem Oriented Screening Instrument for Teens; PALS=Positive Adolescent Life Skills; SD=standard deviation; NS=not significant

not participate for personal reasons, citing a lack of interest, failure to meet the recruitment deadline for application into the program, or a schedule conflict. Specific reasons for nonparticipation are given in table 5.

These data indicate that there are substantial differences between drug-using adolescents who enter drug treatment and those who do not. In the case of PALS, participants were younger, included more Latinas and fewer whites and other ethnic groups, were less likely to drop out of school, and were less likely to live in residential treatment facilities, foster homes, and group homes because these teens were considered to be currently undergoing treatment in another program, one of the project's exclusion criteria. Approximately 1 of 10 nonparticipants was excluded because she was 19 years or older and thus ineligible to participate, accounting in part for the older mean age of the nonparticipant group. School dropouts were less likely to participate because project staff members made a special effort to encourage dropouts to reenroll in school; those not interested in doing so were less willing to participate.

It is uncertain why the program recruited greater-than-expected numbers of Latinas and fewer-than-expected numbers of whites and other ethnic groups. Further analysis of the characteristics of participants and nonparticipants

TABLE 3. *Referral sources of potential adolescent women drug treatment clients, by participation status*

	PALS Participants (N=295)		Nonparticipants (N=802)	
Referral Sources	N	%	N	%
Institutional referrals				
Project PALS representative	14	4.7	50	6.2
UCSD Adolescent Medicine and Teen Obstetric Clinics	99	33.6	427	53.2
County probation office	2	0.6	5	0.6
SANDAPP	17	5.8	56	7.0
Community-based agencies	10	3.4	10	1.2
Child protective services	4	1.4	5	0.6
County health department	3	1.0	16	2.0
City schools	14	4.7	15	1.9
Noninstitutional referrals				
Current or former PALS participants	108	36.6	191	23.8
Self-referral	15	5.1	22	2.7
Parent or guardian	9	3.1	5	0.6
Total institutional referrals	163	55.2	584	72.7*
Total noninstitutional referrals	132	44.8	218	27.1

*Percentages add to only 72.7 because of rounding.

KEY: PALS=Positive Adolescent Life Skills; UCSD=University of California, San Diego; SANDAPP=San Diego Adolescent Pregnant and Parenting Project

TABLE 4. *Comparison of percentage of referrals from UCSD Adolescent Medicine and Teen Obstetric Clinics with percentage of referrals from current or former PALS participants, by wave*

	Wave						
Referral Source	1	2	3	4	5	6	7
UCSD Adolescent Medicine and Teen Obstetric Clinics	96.9	64.7	63.2	54.2	41.3	19.9	22.6
Current or former PALS participants	0.0	8.7	14.8	18.8	28.8	44.7	51.1
Other sources	3.1	26.6	22.0	27.0	29.9	35.4	26.3

KEY: UCSD=University of California, San Diego; PALS=Positive Adolescent Life Skills

TABLE 5. *Reasons of potential adolescent women drug treatment clients for not participating in PALS*

	Nonparticipants (N=640)	
Reasons	N	%
Program-based reasons		
Too old or young	78	12.2
Not living at home	69	10.8
No POSIT risk factors	32	5.0
No parental consent	46	7.2
Out of catchment area	125	19.5
In another program	10	1.6
Poor communication skills	21	3.3
No preassessment	9	1.4
Change in eligibility	5	0.8
Sibling in same program	4	0.7
Other institutional reasons	8	1.2
Participant-based reasons		
Time conflict	26	4.0
Not interested	93	14.5
Missed deadline	114	17.8
Total program-based reasons	407	63.7
Total participant-based reasons	233	36.3

KEY: PALS=Positive Adolescent Life Skills; POSIT=Problem Oriented Screening Instrument for Teens

found no ethnic differences with respect to referral source or reasons for nonparticipation. However, Latinas were less likely to be living outside the home at time of referral than other ethnic groups (x^2=13.4, *df*=1, *p*=0.0002). Moreover, most PALS staff members were either Latinas or bilingual whites and African-Americans, whose ability to speak Spanish was reassuring to both potential participants—even if these participants used English as their primary language—and their parents, many of whom were Spanish speaking only.

Differences in the POSIT domain scores of PALS participants and nonparticipants were to be expected because these psychosocial characteristics were an important determinant of program eligibility. Individuals were identified as being at risk for drug use if they scored 1 or higher on the substance use/abuse questions and 1 or higher on two of four other domains: mental health, family relations, peer relations,

and aggressive behavior or delinquency. Participants were significantly different from nonparticipants in three of these domains: mental health, peer relations, and aggressive behavior or delinquency. Thus, if participants were recruited on the basis of self-reported drug use alone, those recruited would have been viewed as representative of the population of referred adolescent women.

The PALS experience also illustrates the importance of informal or noninstitutional sources of referrals, an experience also found in other RCTs of experimental drug treatment interventions (Ashery and McAuliffe 1992). The percentage of referrals from former and current graduates increased over the course of the project as more teens came into contact with the program and derived positive benefits from participation. Participants also included larger percentages of self-referrals and referrals from parents, both of whom recognized a need for some form of behavioral intervention to treat an existing or potential drug problem. In contrast, teens who were recruited at random from institutional sources, such as the UCSD Adolescent Medicine Clinic, were less likely to participate because they were less likely to have a reason for attending, either because they were not at risk for drug use, not previously identified by the health or social services system as having a drug problem even if such a problem existed, or not having a friend who had previously been through the program to encourage them to participate.

Previous research has found that different types of programs attract different types of clients for different reasons (Deykin et al. 1992; Moise et al. 1981). A study by Carroll and Rounsaville (1992) found that untreated cocaine abusers were more likely to lack social support and pressure to enter treatment than treated abusers. In the case of PALS, the emphasis on social skills development and social network restructuring, combined with the possibility of individual case management and a bilingual staff, had greater appeal to Latinas from more stable and secure living environments, adolescents who already had been identified by the system as having an existing drug problem, adolescents or their parents who recognized the need for program participation, and adolescents who were part of the social networks of current or former PALS participants.

RETENTION

Drug abuse treatment usually requires a long-term commitment on the part of clients. For these programs to demonstrate success, they must maintain high rates of *attendance* and low rates of *attrition*. Although both attendance and attrition determine the level of participant retention, it is important to distinguish between the two. Attrition typically refers

to actual dropout from the study, whereas attendance typically serves as a measure of exposure to treatment. Both attrition and low attendance introduce biases that can have a substantial impact on internal validity and generalizability of the research findings (Howard et al. 1990, pp. 66-79). Reported attrition rates vary from study to study. It is not uncommon for drug abuse treatment studies to experience dropout rates of 40 to 50 percent or higher (Baekeland and Lundwall 1975; Kleinman et al. 1990, pp. 24-38). However, dropout is only part of the problem. Some subjects who remain in the study are exposed to less than the optimum number of treatment sessions.

It is important to consider the relationship between retention and outcome, because those who have low attendance or who drop out may differ in systematic ways from the larger sample of those recruited to participate in drug treatment research. Perhaps even more important than concerns about generalizability of research findings is the notion that there is a "dose-response" effect of treatment (Gainey et al. 1993). However, the relationship between retention and outcome often confounds the exposure to treatment with motivation to attend (Atkins et al. 1990; Kaplan and Atkins 1987). As with subject recruitment, characteristics of the client and the program have been identified as influencing the retention of participants in drug treatment programs. However, little is known of the role of these characteristics in retaining adolescent clients. Identifying predictors of retention can provide important insights for enhancing the effect and success of drug treatment programs for this age group.

Client Characteristics

Severity of Drug Abuse. Severity of use seems to have a somewhat paradoxical effect on treatment retention (Baekeland and Lundwall 1975). For example, long-term and heavier drug use may be associated with other risk factors that are incompatible with completing treatment (Gainey et al. 1993). On the other hand, those with heavier drug use and drug-related problems may have stronger motivations to remain in treatment (Carroll et al. 1991). A recent study by Gainey and colleagues (1993) found that those who used cocaine for shorter periods of time were less likely to remain in treatment. The authors suggest that these particular clients may not have perceived their addictions to be serious enough to warrant drug treatment. Most surprisingly, evidence suggests that the longer one is able to remain abstinent from using drugs like cocaine and alcohol, the more likely it is that he or she will complete treatment (Means et al. 1989).

The particular drug being abused also may affect the likelihood of treatment completion. Those who abuse legal drugs such as alcohol are often less likely to seek and remain in treatment because they do not see themselves

as having a particularly serious problem; therefore, they view treatment as unnecessary. These individuals are also less likely to have been referred to treatment as part of a court-ordered requirement. Such a requirement has been linked to improved retention in drug treatment (Gainey et al. 1993). In general, most findings on severity of drug use do not bode well for teens remaining in treatment. Teens are likely to perceive their drug use problems as less severe, and they are less likely to have external pressure to remain in treatment.

Comorbidity. Epidemiologic evidence suggests that psychiatric comorbidity, such as anxiety, depression, and personality disorders, increases the risk of alcohol or other drug dependence in adolescents (Christie et al. 1988; Deykin et al. 1987; Kandel 1982). The presence of psychiatric comorbidity also has been identified as a potentially important predictor of drug treatment outcomes (McLellan et al. 1983). However, few studies have systematically studied the impact of psychiatric comorbidity on retention in treatment. Although anxiety and depression may provide some motivation to seek and remain in treatment, drug use and abuse also may ameliorate the systems of psychiatric comorbidities and increase the likelihood that subjects will drop out of treatment.

Social and Family Support. The relationships among social and family support, social isolation, and drug abuse treatment retention have been examined in several studies. Social isolation is typically found to be predictive of early dropout from treatment. In a classic review article of treatment dropout, Baekeland and Lundwall (1975) found that social isolation or lack of affiliation predicted early treatment dropout in 19 of 19 studies (100 percent) that addressed the issue. More recently, Gainey and colleagues (1993) found that individuals living alone were three times less likely to complete treatment for cocaine abuse than those living with others. On the other hand, being in a functional relationship or involved in a prosocial organization has been shown to be predictive of longer retention in treatment (Hawkins and Catalano 1985; Means et al. 1989). Thus, positive social support for remaining in treatment can be an important influence.

Program, Therapist, and Environmental Characteristics

Convenience and Incentives. In some cases, characteristics of the treatment environment have been found to be more important motivators for treatment retention than characteristics of the individual seeking treatment (Miller 1985). One of the nonspecific characteristics of the treatment program that has been shown to be important for retention is convenience of the treatment in terms of location and transportation. For example, one study found that the likelihood of attendance was better

predicted by the distance a client had to travel to get to the treatment than by any personality characteristics of the client (Miller 1985). Another important treatment program characteristic is the length of time between referral and program entry. Several studies have demonstrated that both recruitment and retention are negatively correlated with the amount of time the client must wait before being seen in treatment (Miller 1985). Finally, both negative and positive incentives have been shown to be related to treatment retention. Legal pressure has been found to be one of the strongest motivators for treatment retention (Gainey et al. 1993).

Characteristics of Therapists. One of the most widely discussed but perhaps least investigated treatment environment variables is the therapist who delivers the treatment (Miller 1985; Crits-Christoph et al. 1990, pp. 39-49). Studies that attempt to characterize treatment effect typically find at least a moderate difference in treatment effect that can be attributed to the therapist who delivers the treatment (Crits-Christoph et al. 1990, pp. 39-49; McLellan et al. 1988). With regard to early dropout, Baekeland and Lundwall (1975) found that therapist attitude and behavior were more predictive of treatment retention than socioeconomic status and motivation of the client in 35 of 35 studies (100 percent) that assessed therapist variables. Therapist attributes, such as expectation for success, confrontational therapeutic style, and permissiveness, all have been implicated as important. However, there are few systematic studies that explicitly characterize the attributes associated with either successful outcomes or client attrition. Certainly, therapist characteristics are important and should be systematically studied with respect to both client and treatment attributes. Teen clients may be particularly influenced by some of the more salient features of the therapist, such as age, ethnicity, and style of treatment delivery.

Retention of Adolescents in PALS

The experience of PALS offers a unique opportunity to assess some common client characteristics that predict program attendance among a group of adolescent women at high risk for drug abuse and pregnancy. Based on the findings of earlier studies described above and their potential implications for adolescent participation in drug treatment, the authors examined four categories of client characteristics to determine whether one or more of these characteristics were predictive of client attendance at program activities, including classes or scheduled meetings with case managers.

The first category of client characteristics comprised social and demographic characteristics found to be associated with participant recruitment. These included age, ethnicity (Latina or non-Latina), and

school dropout (yes or no) status. In addition, the effect of pregnancy or parenting status (yes or no) on attendance was examined.

The second category of client characteristics included measures of family and social support. The Family Cohesion and Evaluation Scale III (FACES-III) (Olson et al. 1985) was used to obtain a measure of *family cohesion*. The Social Support Questionnaire-Revised (SSQ-R) (Sarason et al. 1983) was used to identify the number of people to whom clients could turn and on whom they could rely in given sets of circumstances (*availability of social support*). The Missouri Peer Relations Inventory (MPRI) (Borduin et al. 1989) consists of 13 five-point scale items relating to various aspects of adolescent peer relations, such as verbal and physical aggression, popularity, respect for others, and sense of humor. This instrument was used to assess levels of *peer emotional bonding*, *aggression*, and *acceptance*. The Children's Report of Parent Behavior Inventory (CRPBI-30)(Schludermann and Schludermann 1988) is a 30-item (for each parent) instrument to assess children's perceptions of their parents' child-rearing style. Three factor scores were derived for each parent: *acceptance*, *psychological control*, and *firm control.* Finally, the *peer relations* and *social skills* domains of the POSIT were used to evaluate nonfamily support.

The third category of client characteristics included measures of potential psychiatric comorbidity. The Children's Depression Inventory (CDI) (Kovacs 1992) was used to assess level of *depressive symptomatology*, and the Revised Children's Manifest Anxiety Scale (Reynolds and Richmond 1979) and anxiety subscale of the Achenbach Inventory (Achenbach 1991) were used to assess the level of *anxiety*. *Self-esteem* was measured with Rosenberg's (1989) scale, and *locus of control* was assessed by means of the Locus of Control Inventory (Nowicki and Strickland 1973).

The fourth category of client characteristics included self-reported measures of *drug use*, *sexual activity*, and *delinquent behavior*. These measures included the substance use/abuse domain of the POSIT, the sexuality and drug abuse subscales of the A-File (McCubbin 1981), and the Self Reported Delinquency Scale (Elliott 1980) and delinquency subscale of the Achenbach Inventory (Achenbach 1991).

Preassessment data for all subjects who enrolled in PALS and for whom complete attendance information was available (waves 1 through 6, n=234) were used for this analysis. Nonparametric correlation coefficients (Spearman's r_s) (Snedecor and Cochran 1980) were calculated to examine the association between each candidate measure and attendance. Generalized logistic regression analyses of attendance (i.e., linear regression of the logit of the probability of attendance based on the

observed number of sessions attended in a fixed number of trials) were run using the Statistical Package for the Social Sciences for Personal Computers (SPSSPC) "Probit Analysis" program (Norusis 1992) with the logit option. Variables were selected for an initial regression if the significance level for the Spearman's r_s was less than 0.10. These variables included general delinquency (r_s=–0.21, p=0.001), POSIT substance use/abuse score (r_s=–0.19, p=0.002), peer acceptance (r_s=0.18, p=0.004), availability of social supports (r_s=0.16, p=0.01), locus-of-control scale score (r_s=–0.14, p=0.025), and family cohesion (r_s=0.12, p=0.06). All these variables had adequate regression coefficients. (The absolute value of the coefficient divided by the standard error was greater than 1.) The remaining variables then were added to this base set, one at a time, and used in separate analyses. Variables found to have an absolute value of the coefficient divided by standard error greater than 1 then were added into the final model. These additional variables were age, pregnancy or parenting status, school dropout status, maternal firm control, sexual activity, and depression.

The final model is presented in table 6. Age, pregnancy or parenting status, school dropout status, family cohesion, locus of control, number of social supports, POSIT substance use/abuse at preassessment, peer acceptance, sexual activity, maternal firm control, and general delinquency were significant independent predictors of program attendance in this cohort. Family cohesion, number of social supports, peer acceptance, maternal firm control, and general delinquency were directly associated with program attendance, whereas age, pregnancy or parenting status, school dropout status, internal locus of control, POSIT substance use/abuse, and sexual activity were inversely associated with attendance. Depression was marginally (p=0.06) associated with attendance.

These results suggest that some client characteristics associated with successful recruitment, including age and school dropout status, also are associated with successful retention. Adolescent women who are older or who have a history of dropping out of school are less likely to attend the treatment sessions. Adolescents who are pregnant or parenting also are less likely to attend, usually citing their pregnancy or child care responsibilities, if parenting, as a reason for failure to attend subsequent sessions.

In part, these results confirm the findings of studies on predictors of retention of older clients in drug treatment programs. For instance, as with retention of older adults, adolescent retention is inversely associated with severity of drug use: Those adolescents most in need of drug treatment are often the least likely to remain in treatment. Self-reported drug abuse prior to treatment was the strongest predictor of program attendance in this cohort.

TABLE 6. *Regression of participant characteristics on program attendance: Results of logit analysis*

Variable	B	SE	B/SE	Significance
Age	–0.0075	0.0025	–2.99	<0.01
Pregnant or parenting*	–0.1711	0.0923	–1.85	<0.05
School dropout*	–0.2163	0.1273	–1.70	<0.05
Family cohesion	0.0156	0.0045	3.49	<0.01
Locus of control	–0.0155	0.0080	–1.94	<0.05
Number of social supports	0.0207	0.0031	6.60	<0.001
POSIT substance use/abuse	–0.1254	0.0154	–8.15	<0.001
Peer acceptance	0.1294	0.0263	4.92	<0.001
Sexual activity	–0.1609	0.0367	–4.39	<0.001
Maternal firm control	0.0225	0.0070	3.19	<0.01
Depression	0.0085	0.0057	1.51	<0.07
General delinquency	0.0003	0.0001	2.13	<0.05

*yes or no

KEY: B=regression coefficient; SE=standard error of regression coefficient

The PALS experience also illustrates that many client characteristics associated with adolescent drug use, including depressive symptoms (Deykin et al. 1987), delinquent behavior (Jessor and Jessor 1977), availability of social support (Wills and Vaughan 1989), relationship with mother (Kokotailo et al. 1992) and peers (Huba and Bentler 1980), and sexual activity (Rosenbaum and Kandel 1990), also are associated with retention in drug treatment. Other studies have demonstrated the importance of social support in predicting retention of women in drug treatment (Huselid et al. 1991). The association between sexual activity and retention in drug treatment may have been confounded by the increased likelihood of drug use among sexually active teens. However, sexual activity predicted low program attendance (B=–0.1609 [see table 6]) independent of drug use. Further research is required to determine the basis for this association.

Similarly, the association between program retention and external locus of control in the PALS cohort appears in contrast to earlier research that linked external locus of control to increased likelihood of drug use (Clark et al. 1982; Diehlman et al. 1984) and decreased likelihood of engaging in health promotion activities such as smoking cessation (Seeman and Seeman 1983). However, a study by Huselid and colleagues (1991) also reported that for women an external locus of control was associated with retention

in a drug treatment program. Further investigation of this association is required.

CONCLUSIONS

A greater understanding of the factors that promote recruitment and retention of adolescents in drug treatment programs in general and drug treatment research in particular is critical to the success of these programs. Both client and program characteristics appear to exert considerable influence on the ability of programs to successfully recruit and retain adolescent clients. With respect to the program, inclusion and exclusion criteria based on experimental design, treatment approach, and practicalities of implementation may be responsible for the exclusion of a substantial percentage of potential clients. With respect to the client, factors such as ethnicity appear to influence recruitment of adolescents in drug treatment. Factors such as severity of drug use, pregnancy or parenting status, presence of psychiatric comorbidity, delinquent behavior, sexual activity, and internal locus of control appear to decrease the likelihood of retaining adolescents in drug treatment. Factors such as age and school dropout status influence the likelihood of successful recruitment and retention. Most important, drug treatment programs for adolescents can improve both recruitment and retention efforts by making greater use of adolescent social support networks.

REFERENCES

Achenbach, T.M. *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont, 1991.

Ashery, R.S., and McAuliffe, W.E. Implementation issues and techniques in randomized trials of outpatient psychosocial treatments for drug abusers: Recruitment of subjects. *Am J Drug Alcohol Abuse* 18:305-329, 1992.

Atkins, C.J.; Senn, K.; Rupp, J.; Kaplan, R.; Patterson, T.L.; Sallis, J.F.; and Nader, P.R. Attendance at health promotion programs in Mexican-American and Anglo families: Baseline predictors and program outcomes. *Health Ed Q* 17:417-428, 1990.

Baekeland, F., and Lundwall, L. Dropping out of treatment: A critical review. *Psychol Bull* 82(5):738-783, 1975.

Borduin, C.M.; Blaske, D.M.; Cone, L.; Mann, B.J.; and Hazelrigg, M.D. "Development and Validation of a Measure of Adolescent Peer Relationships: The Missouri Peer Relations Inventory." Unpublished manuscript, Department of Psychology, University of Missouri, 1989.

Carroll, K.M., and Rounsaville, B.J. Can a technology model of psychotherapy research be applied to cocaine abuse treatment? In: Onken, L.S., and Blaine, J.D., eds. *Psychotherapy and Counseling in the Treatment of Drug Abuse*. National Institute on Drug Abuse Research Monograph 104. DHHS Pub. No. (ADM)91-1722. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Carroll, K.M., and Rounsaville, B.J. Contrast of treatment seeking and untreated cocaine abusers. *Arch Gen Psychiatry* 49:464-471, 1992.

Carroll, K.M.; Rounsaville, B.J.; and Gawin, F.H. A comparative trial of psychotherapies for ambulatory cocaine abusers: Relapse prevention and interpersonal psychotherapy. *Am J Drug Alcohol Abuse* 17:229-247, 1991.

Catalano, R.F., and Hawkins, J.D. Project skills: Preliminary results from a theoretically based aftercare experiment. In: Ashery, R.S., ed. *Progress in the Development of Cost-Effective Treatment for Drug Abusers*. National Institute on Drug Abuse Research Monograph 58. DHHS Pub. No. (ADM)85-1401. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Christie, K.A.; Burke, J.D.; and Regier, D.A. Epidemiologic evidence for early onset of mental disorders and higher risk of drug abuse in young adults. *Am J Psychiatry* 145:971-975, 1988.

Clark, J.G.; MacPherson, B.V.; and Holmes, D.B. Cigarette smoking and external locus of control among adolescents. *J Health Soc Behav* 23:253-259, 1982.

Cox, W.M., and Klinger, E. A motivational model of alcohol use. *J Abnorm Psychol* 97:168-180, 1988.

Crits-Christoph, P.; Beebe, K.L.; and Connolly, M.B. Therapist effects in the treatment of drug dependence: Implications for conducting comparative treatment studies. In: Onken, L.S., and Blaine, J.D., eds. *Psychotherapy and Counseling in the Treatment of Drug Abuse*. National Institute on Drug Abuse Research Monograph 104. DHHS Pub. No. (ADM)91-1722. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Deykin, E.Y.; Buka, S.L.; and Zeena, T.H. Depressive illness among chemically dependent adolescents. *Am J Psychiatry* 149(10):1341-1347, 1992.

Deykin, E.Y.; Levy, J.C.; and Wells, V. Adolescent depression, alcohol and drug abuse. *Am J Public Health* 77(2):178-182, 1987.

Diehlman, T.E.; Leech, S.L.; Lorenger, A.T.; and Horvath, W.J. Health locus of control and self-esteem as related to adolescent health behavior and intentions. *Adolescence* 19:935-950, 1984.

Dusenbury, L.; Khuri, E.; and Millman, R.B. Adolescent substance abuse: A sociodevelopmental perspective. In: Lowinson, J.H.; Ruiz, P.; Millman, R.B.; and Langrod, J.G., eds. *Substance Abuse: A Comprehensive Textbook*. 2d ed. Baltimore, MD: Williams & Wilkins, 1992.

Elliott, D.S. *Report of the National Youth Survey*. Report No. 26. Boulder, CO: Behavioral Research Institute, 1980. 236 pp.

Gainey, R.R.; Wells, E.A.; Hawkins, J.D.; and Catalano, R.F. Predicting treatment retention among cocaine users. *Int J Addict* 28:487-505, 1993.

Hawkins, J.D., and Catalano, R.F. Aftercare in drug abuse treatment. *Int J Addict* 20:917-945, 1985.

Hawkins, J.D.; Catalano, R.F.; and Hall, J.A. "Adolescent History Interview Schedule." Unpublished manuscript, School of Social Work, University of Washington, 1987.

Howard, K.I.; Cox, W.M.; and Saunders, S.M. Attrition in substance abuse comparative treatment research: The illusion of randomization. In: Onken, L.S., and Blaine, J.D., eds. *Psychotherapy and Counseling in the Treatment of Drug Abuse*. National Institute on Drug Abuse Research Monograph 104. DHHS Pub. No. (ADM)91-1722. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Huba, G.J., and Bentler, P.M. The role of peer and adult models for drug taking at different stages in adolescence. *J Youth Adolescence* 9(5):449-465, 1980.

Huselid, R.F.; Self, E.A.; and Gutierres, S.E. Predictors of successful completion of a halfway-house program for chemically-dependent women. *Am J Drug Alcohol Abuse* 17(1):89-101, 1991.

Jessor, R., and Jessor, S.L. *Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth*. New York: Academic Press, 1977. 281 pp.

Kandel, D.B. Epidemiological and psychosocial perspectives on adolescent drug use. *J Am Acad Child Psychiatry* 21(4):328-347, 1982.

Kandel, D.B., and Logan, J.A. Patterns of drug use from adolescence to young adulthood: Periods of risk for initiation, continued use, and discontinuation. *Am J Public Health* 74(7):660-666, 1984.

Kaplan, R.M., and Atkins, C.J. Selective attrition causes overestimates of treatment effects in studies of weight loss. *Addict Behav* 12:297-302, 1987.

Kleinman, P.H.; Woody, G.E.; Todd, T.C.; Millman, R.B.; Kang, S.-Y.; Kemp, J.; and Lipton, D.S. Crack and cocaine abusers in outpatient psychotherapy. In: Onken, L.S., and Blaine, J.D., eds. *Psychotherapy and Counseling in the Treatment of Drug Abuse*. National Institute on Drug Abuse Research Monograph 104. DHHS Pub. No. (ADM)91-1722. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Kokotailo, P.K.; Adger, H.; Duggan, A.K.; Repke, J.; and Joffe, A. Cigarette, alcohol, and other drug use by school-age pregnant adolescents: Prevalence, detection, and associated risk factors. *Pediatrics* 90:328-334, 1992.

Kovacs, M. *The Children's Depression Inventory (CDI) Manual*. North Tonawanda, NY: Multi-Health Systems, 1992.

McAuliffe, W.E.; Breer, P.; Ahmadifar, N.W.; and Spino, C. Assessment of drug abuser treatment needs in Rhode Island. *Am J Public Health* 81:365-371, 1991.

McCubbin, H.I. A-File. In: Olson, D.H., ed. *Inventories Used in a National Survey of Families Across the Life Cycle.* St. Paul, MN: University of Minnesota Press, 1981.

McLellan, A.T.; Luborsky, L.; Woody, G.E.; Druley, K.A.; and O'Brien, C.P. Predicting response to alcohol and drug abuse treatments: Role of psychiatric severity. *Arch Gen Psychiatry* 40:620-625, 1983.

McLellan, A.T.; Woody, G.E.; Luborsky, L.; and O'Brien, C.P. Is the counselor an "active ingredient" in substance abuse treatment? *J Nerv Ment Dis* 176:423-430, 1988.

Means, L.B.; Small, M.A.; Capone, D.M.; Capone, T.J.; Condren, R.; Peterson, M.; and Hayward, B. Client demographics and outcome in outpatient cocaine treatment. *Int J Addict* 24:765-783, 1989.

Miller, W.R. Motivation for treatment: A review with special emphasis on alcoholism. *Psychol Bull* 98:84-107, 1985.

Moise, R.; Reed, B.G.; and Conell, C. Women in drug abuse treatment programs: Factors that influence retention at very early and later stages in two treatment modalities. A summary. *Int J Addict* 16:1295-1300, 1981.

Norusis, M.J. *SPSS/PC+ Advanced Statistics, Version 5.0*. Chicago: SPSS, Inc., 1992. 481 pp.

Nowicki, S., and Strickland, B. A locus of control scale for children. *J Consult Clin Psychol* 97:64-75, 1973.

Olson, D.H.; Portner, J.; and Lavee, Y. *FACES III*. St. Paul, MN: Department of Family Social Science, University of Minnesota, 1985.

O'Malley, P.M.; Bachman, J.G.; and Johnson, L.D. Period, age, and cohort effects on substance abuse among American youth: 1976-1982. *Am J Public Health* 74:682-688, 1984.

Rahdert, E.R. *The Adolescent Assessment/Referral System Manual.* DHHS Pub. No. (ADM)91-1735. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991. 109 pp.

Reynolds, C.R., and Richmond, B.O. Factor structure and construct validity of what I think and feel: The Revised Children's Manifest Anxiety Scale. *J Pers Assess* 43(3):281-283, 1979.

Rosenbaum, E., and Kandel, D.B. Early onset of adolescent sexual behavior and drug involvement. *J Marriage Family* 52:783-798, 1990.

Rosenberg, M. *Society and the Adolescent Self-Image.* Middletown, CT: Wesleyan University Press, 1989. 326 pp.

Sarason, I.G.; Levine, H.M.; Basham, R.B.; and Sarason, B.R. Assessing social support: The Social Support Questionnaire. *J Pers Soc Psychol* 44:127-139, 1983.

Schludermann, S., and Schludermann, E. “Questionnaire for Children and Youth.” Unpublished manuscript, Department of Psychology, University of Manitoba, 1988.
Seeman, M., and Seeman, T.E. Health behavior and personal autonomy: A longitudinal study of the sense of control in illness. *J Health Soc Behavior* 24:144-160, 1983.
Shapiro, S.H., and Louis, T.A., eds. *Clinical Trials: Issues and Approaches.* New York: Marcel Dekker, 1983. 209 pp.
Snedecor, G.W., and Cochran, W.G. Correlation. In: Snedecor, G.W., and Cochran, W.G., eds. *Statistical Methods*. Ames, IA: Iowa State University Press, 1980.
Wills, T.A., and Vaughan, R. Social support and substance use in early adolescence. *J Behav Med* 12(4):321-339, 1989.

ACKNOWLEDGMENT

Preparation of this chapter was supported by National Institute on Drug Abuse grant R18-DA-06911.

AUTHORS

Lawrence A. Palinkas, Ph.D.
Associate Professor
(619) 543-5493 (Tel)
(619) 543-5996 (Fax)
lawrence-palinkas@som-mpf.ucsd.edu (Internet)

Catherine J. Atkins, Ph.D.
Associate Adjunct Professor
(619) 594-5350 (Tel)
(619) 594-3590 (Fax)
catkins@sunstroke.sdsu.edu (Internet)

Pamela Noel, M.S.W., M.P.H.
Staff Research Associate
(919) 914-4920 (Tel)

Christopher Miller, M.S.
Staff Research Associate
(619) 294-3700 (Tel)
(619) 543-5996 (Fax)
cmiller@comm.family-ppals.ucsd.edu (Internet)

Department of Family and Preventive Medicine
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0807

Retention Issues Involving Drug-Abusing Women in Treatment Research

Robert A. Lewis, Deborah L. Haller, Doreen Branch, and Karen S. Ingersoll

INTRODUCTION

Most drug treatment programs in the United States have been developed by men, on men, and for men. Although drug treatment has been available to women, it rarely has been based on women's special needs. As a partial consequence, fewer women than men receive drug treatment. Not only are women an underserved population, but they also experience greater problems with retention once they begin treatment.

Numerous barriers contribute to the high dropout rate of women from drug programs. Unfortunately, many drug agencies will not continue women in treatment when they become pregnant because of the agencies' inability to provide the prenatal and postnatal medical services and other special support services that pregnant and postpartum women need. Some agencies are concerned about the legal liabilities related to treating pregnant women, especially those who inject drugs and may be positive for the human immunodeficiency virus (HIV). Sedative-hypnotic withdrawal can be safely accomplished during pregnancy, but opioid withdrawal among pregnant women is controversial (Jarvis and Schnoll 1994). There is no pharmacologic treatment available to wean women from cocaine. However, despite the availability of medical treatment, many programs serving women have inadequate funds to support a medical staff.

ECONOMIC BARRIERS

Many drug-abusing women drop out of treatment because of economic problems. First, drug treatment is not always affordable to women, especially those on public assistance and those with few monetary resources who require treatment that is either free or inexpensive. Drug treatment, especially residential care, can cost thousands of dollars and may not be covered by medicaid. Women with alcohol or other drug addiction often find few places where they can seek treatment, and the agencies that are accessible to them usually have long waiting lists.

Second, adequate child care and transportation are critical for most women who have young children, work outside the home, and attempt to stay in drug therapy. In a demonstration study, the Phoenix, Arizona, project, the most frequent reasons given by women for discontinuing drug therapy sessions were the lack of child care and transportation. If the women work, they have low-paying jobs, often part time, with no flextime, no job security, and few or no health benefits (Lewis 1992).

When women are not only the mothers of young children but also the primary caretakers of elderly parents or disabled husbands or partners, they usually cannot enter, much less remain in, residential (inpatient) treatment programs. Significant others and nuclear family members may resist a woman's staying in residential (or even outpatient) treatment because this disrupts the family system.

LOGISTICAL BARRIERS TO RETENTION

Child Care Barriers

Lack of child care is a major barrier to many women continuing in drug treatment. Substance-abusing women who head single-parent households often have low incomes, unstable housing, and few social resources to help with child care. Unless drug programs provide child care services, many mothers are unable to continue in treatment. Providing onsite child care affords mothers not only the means to continue attending treatment but also the opportunity to learn parenting skills and improve parent-child bonding.

The Phoenix project's intensive outpatient program found that mothers could not find adequate, affordable child care (Lewis 1993) and that the women's husbands or partners were often unavailable or unwilling to provide the necessary child care during the women's 11 outpatient treatment hours each week. Many women had no other family members available to help care for their children. Sometimes, even when family help was available, women were reluctant to use that help because the resulting dependence could keep them involved in unhealthy ways with their families, especially if family members did not support the women's efforts toward recovery.

Retention of mothers in treatment was greatly improved in the Phoenix project once free babysitting services and transportation were provided to all women clients (Lewis 1993). In contrast, the Landover, Maryland, and Richmond, Virginia, projects provided child care for both their outpatient and residential programs from the beginning of the projects. These services

proved to be a significant factor in attracting clients because many women did not want to be separated from their children during treatment.

Many drug treatment agencies have no history of providing child care to their clients. In Phoenix, administrators were told that there were insurmountable obstacles to providing onsite child care. The process of obtaining State child care certification was long and expensive. However, if parents remained onsite during the babysitting service, the agency was not required to be licensed. Similarly, in the Richmond project at the Virginia Commonwealth University Center for Perinatal Addiction (CPA), onsite babysitting was provided when the mothers were onsite. However, this project became State licensed despite the bureaucratic obstacles.

Another impediment to providing affordable babysitting to the Phoenix mothers was that babysitting costs had not been requested in the original grant from the National Institute on Drug Abuse. Fortunately, project staff members eventually recruited a small cadre of community volunteers for babysitting. These efforts were quickly rewarded by noticeable increases in women's treatment retention (Lewis 1993).

In summary, Finnegan (1979, pp. 121-131), Beschner and Thompson (1981), and others, in describing the unique needs of women (especially pregnant women) in drug treatment, have strongly recommended providing child care, socialization and support skills for mother-child bonding, nursery services for high-risk infants born to substance-abusing mothers, parenting skills training, outreach, and followup care for mothers and children. In short, gender-sensitive programs that provide child care services have found that this option not only improves retention but also enhances the effectiveness of the program as a whole (Reckman et al. 1984).

Transportation Barriers

Lack of transportation also influences treatment retention for many women. Drug programs rarely are found in the women's own neighborhoods. Public transportation is not always provided in suburban areas, and traveling at night, even in one's own vehicle, can be dangerous. The cost of providing transportation for a mother and her children can be prohibitive, but providing clients with free or low-cost, door-to-door transportation often improves retention.

In the Landover project, a 15-passenger van was purchased to take clients to and from the outpatient program. A car telephone also was bought and found to be useful in alerting clients to the time that the van was arriving to pick them up. For those clients who did not have access to a telephone,

a scheduling pattern was established so that the women knew when to expect the van.

As a result of available transportation and other resources, the Landover project was viewed by its main client referral source as "the best game in town." This was because the women were not put on a waiting list, did not need medical insurance to participate, and were provided child care and transportation services to and from the treatment programs and even for home visits.

The Phoenix project, in contrast, inherited transportation problems characteristic of many large, urban agencies, namely, insufficient public transportation and great distances between a woman's community and the treatment center. Although a public bus system existed, greater Phoenix bus routes were relatively few. In addition, scheduled buses were often not available in the evenings when many mothers were free to attend treatment meetings.

Although several Perinatal-20 projects provided some clients with transportation by hiring or buying vans, the Phoenix project did not have such resources and therefore had to turn instead to buying bus tokens and developing relationships with taxicab companies through a voucher system for the occasional transportation of certain clients—women who were unable to pay anything toward their transportation to drug treatment. CPA also provided bus tickets for treatment-related transportation and taxicab vouchers for women in medical crises or in labor.

OTHER BARRIERS TO TREATMENT

Programs that create a user-friendly environment increase the likelihood that women will access available services and remain in treatment. However, child care and transportation are not the only factors that can influence treatment dropout for this population. Environmental, patient, and program variables also play roles.

Impoverished Environments—Housing

For example, during the 5 years of CPA's operation, the majority of patients came from impoverished environments characterized by high rates of drug use, crime, and family violence. As substance abuse clinicians know, individuals who are successful at recovery must frequently remove themselves from situations that place them at risk for relapse. However, this is a difficult task for addicted women who lack the necessary resources for independent living.

This problem was addressed through provision of "transitional housing" for addicted women entering intensive outpatient treatment at CPA. Housing was made available on a first-come, first-served basis to anyone whose circumstances warranted her separation from her home environment. This included women who had actively drug-using or abusive family members or partners and those who were homeless or living in shelters. Women were able to bring their children with them to the transitional housing unit (THU). In this way, families were kept together so that parenting skills could be taught and the foster placement of project children could be avoided. The women who received transitional housing services also obtained enhanced case management from residence staff members who assisted them in locating permanent, safe, drug-free housing. Aside from their limited stay in the THU (during pregnancy and up to 3 months postpartum), these women received the same medical, psychotherapeutic, and support services (child care and transportation) as those in intensive outpatient treatment.

Two studies on the housing variable were conducted by the Richmond project (Haller et al., in press; 1993, p. 303). In both studies women's retention rate was positively influenced by their residing in the program operated by the THU. In the first study (n=48), only 20 percent of those not receiving housing services were still in the program at the 60-day mark compared with nearly 60 percent of those residing in the THU. In the second study (n=52), the dropout (survival) curve leveled off at 50 percent after 80 days in treatment for those living in the THU, whereas it plateaued at 20 percent at day 60 and then fell to 10 percent by day 120 for those not living in the THU. The relative risk for dropout for those living in the community was 2.21 times that of those living in the THU. It should be noted that, from a clinical perspective, the THU is an obvious "draw" to the clinical programs, facilitating recruitment as well as retention. Therefore, the waiting list for acceptance into the THU tends to be long. However, those women scheduled for admission to the THU are more likely to start treatment than are those not expecting to enter the THU.

The problem with this finding is that such supervised, drug-free housing for addicted pregnant women and their children is expensive and not typically available. In addition, once women are ready to leave the THU (even while they remain engaged in intensive outpatient treatment), they cannot find acceptable housing within the community. Most subsidized housing units are located in undesirable neighborhoods. The prospect of returning patients to their pretreatment living environment is equally unacceptable. Those who return home are viewed as being at higher risk for relapse and treatment dropout. Accordingly, followup housing was given much attention in the CPA program. Staff members considered it

to be the patients' primary problem. Residence staff members and case managers now work with various community agencies to create housing opportunities for their patients to secure their continuation in treatment and to sustain recovery. The recommendation of staff members is for drug treatment programs to form partnerships with local housing authorities that will establish facilities (group homes, shared apartments, etc.) specifically for perinatal patients and their dependents. This allows the agencies to work cooperatively with CPA and with women who are actively engaged in a recovery program.

Patient Barriers

Self-Volunteers vs. Court Referrals. The patients' legal status also may affect retention. Interagency cooperation is frequently important for women drug abusers who are involved with the criminal justice system. Several studies of substance abusers have found legal coercion to positively affect retention (Collins and Allison 1983; McFarlain et al. 1977). Recently, CPA discovered that those with a "clear" legal status were more likely to become treatment dropouts than those who were either on probation or parole or who had court dates pending (Haller et al. 1993, p. 303). The survival curve for those with legal involvement suggested that about 40 percent of these patients were retained through day 220. In contrast, only about 10 percent of those not involved in the legal system were retained by day 120. Although CPA did not accept court-mandated patients, it did take those who were court referred. Often, these women knew that they would be jailed if they dropped out of treatment. Therefore, the positive influence of legal involvement suggests that treatment providers who wish to be successful (and those interested in the recruitment and long-time retention of research subjects) should align themselves with the courts and legal system.

Patient Demographics. Although several demographic factors have been found to influence the treatment retention of substance abusers, findings have varied among populations and studies. Of particular relevance are findings that women as a group (Sansone 1980) and unemployed, single African-American women who are polysubstance abusers (Steer 1980) have lower treatment retention rates.

In the Richmond project, researchers identified several demographic variables that seemed to predict retention. For their sample of mostly female, single African-American cocaine/marijuana/alcohol/nicotine users, age and number of prior treatment attempts were associated with retention, whereas educational level and neighborhood quality (with regard to crime and drugs) were not (Haller et al. 1993, p. 303). Younger patients and those who were inexperienced with treatment were more likely to leave

treatment prematurely. Indeed, nearly 60 percent of those older than age 29 were retained at day 200, whereas all patients younger than 23 years had dropped out by day 120.

In the Landover project, clients younger than 26 years were most likely to drop out of the project compared with the 26- to 30-year-olds or those older than 30. Sixty-eight percent of those younger than 26 years old remained in the project by the sixth month (n=42, p=0.007), and 89 percent of the 26- to 30-year-olds remained in the project by the sixth month (n=30, p=0.002). Interestingly, 70 percent of the older-than-30 clients remained in the project; however, this was not significant. Clients younger than 26 were more likely to leave treatment by the sixth month of participation than were clients ages 26 to 30, who were most likely to remain.

Similarly, CPA found that nearly 60 percent of those with three or more prior treatment episodes were retained at day 200 (no dropout was evidenced after day 60), whereas all the first-time patients had left treatment by day 60. It may be that the younger, more treatment-naive patients had not suffered sufficient consequences of addiction to convince them of the need for intensive treatment. For those in treatment for the first time, the program may have been viewed as too restrictive. These findings are of particular importance because treatment-matching based solely on clinical need may fail if age and number of prior treatment attempts are not taken into account. Thus, even if a woman is deemed to need intensive treatment, it might be necessary to make adjustments in the recommended treatment plan to maintain her in the treatment delivery system.

One additional predictor of retention in the CPA project seemed to be a woman's particular pregnancy trimester. For instance, those who were admitted to CPA in their first trimester were least likely to drop out, whereas those who had entered postpartum were most likely to leave treatment prematurely. Of those who entered treatment early in pregnancy, 60 percent were retained, compared with fewer than 10 percent who entered postdelivery. This suggests the need to recruit pregnant addicts into treatment as early as possible when motivation to remain drug-free through the rest of the pregnancy is greater.

Program Barriers

Detoxification. The Phoenix project spent considerable effort examining various program-related barriers to the retention of women in its outpatient treatment program. An analysis of the women's attrition, by stage of treatment, revealed numerous sources of attrition over time. First, only slightly more than one-third (38 percent) of the Phoenix women who received an intake interview into the project completed 1 or more weeks

of drug detoxification and proceeded into the next treatment phase, despite the provision of free detoxification services for all women who needed them.

Fortunately, high relapse rates were not grounds for research staff apathy. Instead, concerted efforts led to increasing client-friendly measures for all project women during their detoxification experience, including one-on-one contact by a specified counselor instructed to give personal attention and support to each woman passing through the outpatient detoxification laboratory. These supportive efforts by the counselors seemed to account for a decided increase in retention at this stage. That is, there was a 24-percent improvement in retention over the first year.

Waiting List Barrier. Only 42 percent of women entering the Phoenix project and successfully completing detoxification during the first year were retained long enough to remain through the waiting list period and finally be assigned to a group for drug abuse therapy and perhaps couples therapy.

The waiting period was much shorter than the typical 1 to 6 months for women who requested drug abuse therapy in the first agency to cooperate with the Phoenix project. The agency provided a pretherapy introduction program (group socialization program) for those awaiting therapy instead of the typical no-therapy waiting list. However, it was apparent that even a waiting period of a few weeks reduced the list of women who had requested treatment. Therefore, the Phoenix project hired a female intake worker/counselor who spent substantial time "joining with" the women immediately after their intake or detoxification. She nurtured them throughout the weeks they attended the pretherapy orientation meetings while they waited for a slot to open in a drug treatment group. This female counselor made weekly contacts with each woman and served as a supportive resource. The researchers discovered that this additional supportive effort kept some of their women clients from dropping out during this early stage of pretreatment. In quantitative terms, during the second year of its use, Phoenix project retention rates were nearly doubled (a 44-percent improvement over the first-year retention rate) after the supportive efforts of pretherapy treatment were initiated (MacDermid et al. 1994).

Gender-Sensitive Programming. During the second year of the Phoenix project, the research team instituted more gender-sensitive programming, including the provision of babysitting and transportation. The team also considered the option of offering women-only drug treatment groups at the cooperating agency because it was apparent that some women were not comfortable in coeducational therapy groups.

As one of the women told the program coordinator, "I just cannot relax being around so many men who have extensive body tattoos." The idea for this option came from the suggestions of several women clients who thought that more of their special needs could be met in a women-only treatment group. However, administrators at the agency had concerns about developing such groups. In particular, they were worried that a women-only group would draw women away from their existing coeducational groups. In lieu of developing women-only groups, the research team worked with the agency's administration to provide an additional treatment group that would meet in the evenings. This new mixed-sex treatment group quickly provided many new treatment slots for women who had been unable to attend daytime programs and apparently helped increase the retention of many women clients (MacDermid et al. 1994).

CONCLUSIONS

Staff Suggestions

In the CPA project, a meeting of the entire staff (including psychiatric nurses, social workers, psychologists, substance abuse counselors, residence staff members, and child care staff members) was held to summarize the major retention barriers for women and offer solutions for these problems. According to meeting attendees, the single most important factor for aiding the retention of women was adequate housing (both during and after treatment). The therapists' top priority on a "wish list" was for the city to develop low-income, semisupervised, recovery-oriented housing (such as Oxford House, a self-governed, aftercare house for persons in recovery) that could accommodate female addicts and their children on a long-term basis. Currently, the only Oxford House is costly and does not permit children. Ideally, facilities would also provide cooperative child care. Thus, a 2- to 3-year stay in such housing would help an addicted woman complete her education, begin a viable career path, and become self-sufficient, ultimately breaking the cycle of dependence on society, improving the individual's self-esteem, and maintaining her sobriety.

The second recommendation to prevent dropout was to provide child care for all treatment-related activities, which the CPA project had been doing. Offering clinical services to addicted mothers without offering child care was considered "ludicrous" and "unrealistic" by CPA staff members. Providing child care sends several messages: that the treatment team acknowledges the patients as mothers, values them and their children, understands and accepts the context of the mothers' lives, and understands that treatment for women needs to be different from treatment for men.

However, housing and child care are both expensive to provide. Therefore, CPA staff members also recommended many less costly but meaningful additional incentives for retention. For example, the staff members described perinatal substance abusers as deprived, both emotionally and materially. Because the women have a history of impoverishment, tangible things are meaningful to them. One therapist thus recommended that perinatal addiction programs give frequent material incentives. Specifically, staff members recommended providing women with material items that meet real needs but that carry therapeutic messages as well. For example, condoms should be distributed freely and frequently throughout the treatment week. Practically, these protect against pregnancy and sexually transmitted diseases, including HIV. Metaphorically, they indicate that CPA staff members value the women and want to help protect them.

Another recommendation was that new patients be given an appointment calendar on the first day of treatment, a calendar similar to those kept by many professionals. Practically, the calendar can help to orient an addict to her treatment schedule and help her to learn to better organize her time and appointments. Metaphorically, the calendar says "You belong somewhere and are important enough to have a schedule of appointments." The calendar serves an additional function if it is filled with affirmations, for example, 12-step sayings or spiritual encouragements.

One suggestion, which was debated among staff members, was to provide hot lunches to participants and their children on treatment days. All staff members agreed that providing meals would likely enhance retention; however, some members were concerned about possible unintended effects. That is, some staff members felt that, even if women attended treatment merely to be fed, this would afford them a better chance of staying in recovery. Others felt that feeding patients would foster overdependency and passivity.

Finally, staff members recommended that treatment programs employ some personnel who are in recovery. This can enhance an addict's perception that the therapists understand them because of common experiences. Similarly, employing staff members of representative ethnic backgrounds also was seen as important in providing a welcoming atmosphere and role models for the achievement of recovery.

Overall, the multidisciplinary staff at CPA felt that the provision of housing and child care was essential to retaining perinatal substance abusers in treatment. In addition, they recommended providing tangible items to patients as a means of reducing dropout and enhancing participation in

treatment. The gift items were seen as having both practical and symbolic value.

Recommendations for Programs

This chapter reviews the barriers to retention of women in drug treatment programs. Many of these barriers are economic or logistical (housing, child care, transportation), patient related (legal status, age, experience with treatment, stage of pregnancy), or program related (availability of medical detoxification, gender-sensitive vs. generic treatment, waiting lists). To avoid the grim but common attrition of women in drug treatment, the authors recommend the following: Drug treatment programs should anticipate and meet the special, multiple needs of women, be affordable, be client friendly, and be gender sensitive and culture sensitive.

REFERENCES

Beschner, G., and Thompson, P. *Women and Drug Abuse Treatment: Needs and Services*. National Institute on Drug Abuse Services Research Monograph Series. DHHS Pub. No. (ADM)81-1057. Rockville, MD: National Institute on Drug Abuse, 1981. 25 pp.

Collins, J.J., and Allison, M. Legal coercion and retention in drug abuse treatment. *Hosp Community Psychiatry* 34(12):1145-1149, 1983.

Finnegan, L.P. Women in treatment. In: DuPont, R.L.; Goldstein, A.; and O'Donnell, J., eds. *Handbook on Drug Abuse*. Rockville, MD: Department of Health, Education, and Welfare, Alcohol, Drug Abuse, and Mental Health Administration, National Institute on Drug Abuse, 1979.

Haller, D.L.; Elswick, R.K., Jr.; Dawson, K.S.; Knisely, J.S.; and Schnoll, S.H. Retention in treatment of perinatal substance abusers. In: Harris, L., ed. *Problems of Drug Dependence, 1992: Proceeding of the 54th Annual Scientific Meeting. The College on Problems of Drug Dependence, Inc*. National Institute on Drug Abuse Research Monograph 132. NIH Pub. No. 93-3505. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1993.

Haller, D.L.; Knisely, J.S.; Elswick, R.K., Jr.; Dawson, K.S.; and Schnoll, S.H. Perinatal drug abusers: Substance abuse treatment. Some factors influencing treatment retention. *J Subst Abuse Treat,* in press.

Jarvis, M.A., and Schnoll, S.H. Methadone treatment during pregnancy. *J Psychoactive Drugs* 26(2):155-161, 1994.

Lewis, R.A. “Barriers to the Treatment of Women Drug Abusers.” Keynote address to the annual meeting of the Ohio chapters of the American Association for Marriage and Family Therapy, National Council on Family Relations, and Sex Information and Education Council of the United States, Dublin, OH, September 25, 1992.

Lewis, R.A. “Overcoming Barriers to the Treatment of Drug-Abusing Women.” Paper written for the Central European Conference on Women, the Family, and Social Change. Warsaw, Poland, and Prague, The Czech Republic, May 16-22, 1993. 21 pp.

MacDermid, S.; Dickey, M.; and Lewis, R. “Clients’ Attendance as a Function of Women’s Retention in Drug Treatment Programs.” Unpublished manuscript, Purdue University, 1994. 29 pp.

McFarlain, R.A.; Cohen, G.H.; Yoder, J.; and Guidry, L. Psychological test and demographic variables associated with retention of narcotic addicts in treatment. *Int J Addict* 12(2-3):399-410, 1977.

Reckman, L.W.; Babcock, B.G.; and O’Bryan, T. Meeting the child care needs of the female alcoholic. *Child Welfare* 68:541-546, 1984.

Sansone, J. Retention patterns in a therapeutic community for the treatment of drug abuse. *Int J Addict* 15(5):711-736, 1980.

Steer, R.A. Psychosocial correlates of retention in methadone maintenance. *Int J Addict* 15(7):1003-1009, 1980.

AUTHORS

Robert A. Lewis, Ph.D.
Distinguished Professor of Family Studies
Child Development and Family Studies
Purdue University
106 Child Development and Family Studies Building
West Lafayette, IN 47907-1267
(317) 494-2931 (Tel)
(317) 494-1037 (Fax)
lewisr@cdfs.cfs.purdue.edu (Internet)

Deborah L. Haller, Ph.D.
Associate Professor of Psychiatry, Medicine, and Anesthesiology
Division of Substance Abuse Medicine
(804) 828-9925 (Tel)
(804) 828-9906 (Fax)
dhaller@gems.vcu.edu (Internet)

Karen S. Ingersoll, Ph.D.
Assistant Professor of Psychiatry
(804) 828-7456 (Tel)
(804) 828-9906 (Fax)
kingersoll@gems.vcu.edu (Internet)

Medical College of Virginia
Virginia Commonwealth University
P.O. Box 980109
Richmond, VA 23298-0109

Doreen Branch, M.S.
Research Associate
National Public Services Research Institute
Suite 220
8201 Corporate Drive
Landover, MD 20785
(301) 731-9891 (Tel)
(301) 731-6649 (Fax)

The Psychosocial History: An Interview for Pregnant and Parenting Women in Substance Abuse Treatment and Research

Marilee Comfort and Karol A. Kaltenbach

INTRODUCTION

Psychosocial evaluations are commonly required of individuals at enrollment in substance abuse treatment and research programs. The purpose of these assessment procedures is to describe the demographics, personal characteristics, pertinent history, and current biopsychosocial status of those seeking admission to treatment. The information can be used to determine eligibility for a treatment program or research project or to route clients to the services that best suit their needs. In addition, intake data can be used to establish a baseline description of client for treatment planning and review or of the treatment population for program accountability and evaluation.

Several instruments have been developed to document the psychosocial characteristics of clients (Haller and Ingersoll, this volume), including the extent of alcohol or other drug use. Many assessment tools that measure substance use were based primarily on male subjects, who represented the predominant treatment population at the time the tools were developed (Selzer et al. 1975; Skinner and Horn 1984; Wanberg and Horn 1983). One of the most widely used instruments, the Addiction Severity Index (ASI) (McLellan et al. 1980), was tested initially with a population of male armed services veterans. The ASI was designed to use a minimal set of items across repeated assessments to provide information for clinical treatment planning and treatment evaluation. These two purposes necessitated measurement of a wide range of baseline data and potential outcomes in six areas of personal functioning that might be influenced by substance abuse treatment. The comprehensive nature of the ASI and its capacity for followup assessment make it an attractive instrument for treatment and research settings that are concerned with the multiple problems experienced by individuals who abuse substances.

The most recent edition of the ASI (McLellan et al. 1992) includes new items about important life outcomes, such as physical and sexual abuse and long-term personal relationships. It has been tested with several special populations, including pregnant women, drug-dependent inmates,

and dually diagnosed substance abusers. The ASI has been used widely in substance abuse treatment studies with various populations (McLellan et al. 1992). It was adopted as a measurement instrument by community demonstration projects funded by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) in 1988, which served homeless alcohol- or other drug-dependent men and women (Argeriou and McCarty 1990). Several Perinatal-20 research demonstration projects, funded by the National Institute on Drug Abuse (NIDA) in 1989 and 1990, adapted the ASI as a measurement instrument for pregnant and parenting drug-dependent women. Before incorporating the ASI into their protocols, a few of these projects added supplementary questions that broadened the scope of the instrument to address more fully the special needs of women (Comfort et al. 1991; Ridlen et al. 1990, pp. 99-109; Brown 1990). This chapter describes the development of the supplemented version of the ASI for women, the Psychosocial History (PSH), developed at the Family Center at Thomas Jefferson University in Philadelphia.

RATIONALE FOR DEVELOPING THE PSYCHOSOCIAL HISTORY

During the past 15 years numerous conferences and publications have addressed the special needs and characteristics of drug-dependent women of childbearing age who are undergoing substance abuse treatment (Beschner et al. 1981; Blepko 1991; Kilbey and Asghar 1992; Mathias 1995; Reed et al. 1982). Reed and Leibson (1981) studied the differences in finances, criminal activity, social support, and living situations between black and white women who attended women's demonstration treatment programs and mixed-gender programs. Subsequently, Reed (1987) investigated the reasons for the limited progress in offering gender-sensitive treatment programs for drug-dependent women who demonstrate behavioral patterns and coping styles different from those of men. Treasure and Liao (1982, pp. 137-212) concluded that substance-abusing women need to focus on self-management, self-development, self-esteem, and self-confidence. These personal improvements are fundamental to the development of functional lifestyles that require new tasks and activities in daily living, improvement of interpersonal skills, and examination of gender-role expectations. Hagan and colleagues (1994) discussed the cycle of dependence faced by urban women of low socioeconomic status. The realities and frustrations associated with basic survival for these women and their families may contribute to self-medicating escape by means of substance abuse.

Assessment of women's personal histories may reveal the presence of dysfunctional family relationships during childhood that can interrupt

healthy personal and family development and lead to aberrant patterns of daily adult functioning. Woodhouse (1992, p. 262) noted common "themes of violence, male dominance, dependence, motherhood issues and depression" throughout life history interviews, focus groups, "life line" drawings, and written exercises she conducted with women in substance abuse treatment. Root (1989) proposed the role of sexual victimization and the use of substances in mitigating posttraumatic stress disorder. Treatment failures may reflect a lack of understanding of the long-lasting sequelae of symptoms following such traumatic experiences. Regan and coworkers (1987) suggested that the history of violence and abuse associated with the lifestyles of female substance abusers, their family, and social environments, and heightened depression place these women and their children at risk for parenting problems, child abuse or neglect, and foster placement.

In a report on drug-exposed infants, the U.S. General Accounting Office (1990) enumerated barriers to women's substance abuse treatment based on information obtained from drug treatment, health, and social service providers. The findings focused on external social and systemic barriers, such as gender and cultural insensitivity, negative community and professional attitudes, fear of prosecution for child abuse, limitations in insurance coverage, and lack of transportation and child care. In a qualitative study of women's perceptions of treatment effectiveness (Nelson-Zlupko et al., in press), women in recovery identified the particular needs that influenced their engagement and outcomes in treatment. These included child care; parenting support; information on and discussion of concerns regarding sexuality, victimization, and gender roles; and assistance with interpersonal relationships.

As part of a NIDA research demonstration project, women in treatment for cocaine dependence were filmed as they described their personal stories of addiction, treatment, and recovery (Kaltenbach et al. 1994). These life story videotapes revealed personal barriers to treatment, with common themes related to addiction and female socialization. Addiction issues such as denial, medication of feelings, control and responsibility, the need for informal support, and the view of relapse as failure recurred as focal points throughout the treatment process. They were accompanied by issues concerning female socialization, such as dependence; expression of feelings; the viewing of a woman's needs as secondary; management of personal, interpersonal, and family responsibilities; and the seeking of treatment for the benefit of others.

To engage women successfully in substance abuse treatment, programs must elicit individual client perspectives, acknowledge their value, and facilitate resolution of needs and issues voiced by clients. Women's

characteristics and the circumstances they face in daily life must be addressed concurrently with addiction in substance abuse treatment programs. To address women's special needs and strengths, substance abuse treatment programs must assess women comprehensively as they enter and proceed through the programs.

CONTEXT FOR DEVELOPMENT OF THE PSYCHOSOCIAL HISTORY

Family Center is a treatment program that offers comprehensive medical and psychosocial services for pregnant and parenting drug-dependent women and their children. Services are provided at both an intensive outpatient program and a long-term residential treatment program, with parent-child centers at both facilities. Family Center uses a multidisciplinary approach, with a team that includes an obstetrician, psychiatrists, nurses, therapists, certified addiction counselors, case managers, and early childhood specialists.

At Family Center psychosocial evaluation is part of a multidisciplinary assessment process conducted at enrollment for treatment and research. Each woman is introduced to the program services, participates in a psychosocial interview, and signs consents for treatment and research with an intake counselor. She also receives a prenatal nursing assessment and a physical examination by an obstetrician. A psychiatric interview is scheduled for each outpatient and residential applicant. A woman who requests entry to the residential program attends an interview with the program staff, during which she visits the residence and is screened for abilities to care for herself and her children and to accept the responsibilities of living in a community environment. Followup psychosocial interviews are conducted with women who participate in research studies at Family Center.

Family Center's outpatient and residential programs together maintain an average monthly census of 100 clients. The women are pregnant (typically early in the second trimester) or parenting at enrollment. The majority reside in the Philadelphia metropolitan area. Approximately 45 percent of the women are African-American, 40 percent are Caucasian, 10 percent are Hispanic, and 5 percent are biracial. They range in age from 19 years to the early forties, with the majority in the 25- to 34-year-old range. Most of the women have completed 11 years of education, and nearly 90 percent have had job experience. The majority of the women receive public and medical assistance and are single heads of households with two or three children per family. Typically, half the children are living with their mothers.

Approximately 70 percent of the clients at Family Center have abused opiates in addition to alcohol and other drugs (e.g., barbiturates, benzodiazepines, amphetamines, propoxyphenes, cocaine, marijuana, nicotine) and are currently maintained on methadone. The primary drug of choice for nearly 25 percent of the clients is cocaine, often used with alcohol, marijuana, and/or tobacco. The other 75 percent are being treated for abuse of opiates or alcohol. At enrollment in Family Center, 87 percent of the women report previous episodes of alcohol or other drug treatment. The majority share living accommodations with family members or partners and their children. Intake interviews with cocaine-dependent clients indicate that 78 percent of the women have been victims of violence during their lifetimes. Rape (50 percent) and domestic violence (48 percent) are the leading types of violence reported.

The average length of time in outpatient treatment for clients at Family Center is 9 months for opiate-dependent women maintained on methadone (range=1 week to 42 months) and 5.8 months for cocaine-dependent women (range=3 weeks to 22 months). Average retention time for cocaine-dependent women in the residential program is 6.3 months (range=2 weeks to 16 months).

STEPS IN DEVELOPING THE PSYCHOSOCIAL HISTORY

The developers of the ASI have emphasized that their semistructured interview can be adapted to the informational needs of programs and to the styles of individual interviewers and clients (McLellan 1992; National Institute on Drug Abuse 1993). They also have encouraged development of supplemental questions that meet the needs of specific populations (A.T. McLellan, personal communication, June 1990; McLellan 1992). With the knowledge and support of the ASI creators, the PSH was designed to meet the clinical requirements of a comprehensive drug treatment program for pregnant and parenting women as well as comply with the research purposes of the NIDA Perinatal-20 research demonstration project. Both the treatment program and the research project required a baseline description of the pregnant woman, her history, current status, and individual needs as she enrolled for services.

Although the ASI tapped all the fundamental areas of life functioning, it seemed to stop short of asking many essential questions concerning women, such as pregnancy status and history, housing and caregiving for children, status of intimate relationships, and history of violence and victimization. Family Center clinical and research staff members developed an instrument that retains the fundamental structure, administrative procedures, and original items of the ASI but expands

on the measure by probing areas relevant to barriers to treatment, women's needs and strengths, and substance abuse treatment outcomes. Examples of supplemental areas of interest include family history and relationships, relationship with partner, responsibilities for children, current health problems, pregnancy history, drug use during pregnancy, previous treatment experiences, history of violence and victimization, and family legal issues.

The development of the PSH required several steps. Initially, instruments were gathered from other research projects that had used the ASI, in adapted form, to assess women entering substance abuse treatment. The Families of Recovering Mothers (FORM) project, funded by NIAAA in Philadelphia (Comfort et al. 1991), and the Family Life Program, a NIDA Perinatal-20 project in Boston, had created substantive modifications that were relevant to the Family Center population. The FORM project integrated new items into the ASI and administered a separate housing and family interview (Comfort et al. 1991), whereas the Family Life Program supplemented primarily the medical and drug sections of the ASI to include pregnancy information (E. Brown, personal communication, July 1990). Items adapted from each of these instruments were woven into the appropriate sections of the PSH to create a logical conversational flow during the interview. The format of the PSH is considerably longer than the ASI because, in addition to extra items, coding instructions and space to record the responses to open-ended narrative questions were incorporated into the instrument. New patient and severity ratings are created to represent the additional areas of life functioning that are included in the PSH.

The original version of the PSH was first administered in November 1990 to pregnant women entering the Cocaine, Pregnancy and Progeny (CPP) project at Family Center, a member of the Perinatal-20 project. This research demonstration project compared the characteristics and outcomes of pregnant cocaine-dependent women and their infants participating in residential or outpatient treatment. The ASI administration manual (McLellan et al. 1985) and training videotape formed the basis of training for the PSH administration. The videotape prepared for the third edition of the ASI (McLellan et al. 1985) was used initially, accompanied by discussion of the manual and role-play activities involving the supplementary PSH items for women. After training was completed, the PSH was integrated into the Family Center intake process for cocaine-dependent women involved with the CPP project. Ongoing discussion continued among the intake specialist, counselors, and the research coordinator to clarify the intent of ASI items and supplementary PSH items and to refine coding details.

The next step in the development of the PSH began when Family Center clinical staff members expressed an interest in extending administration of the PSH to all clients who enrolled at Family Center as a step toward quantifying intake information for the total program. Research and clinical staff members agreed to collaborate in revising the PSH for joint use. It was necessary to retain the items essential to describe the heterogeneity and multiple needs specific to pregnant cocaine-abusing women, as proposed for the research demonstration project. At the same time, information had to be incorporated that had been collected by means of referral forms and a narrative psychosocial interview to create treatment plans and to fulfill the State licensing requirements of the Pennsylvania State Office of Drug and Alcohol Programs (ODAP). The feedback from extensive review and negotiation of PSH versions within Family Center were coupled with outside reviews by an ASI consultant (D. Zanis, personal communication, March 1992), the Philadelphia City Office of Drug and Alcohol Programs, and the Pennsylvania State ODAP (J. Peterson, personal communication, January 1991) to create a revised PSH in 1992. The Family Center intake specialist began administering the revised PSH to all women who enrolled at Family Center in March 1992.

The change in psychosocial assessment procedures from a narrative to a semistructured, coded interview presented advantages and disadvantages for the staff, clients, and program. The PSH provided extensive quantitative and qualitative information for research and clinical use about each woman's personal characteristics, needs, substance use, and life situation as well as descriptive data on the service population in general. It also established a standard set of questions on which data were collected at intake. Unfortunately, the PSH required double the usual psychosocial assessment time from clinical staff members and clients, which in turn decreased the potential number of intake interviews per month. In addition, the increased length of the instrument necessitated many hours of data entry and analysis to obtain meaningful information for client description, program evaluation, and research. Weighing the advantages against the disadvantages, the investigators decided to allocate resources by administering the PSH to all women at Family Center yet entering data in computerized databases only for women enrolled in the CPP project so that both clinical and research commitments were met. Meanwhile, use was made of the experience gained from PSH interviews to refine the instrument by gleaning feedback from intake and research staff members and from data entry personnel and counselors as they gathered information from clients, coded responses, and developed treatment plans. Minor additions, clarifications, and reformatting of the PSH have continued periodically to incorporate new items from the fifth edition of the ASI and suggestions from clinical and research staff members as program evaluation requirements have evolved.

TRAINING FOR THE PSYCHOSOCIAL HISTORY ADMINISTRATION

Family Center was invited to introduce the women's supplements to the ASI on a training videotape developed by NIDA as part of a technical training package (National Institute on Drug Abuse 1993). The module offers two videotapes of simulated ASI interviews and a step-by-step resource manual that details administration and scoring procedures. The ASI training videotapes provide an introduction and comments interspersed with role-plays of ASI interviews by male and female interviewers and clients of various cultural backgrounds, life histories, and interpersonal styles. Excerpts from a PSH interview demonstrate the assessment of the client's pregnancy history, relationship with the father of the baby, current housing arrangements, and responsibilities for children. A brief discussion of the PSH modifications underscores the need to assess the multifaceted problems and strengths of women who participate in substance abuse treatment and research.

Training for PSH interviewers at Family Center begins with the ASI training videotapes. These are viewed along with careful study of the ASI administration manual (McLellan et al. 1990). In addition, specialized training includes discussion of how to organize and ensure the flexibility of the PSH interview to respond to the personal styles and individual capabilities of clients. This session also includes instructions for developing interviewer severity ratings and coding PSH items. Staff members are paired with experienced PSH administrators to observe a PSH interview in process.

PSYCHOSOCIAL HISTORY/ADDICTION SEVERITY INDEX RELIABILITY AND VALIDITY STUDY

Family Center has made the PSH available in its preliminary versions to interested treatment programs and research projects across the country, including the NIDA Perinatal-20 projects and several projects funded by the Center for Substance Abuse Prevention. Before distributing the PSH more widely, Family Center has initiated a collaborative study with the authors of the ASI at the Center for Studies of Addiction, University of Pennsylvania/Philadelphia Department of Veterans' Affairs Medical Center. The study examines the reliability and validity of the PSH in relation to the ASI and tries to determine whether the supplementary items for women provide more complete information about women's treatment needs. In preparation for this study, the authors designed a more focused version of the PSH based on the data and experience gathered during 3 years of development and use of the instrument. This version of the PSH was

based on the recommendations of a committee of Family Center clinicians and researchers who had at least 2 years of experience in using the PSH and on suggestions from another researcher who had used portions of the instrument in a neighboring Philadelphia perinatal addiction program (S. Course, personal communication, January 1994). The revisions take into consideration the content of the ASI, significance of information for clinical or research purposes, ease of item administration and recording of responses, clients' responses to questions, and clarity for data entry, analysis, and interpretation of results.

EXCERPTS FROM THE PSYCHOSOCIAL HISTORY

The PSH supplements to the ASI pertain to substance use during pregnancy, family legal issues, history of violence and victimization, budgeting needs, employment history, perinatal medical status, family history, partner relationships, child care, and current housing arrangements. For most of these topics, additional questions were integrated into each section of the ASI. When several additions were required for a specific topic, such as pregnancy history, child care, and housing arrangements, separate PSH sections were created and interspersed with sections of the ASI to ensure a logical flow in the psychosocial interview. The PSH includes open-ended questions as well as categorical items. As they do with the ASI, interviewers are encouraged to rephrase items, probe the client's responses, and comment freely on the record form to ensure that the intent of the questions and the responses are clearly understood. Excerpts from supplementary PSH items are listed in appendices 1, 2, and 3 at the end of the chapter.

ILLUSTRATIONS OF SELECTED PSYCHOSOCIAL HISTORY DATA

The following data were collected during PSH interviews with 64 pregnant cocaine-dependent women as part of intake appointments at Family Center. They have been selected to illustrate PSH information that supplements that of the ASI. Distinctions are made between data gathered with ASI and PSH items.

Analysis of ASI demographic data showed that the women had a mean age of 28 years and were 89 percent African-American, 9.5 percent Caucasian, and 1.5 percent Hispanic. The women reported an average of 11.3 years of education, and 43 percent had received job or technical training. PSH items showed that 86 percent had been employed at some point in their lives, but 83 percent currently were receiving public assistance. The usual

employment patterns in the past 3 years, according to a combination of ASI and PSH data, were 42 percent unemployed, 32 percent employed full time, and 24 percent employed part time; 2 percent were in a controlled environment with no opportunity to work. PSH responses indicated that 30 percent of those unemployed chose not to look for work because of pregnancy, child-rearing responsibilities, or other reasons. With regard to current marital status, the ASI indicated that 4.9 percent were married, 67.2 percent were never married, and the remainder were separated, divorced, or as coded on the PSH, in a long-term relationship (longer than 1 year) with a partner.

Responses to ASI questions concerning substance use and treatment indicated that during their lifetimes women had used, on average, alcohol for 7.7 years, cocaine for 6.2 years, and marijuana for 5.3 years. In addition, the PSH supplements showed that, at the time of the interview, 73 percent of the women admitted to smoking cigarettes and had done so for an average of 12 years. The women had experienced an average of 2.4 previous episodes of treatment, with 83 percent reporting at least 1 episode. Substance use during pregnancy is of particular interest to maternal addiction treatment programs because of the potential effects on the fetus. In response to PSH items, women reported days of use during pregnancy as 73.5 for tobacco, 43.4 for cocaine, 16.9 for alcohol, and 5.3 for marijuana.

The living arrangements of substance-abusing women tend to be transitory or dependent and may involve domestic crises or violence. Therefore, it is important to understand both past living situations, as documented in the ASI, and current living arrangements that are probed in the PSH. At intake, ASI data indicated that 64 percent of the women had been dissatisfied with their living arrangements during the past 3 years. With the addition of several PSH response codes, the data showed that the women usually had lived with partners (10 percent), parents or other family members (8 percent), alone (8 percent), or in several of these living arrangements over the 3-year period (75 percent). Responses to PSH questions regarding current housing demonstrated that 59 percent of the women were dissatisfied with their current housing situations. These living arrangements were more varied than the 3-year patterns—46 percent lived with parents, other family, or friends; 19 percent lived with partners; 19 percent lived in shelters or residential programs; 7 percent had no stable housing; and the remainder reported several of the aforementioned or other living situations.

The legal status section of the PSH includes questions regarding victimization to assess the history of violence experienced by drug-dependent women, in addition to the history of charges and arrests

queried by the ASI. ASI data from cocaine-dependent women at Family Center revealed that 28 percent reported ever having been charged or arrested and 16 percent having been incarcerated. Supplemental information from the PSH indicated that the women had experienced an average of two types of victimization in their lifetimes. These women represented 81 percent of the respondents. The most frequent types of victimization reported include domestic violence (55 percent), rape (49 percent), child abuse/neglect (36 percent), sexual exploitation (30 percent), assault (22 percent), and robbery (21 percent). Thirty-eight percent of the women reported having family-related legal problems.

For pregnant and parenting women, the PSH devotes a section of the interview to questions regarding pregnancy, prenatal care, and history of complications. Cocaine-dependent women at Family Center reported being an average of 19 weeks into pregnancy at enrollment. Their responses showed mean values of gravida to be 4.9, parity 2.4, therapeutic abortions 0.94, and spontaneous abortions 0.60, and they recalled an average of 1.3 previous pregnancies with complications. Of the 42 women who responded to this item, 61 percent had received some prenatal care prior to enrollment at Family Center. When asked how they felt about the current pregnancy, most women discussed assorted ambivalent feelings. Fifty-nine percent were happy to be having a baby, 48 percent felt that this baby would provide a chance to make their lives worthwhile, and 42 percent were happy to be pregnant. On the other hand, 22 percent did not want another baby, 22 percent disliked being pregnant, 28 percent were upset about the conception circumstances, 71 percent felt guilty about using drugs during the pregnancy, and 69 percent worried about the baby having problems.

CONCLUSION

The PSH data presented in this chapter illustrate the breadth of information available during an initial interview through administration of this supplemented version of the ASI. Substance abuse is usually only one of numerous problems that exist in the lives of alcohol- or other drug-dependent women. Their strengths are equally important to identify as treatment begins. A comprehensive assessment is essential to promote treatment and research tailored to the individual needs and strengths of women and their children. The PSH was developed to serve both clinical and research purposes. For clinical use it organizes information on current and historical psychosocial and health status into a single instrument and provides abundant material to aid in the understanding of a woman's history and needs. The PSH thereby facilitates identification of strengths, problem areas, and goals for treatment.

At the most elementary level, PSH data provide researchers with detailed descriptive information about individual women that can be aggregated as needed for program evaluation and research investigations in conjunction with other assessments. These client data also may identify psychosocial and economic realities related to enrollment, use of services, and retention of women in treatment and their ability to achieve and maintain recovery. Although the PSH requires an investment of staff time in training, client interviews, and data entry and analysis, in return it provides a detailed portrait of the characteristics, history, and needs within multiple areas of the lives of drug-dependent women seeking substance abuse treatment.

REFERENCES

Argeriou, M., and McCarty, D., eds. *Treating Alcoholism and Drug Abuse Among Homeless Men and Women: Nine Community Demonstration Grants*. New York: Haworth Press, 1990. 164 pp.

Beschner, G.M.; Reed, B.G.; and Mondanaro, J., eds. *Treatment Services for Drug Dependent Women*. Vol. 1. National Institute on Drug Abuse Treatment Research Monograph Series. DHHS Pub. No. (ADM)81-1177. Rockville, MD: National Institute on Drug Abuse, 1981. 503 pp.

Blepko, C., ed. *Feminism and Addiction*. New York: Haworth Press, 1991. 224 pp.

Brown, E. "Maternal Measures." Discussion conducted at the National Institute on Drug Abuse Perinatal-20 Treatment Research Demonstration Program Meeting, Bethesda, MD, November 9, 1990.

Comfort, M.; Hagan, T.; and Kaltenbach, K. "Psychosocial History." Unpublished assessment, Family Center, Thomas Jefferson University, 1995.

Comfort, M.; Richlin, L.; Shipley, T.E.; White, K.; and Shandler, I. "Final Report of the F.O.R.M. Project." Diagnostic and Rehabilitation Center of Philadelphia. National Institute on Alcohol Abuse and Alcoholism grant R18-AA-07966. October 1991. 56 pp.

Hagan, T.A.; Finnegan, L.P.; and Nelson-Zlupko, L.L. "Impediments to a Comprehensive Drug Treatment Model for Substance Abusing Women: Treatment and Research Questions." Unpublished manuscript, Family Center, Thomas Jefferson University, 1994. 24 pp.

Kaltenbach, K.; Comfort, M.; and Balis, S. "Final Report: Cocaine, Pregnancy and Progeny Project: Treatment and Evaluation." National Institute on Drug Abuse grant R18-DA-06363. Report submitted to the National Institute on Drug Abuse, 1994. 98 pp.

Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children.* National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. 399 pp.

Mathias, R. NIDA expands its research on addiction and women's health. *NIDA Notes* 10(1):1,4-5, 1995.

McLellan, A.T. "Planning Committee Document for the ASI Technology Transfer Package." Discussion at Planning Session for National Institute on Drug Abuse ASI Training Package, Washington, DC, March 16-17, 1992.

McLellan, A.T.; Cacciola, J.; Griffith, J.; McGahan, P.; and O'Brien, C.P. *Guide to the Addiction Severity Index: Background, Administration, and Field Testing Results*. 3d ed. National Institute on Drug Abuse Treatment Research Report. DHHS Pub. No. (ADM)88-1419. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute on Drug Abuse, 1985. 51 pp.

McLellan, A.T.; Kushner, H.; Metzger, D.; Peters, R.; Smith, I.; Grissom, G.; Pettinati, H.; and Argeriou, M. The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat* 9:199-213, 1992.

McLellan, A.T.; Luborsky, L.; O'Brien, C.P.; and Woody, G.E. An improved evaluation instrument for substance abuse patients: The Addiction Severity Index. *J Nerv Ment Dis* 168:26-33, 1980.

McLellan, A.T.; Parikh, G.; Bragg, A.; Cacciola, J.; Fureman, B.; and Incmikoski, R. "Addiction Severity Index Fifth Edition Administration Manual." Unpublished manuscript, Center for Studies of Addiction, University of Pennsylvania/Philadelphia Veteran's Administration Medical Center, 1990. 42 pp.

National Institute on Drug Abuse. *Assessing Client Needs Using the ASI: Resource Manual and Training Videotapes*. NIH Pub. No. 93-362. Rockville, MD: U.S. Department of Health and Human Services, 1993. 150 pp.

Nelson-Zlupko, L.; Dore, M.M.; Kauffman, E.; and Kaltenbach, K. Women in recovery: Their perceptions of treatment effectiveness. *J Subst Abuse Treat*, in press.

Reed, B.G. Developing women-sensitive drug dependence treatment services: Why so difficult? *J Psychoactive Drugs* 19:151-164, 1987.

Reed, B.G.; Beschner, G.M.; and Mondanaro, J., eds. *Treatment Services for Drug Dependent Women.* Vol. 2. National Institute on Drug Abuse Treatment Research Monograph Series. DHHS Pub. No. (ADM)82-1219. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1982. 618 pp.

Reed, B.G., and Leibson, E. Women clients in special women's demonstration drug abuse treatment programs compared with women entering selected co-sex programs. *Int J Addict* 16(8):1425-1466, 1981.
Regan, D.O.; Ehrlich, S.M.; and Finnegan, L.P. Infants of drug addicts: At risk for child abuse, neglect, and placement in foster care. *Neurotoxicol Teratol* 9:315-319, 1987.
Ridlen, S.; Asamoah, Y.; Edwards, H.G.; and Zimmer, R. Outreach and engagement for homeless women at risk of alcoholism. In: Argeriou, M., and McCarty, D., eds. *Treating Alcoholism and Drug Abuse Among Homeless Men and Women: Nine Community Demonstration Grants*. New York: Haworth Press, 1990.
Root, M.P. Treatment failures: The role of sexual victimization in women's addictive behavior. *Am J Orthopsychiatry* 59:542-549, 1989.
Selzer, M.L.; Vinokur, A.; and van Rooijen, L. A self-administered Short Michigan Alcoholism Screening Test (SMAST). *J Stud Alcohol* 36(1):117-126, 1975.
Skinner, H.A., and Horn, J.L. *Alcohol Dependence Scale User's Guide.* Toronto: Addiction Research Foundation, 1984.
Treasure, K.G., and Liao, H. Survival skills training for drug dependent women. In: Reed, B.G.; Beschner, G.M.; and Mondanaro, J., eds. *Treatment Services for Drug Dependent Women.* Vol. 2. National Institute on Drug Abuse Treatment Research Monograph Series. DHHS Pub. No. (ADM)82-1219. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1982.
U.S. General Accounting Office. *Drug-Exposed Infants: A Generation at Risk.* A report to the chairman, Committee on Finance, U.S. Senate. Washington, DC: U.S. General Accounting Office, 1990. 48 pp.
Wanberg, K.W., and Horn, J.L. Assessment of alcohol use with multidimensional concepts and measures. *Am Psychol* 38:1055-1069, 1983.
Woodhouse, L.D. Women with jagged edges: Voices from a culture of substance abuse. *Qual Health Res* 2:262-281, 1992.

AUTHORS

Marilee Comfort, Ph.D., M.P.H.
Research Director
Family Center
Assistant Professor of Pediatrics
(215) 955-1954 (Tel)
(215) 568-6414 (Fax)
comfort1@jeflin.tju.edu (Internet)

Karol A. Kaltenbach, Ph.D.
Director
Family Center
Clinical Associate Professor
Departments of Pediatrics and Psychiatry and Human Behavior
(215) 955-4068 (Tel)
(215) 568-6414 (Fax)

Thomas Jefferson University
Ninth Floor
1201 Chestnut Street
Philadelphia, PA 19107-4192

APPENDIX 1. *Excerpts from psychosocial history family: Housing and children*

4a. How many adults (including client) live in your current household? _____

4b. How many children? _____

4c. Total number of people in household: _____

5. What type of housing do you presently live in? _____
   a. apartment
   b. house
   c. condominium
   d. room
   e. shelter
   f. supervised group living
   g. no stable living arrangements
   h. other, specify _____

6. How large is the current place where you live?
   a. Total # rooms: _____
      (Count kitchen, dining & living rooms, bedrooms, NOT baths)
   b. Total # of bedrooms: _____

6.1. Do you pay the full cost of your current housing and utilities?
(Check applicable levels of payment and assistance below.
Also explain source of housing assistance.)

_____ NA, no current housing

_____ NA, no housing payments required, explain _____

	a. Level of Housing Payments	**b. Level of Utility Payments**	**c. Comments on Source and Type of Assistance (Note person, organization; type of voucher, subsidy or grant)**
Full	_____	_____	_____
Partial or Reduced	_____	_____	_____
None	_____	_____	_____

Housing History Matrix
(for items #7 - 12)

Housing Past 3 Years (Begin with most recent housing.)	**Length of Time**	**With Whom?**
1. _____		
2. _____		
etc.		

7. How many places have you lived in the past 3 years? _____

8. With which other adults have you *usually* lived in the past 3 years? (Circle all that apply.)
   a. alone
   b. with my parents
   c. with other family members (not parents)
   d. with father of this baby
   dd. with the father of (at least) one of my children
   e. with another sexual partner (not FOB, not father of child)
   f. with friends
   g. in a residential program (e.g., jail, hospital), specify _____
   h. no stable arrangements
   i. in shelter(s), specify_____
   j. other, specify _____
   k. several of the above (Circle k. and all that apply.)

9. How long did you live in the longest of these arrangements?
   yrs. _____ mos. _____

10. Have you been satisfied with any of these living arrangements over the past 3 years?
   _____ Yes, explain _____
   _____ No, explain_____
   _____ Indifferent, explain _____

11. Have you been without a place to stay (homeless) in the past 3 years? (Define homeless as having no place to stay for more than 1 month.)
   _____ Yes _____ No

12. Where did you stay during your homeless period? _____ NA (Circle all that apply.)
   a. in shelter(s)
   b. with extended family
   c. with friends
   d. in a car
   e. in a crack house
   f. in a vacant building
   g. outside on the street
   h. other, specify _____

13. If currently homeless, how long has it been since you had a place to live that you consider to be a permanent home?
   yrs. _____ mos. _____ wks. _____ NA _____

SOURCE: Comfort et al. 1995

APPENDIX 2. *Excerpts from the psychosocial history medical status/ pregnancy history*

8. How many times in your life have you been hospitalized for problems with *past* pregnancy(ies)?
   ____ number of times____ NA (if 1st pregnancy)

9. How many pregnancies have you had with complications? _____
   ____ NA (never pregnant)

10. Please tell me about the complications.
    ____ NA (never pregnant)

    (Please specify complications below from first to last pregnancy. Be sure that all pregnancies are accounted for. Please use back of page if there are more than 10 pregnancies.)

	Specify Problem	**No Comp.**	**No Preg.**
a.	Preg. 1 Complication: _____	_____	NA _____
b.	Preg. 2 Complication: _____	_____	NA _____
	etc.		

11. How many of your babies have had serious health problems at birth or later in childhood?
    ____ NA (never pregnant) ____ NA (first pregnancy)
    a. Total at Birth ____ b. Total Later in Childhood ____

12a. Please tell me about your babies' health problems at birth.
    ____ NA (never pregnant)

    (Please specify from oldest child [Child 1] to youngest child. Please be sure that all births, live and stillborn, are accounted for. Use back of page if there are more than 10 children.)

	Specify Problem	**No Comp.**	**No Preg.**
a.	Child 1's problem: _____	_____	NA _____
	_____	_____	(1st Preg.)
b.	Child 2's problem: _____	_____	NA _____
	etc.		

12b. Please tell me about your children's health problems later in childhood.
    ____ NA (never pregnant)

    (Please specify from oldest child [Child 1] to youngest child. Please be sure that all children are accounted for. Use back of page if there are more than 10 children.)

	Specify Problem	**No Comp.**	**No Preg.**
a.	Child 1's problem: _____	_____	NA _____
	_____	_____	(1st Preg.)
b.	Child 2's problem: _____	_____	NA _____
	etc.		

13a. Have you been hospitalized overnight during this pregnancy at all?
    Yes____ No____ NA ____ (not pregnant)

13b. If yes, why? _____ NA ____

14a. Have you had any prenatal care for this pregnancy yet?
    _____ Yes _____No

14b. If yes, where? _____ NA ____

15. How far along in your pregnancy were you when you started receiving prenatal care?
    _____ number of weeks _____ NA (not pregnant)

16. How many prenatal visits have you attended so far for this pregnancy?
    _____ number of visits
    _____ NA (pregnant, but haven't started prenatal care)
    _____ NA (not pregnant)

17. Did you use drugs and/or alcohol during pregnancy with any of your children?

    (Please specify from oldest [Child 1] to youngest child. Be sure that all births, live and stillborn, are accounted for.)

	NA	NR	DK	Yes	No
a. Child 1 ____	____	____	____	____	____
b. Child 2 ____	____	____	____	____	____
etc.					

18. How do you feel about this pregnancy? (Circle letters of all that apply.)
    ____ NA (not pregnant)
    a. don't want (another) baby
    b. feel guilty about using drugs (and alcohol) during my pregnancy
    c. feel worried about the chance of my baby having problems
    d. dislike being pregnant
    e. happy to have baby
    f. upset due to circumstances of conception
    g. feel that this baby is my chance to make my life worthwhile
    h. happy to be pregnant
    i. other, specify: _____
    j. several of above (Circle j. and all letters that apply.)

For the following, ask Client to use the Client Rating Scale (0-4 scale)

19. How concerned have you been about these medical problems related to your pregnancy in the past 30 days? ___ ___ NA (not pregnant)

20. How important to you NOW is treatment for these medical problems related to your pregnancy? ___ ___ NA (not pregnant)

Interviewer Severity Rating (0-9 Scale)
(defined as need for treatment *beyond* what client is currently receiving)

21. How would you rate the client's need for prenatal care?
    _____ _____ NA (not pregnant)

Confidence Ratings

Is the above information significantly distorted by:

22. Client's misrepresentation? Yes ____ No ____

23. Patient's inability to understand? Yes ____ No ____

24. If yes for #22 or 23, please explain: _____

SOURCE: Comfort et al. 1995

APPENDIX 3. *Excerpts from psychosocial history legal status*

2.How many times have you experienced one of the following:
(If impossible to list number of times, note as continuous over extended period and explain)

_____ NA (never experienced any of the following)
a. _____ assault
b. _____ arson
c. _____ rape
d. _____ domestic violence
e. _____ child abuse, neglect
f. _____ robbery
g. _____ sexual exploitation (forced to provide sexual favors)
h. _____ other, specify

3/4. In the past year, how often have you:

3. Provided sex for money _____
   a. _____ never
   b. _____ once
   c. _____ 2-5 times
   d. _____ 6-10 times
   e. _____ once per month
   f. _____ 2-3 times per month
   g. _____ 1-2 times per week
   h. _____ 3-5 times per week
   i. _____ once a day or more

4. Traded sex for drugs _____
   a. _____ never
   b. _____ once
   c. _____ 2-5 times
   d. _____ 6-10 times
   e. _____ once per month
   f. _____ 2-3 times per month
   g. _____ 1-2 times per week
   h. _____ 3-5 times per week
   i. _____ once a day or more

5a.Have you ever gotten an order of protection? _____Yes _____No

5b.If yes, how many have you gotten? _____ _____NA

Comments:

6a. Do you have any legal problems *now* involving your family?
(e.g., Family or Juvenile Court issues, child custody involving DHS, divorce, domestic violence, etc.)
____ Yes, current ____ Yes, past ____ No

6b. Please describe the nature of CURRENT problems and department
(e.g., DHS, Family Court, etc.)
_____ NA

6c. Please describe the nature of PAST problems and department
(e.g., DHS, Family Court, etc.)
_____ NA

SOURCE: Comfort et al. 1995

Assessment of Perinatal Substance Abusers: Experiences of One Perinatal-20 Project

Deborah L. Haller and Karen S. Ingersoll

INTRODUCTION

This chapter identifies and discusses issues relevant to the psychosocial evaluation of perinatal substance abusers. The conclusions are based on 5 years of experience with 133 women treated at Virginia Commonwealth University's National Institute on Drug Abuse (NIDA)-funded Treatment Research Unit, the Center for Perinatal Addiction (CPA). The chapter begins with a brief overview of measurement methodology relevant to clinical research and then describes the initial rationale for test selection and presents some measurement problems that have been encountered. After delineating weaknesses in the test protocol, the authors recommend adjustments and suggest areas of exploration for future clinical research projects.

GENERAL ISSUES OF MEASUREMENT

The first step in selecting any assessment instrument is to clarify the purpose of testing (Anastasi 1982; Cronbach 1984). In a research project, the purpose will emanate from the research questions. For example, an investigation of the base rates for mental illness in a given population might necessitate use of an instrument that makes categorical psychiatric diagnoses (e.g., the Structured Clinical Interview for *DSM-III-R* [SCID] [Spitzer et al. 1988]) using the *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised) (DSM-III-R)* (American Psychiatric Association 1987). If a researcher wishes to address change in functioning as a consequence of treatment, a different kind of measurement tool is indicated, one that yields scalable data (e.g., the Minnesota Multiphasic Personality Inventory-2 [MMPI-2] [Hathaway and McKinley 1989] and the Millon Clinical Multiaxial Inventory-II [MCMI-II] [Millon 1990]). Thus, for instrument selection to be most effective, research objectives must be specific. The investigator also must consider the psychometric properties of available tests. The two most important properties are reliability (consistency of the test) and validity (usefulness of the test). Anastasi (1982) gives a thorough explication of these issues; a brief review follows.

Reliability

Three types of reliability are relevant: internal consistency, temporal stability, and interrater agreement. If a test purports to measure a single domain, its internal consistency is important. This concept is less relevant to longer measures with multiple scales; however, good internal consistency within scales is expected. Tests also should possess moderate-to-high test-retest reliability, when readministered within a short timeframe, if their aim is to measure an abiding trait. In contrast, test-retest reliability is irrelevant when a test is used to measure a state rather than a trait. If observational or rater-scored measures are used, interrater reliability becomes important. For example, scoring of the SCID relies on the judgment of trained raters who complete a lengthy interview with the subject. If interrater reliability is poor, the investigator would have little confidence in the diagnoses generated because they would be more reflective of the raters than of the subject.

Validity

Once a test is found to possess adequate reliability, validity must be considered. Validity is the degree to which a test measures what it purports to measure. Four types of validity are relevant: construct-related, criterion-related, content, and face validity. *Construct-related* validity is the degree to which a test measures a hypothesized construct, such as "addiction." Construct-related validity is inferred from examining the relationships among tests purporting to measure the same construct (providing convergent validity) and those purporting to measure different constructs (providing discriminant validity). Stated simply, if a given measure of addiction correlates well with established tests of addiction and does not correlate well with tests of less related constructs (e.g., intelligence), it probably possesses adequate construct validity. *Criterion-related* validity generally refers to a test's predictive ability. Tests found to have good predictive validity may be useful as screening tools to match patients to treatments. *Content* validity is usually established by a panel of expert raters who "judge" that a given test adequately measures a domain. For example, to be useful, an achievement test should accurately reflect an individual's performance. *Face* validity is the test's acceptability to the user and the subject as a relevant-appearing instrument. Face validity is clinically important because research subjects quickly tire of tests that seem irrelevant to their lives. Subjects are more likely to enact a negative response set (a test-taking stance or attitude that influences responding, such as exaggerated responding in an attempt to convince others of severity of symptoms) or refuse to continue testing if the instrument lacks face validity. On the other hand, face-valid tests can lead to defensive or exaggerated responding by subjects who wish to create a

particular impression. In some instances, a non-face-valid (i.e., empirically derived) instrument, such as the MMPI-2, may be of greater utility for these reasons.

Standardization

Identifying appropriately normed, well-standardized tests that are suitable for studying perinatal substance abusers is difficult. Few appropriately normed instruments are available to study women. There are even fewer such tests to target women who are pregnant, belong to an ethnic minority, or are addicted to cocaine. Most standardized tests were developed using older males as subjects. Early studies often failed to include women in the standardization group or to provide descriptive statistics for the standardization sample by sex or race.

In general, researchers should build on previous research findings rather than design new instruments; however, this is not always possible in the perinatal substance abuse field. Accordingly, some investigators have developed new tests that tap domains of particular relevance to this population, whereas others have modified existing instruments to meet their needs. Findings for pregnant substance abusers from instruments established with men (such as the Addiction Severity Index [ASI] [McLellan et al. 1980]) also are being collected and published; this means that population-specific norms should be available in the future. It is crucial that limitations in measurement of perinatal substance abusers be discussed in the literature and in public forums so that investigators can make informed decisions regarding psychosocial assessment of this population. Along these lines, Haller and colleagues (1993*a*) recently noted possible sex-related differences for perinatal substance abusers on the MMPI-2 and MCMI-II, two widely used objective personality inventories. However, until large samples are collected from multiple sites, interpretations of many tests must remain tentative.

Test Administration Issues

Other issues of importance in selecting measurement instruments include test version, ease of administration, cost, and ease of scoring and interpretation. Some tests, such as the MMPI (Hathaway and McKinley 1967), have large bodies of established norms and interpretive data available for older test versions. An investigator may be more comfortable with interpretation of an older version of a test, already have a large data set using the older version, or own a computer-scoring program for an older version, making continued use of this version highly attractive. In comparison, another researcher may be unencumbered by such restraints and may therefore elect to use the newest version of the test. When the

Perinatal-20 projects were initiated, it was unclear whether recent test versions for several prominent instruments (including the MMPI and MCMI-II) would assess the same domains or be interpretable using the same assumptions as the older versions. As a result, investigators generally selected test versions for idiosyncratic reasons. It is reassuring that recent studies have found the consistency of MMPI and MMPI-2 profile shapes, elevation, and scatter to be good (Harrell et al. 1992). In addition, most clinicians have rated their interpretations of MMPI and MMPI-2 to be similar (Clavelle 1992), although the issue of codetype[1] equivalence remains unresolved.

Types of Assessment Instruments

Structured interviews, self-report measures, and behavior ratings are the most common types of measurement. All have costs and benefits to be considered, assuming relative equivalence of relevance to the purpose of testing and psychometric standing. *Structured clinical interviews* are among the most costly of techniques, requiring several trained raters who first must achieve solid interrater agreement. Typically, they will spend several hours interviewing a subject and coding the results. Test interpretation relies on the accuracy of coding. Structured interviews often yield truncated, limited information; however, as the result of strict standardization, they generally provide high-quality categorical information. Structured clinical interviews used in the CPA project included the SCID, ASI, and Structured Interview for DSM-III-R Personality-Revised (SIDP-R) (Pfohl et al. 1989).

In contrast, *self-report scales* are low cost, readily available, generally well standardized, and easily administered. Many can be interpreted by a single psychometrician, resulting in lower personnel costs. Computer-scoring and -interpreting programs also are readily available for many of these instruments. However, self-report scales can be vulnerable to subject biases and may not be appropriate for use with subjects with limited verbal skills. A crucial but sometimes neglected issue is the likelihood of subject acceptance or resistance. For example, many subjects balk at completing the MMPI-2 because of its length.

Behavior ratings include observational measures administered by trained raters; informant reports from raters, therapists, and peers; and performance tests in which the subject performs a sample behavior. *Observational measures* are costly in terms of personnel and the need

---

[1]In the MMPI-2, the codetype is determined on the basis of high points (i.e., the scales that are most elevated). These scales define the codetype. For example, someone with the two highest scale elevations on scales 2 and 4 has a 2-4 codetype, with the highest scale named first.

for constant supervision to ensure that standardized procedures are consistently followed. Although they may yield information with better external validity than self-report measures, observational measures can be difficult to interpret. *Informant reports*, although inexpensive, are vulnerable to response sets, inconsistent reporting, and extreme tendencies. *Performance tests* may be more or less costly depending on the complexity of the behavior to be assessed; they are also vulnerable to response sets, coding errors, and lack of standardization.

Summary

Researchers should consider the relative costs and benefits associated with available instruments before making test selections. The first consideration should be the instrument's ability to address the research questions. Next, psychometric properties, specifically reliability and validity, must be considered. Although test standardization that includes the target population is important, norms for perinatal substance abusers are generally unavailable. Similarly, norms for women (particularly ethnic women) are rare. Investigators thus are left with the options of developing new instruments or using existing ones. The advantage of the latter approach is the opportunity for norms to be established for the particular population of interest and for direct comparison of the study population with related populations, such as male substance abusers. The final decision in the test selection process also should take into consideration test version, ease of administration, cost, and ease of scoring and interpretation.

The advantages of investigators working in a given field using the same instruments and the same test versions are obvious. When each project makes independent decisions regarding test protocol, this inadvertently creates a barrier to multisite comparisons and the development of new norms for substance-abusing women based on a large available sample.

DEVELOPMENT OF THE INITIAL CENTER FOR PERINATAL ADDICTION TESTING PROTOCOL

When the Perinatal-20 research projects were initiated, little was known about the target population of pregnant and recently postpartum substance abusers. Accordingly, CPA researchers decided to take a "shotgun" approach to assessment. This strategy was adopted in hopes that sufficient descriptive data would be obtained to characterize this population. A broad-based testing program also was expected to facilitate identification of specific areas where assessment efforts should be focused.

The testing program was designed to collect data relevant to studying treatment outcomes. To accomplish this task, the authors included instruments that previously had been used in outcome studies with related populations. There was special interest in comparing findings for female substance abusers with those for male substance abusers and female psychiatric patients. Accordingly, measures were selected that had been widely used in mental health and substance abuse treatment research (e.g, the ASI, MMPI, SCID, and Beck Depression Inventory [BDI] [Beck 1978]). As a general strategy, the authors focused on tests with good psychometric properties and large bodies of literature to support their efficacy.

Another objective was to compare the psychometric performance of several commonly used instruments purported to measure the same domains. Through this comparison process, the authors hoped to identify specific tests that were both acceptable to subjects and capable of providing consistent, useful information in a cost-effective manner. To accomplish this task, multiple measures of personality disorder, depression, and substance abuse were administered. Unfortunately, at that time, none of the instruments had norms available for perinatal substance abusers. As a general strategy, the authors supported the development of perinatal norms for use with existing tests as a more effective and cost-effective approach than new test development.

The initial domains of interest included cognitive functioning, psychiatric diagnosis (both *DSM-III-R* Axis I [acute disorders] and Axis II [personality disorders] [American Psychiatric Association 1987]), personality functioning, addiction and psychiatric severity, global psychiatric functioning, and family functioning.

Cognitive Functioning

With regard to cognitive functioning, the authors used a screening tool, the Shipley Institute of Living Scale (SILS) (Shipley 1967), which yields an intelligence quotient (IQ) estimate correlating with the Wechsler Adult Intelligence Scale Full-Scale Intelligence Quotient (WAIS IQ) (Zachary 1986). We did not anticipate many subjects with subnormal intelligence in this population. Unfortunately, approximately one-third of the sample evidenced some cognitive impairment (Haller et al. 1993*b*). For these cases, it proved necessary to supplement the screening instrument with a psychologist-administered Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler 1981) to better define the impairment and to potentially qualify the subjects for community-based mental retardation services. This two-tiered assessment strategy has worked reasonably well; it provides general information on those who are functioning within

the normal range of intellectual ability and more specific information about the type and extent of impairment for those who are not. However, administration of the WAIS-R requires a psychologist or skilled psychometrician because test interpretation and scoring are complex, and the entire procedure is both time consuming and expensive.

Psychiatric Diagnosis

From the outset, the authors distinguished between personality functioning or "style" and psychiatric diagnosis (*DSM-III-R*) in the assessment process. One significant drawback to the diagnosis approach is that no information is available for those who fail to meet diagnostic criteria (i.e., who fall below the threshold for detection). Unfortunately, the yes-no categorization typical of diagnosis-oriented instruments such as the SCID or the SIDP-R provides little useful information beyond base rates for the various mental disorders. Also, instruments such as these are not particularly helpful in conducting treatment outcome research because that change is difficult to detect unless someone recovers completely or, conversely, develops a mental disorder during treatment.

The authors employed multiple measures designed to yield psychiatric diagnoses. These included the SCID, SIDP-R, MCMI-II, and MMPI-2. The SCID is a structured clinical interview that identifies Axis I psychiatric disorders. Similarly, the SIDP-R is a structured clinical interview that identifies Axis II psychiatric disorders. In contrast, the MCMI-II comprises numerous scales paralleling *DSM-III-R* diagnoses, both Axis I and Axis II. Finally, scores on several clinical and research scales derived from the MMPI item pool (including the MMPI Personality Disorder Scales [Morey 1988]) were compared with findings for the other tests that measure similar domains (e.g., SIDP-R and MCMI-II).

Data derived from these multiple sources have allowed the authors to begin to characterize the population in terms of psychopathology. Relatively high rates of emotional disturbance were found across the board (Haller et al. 1993*b*). For the substance abuse disorders (and most other Axis I disorders), there has been relatively good agreement across tests, suggesting little need to employ multiple instruments. One exception is the MCMI-II, which appears to underestimate the prevalence of drug disorders (Haller and Dawson 1994). On the other hand, clinical diagnosis of substance abuse disorders was as effective as SCID diagnosis (93 vs. 94 percent). In comparison, non-substance-abuse Axis I disorders were grossly underdiagnosed by clinicians compared with all the assessment tools used (e.g., 12 percent for depressive disorders compared with 29 percent by the SCID and 37 percent by the MCMI-II). The bias against diagnosing non-substance-abuse Axis I disorders is

curious because staff members were trained mental health professionals with at least a master's degree. One explanation is that people who work in the addiction field are conservative about making Axis I diagnoses until they have had ample opportunity to observe the subject drug-free.

In comparison, intertest agreement for Axis II disorders was poor. Base rates for the various disorders fluctuated widely, depending on the instruments being compared (Haller 1995). Interestingly, some personality disorders were measured more reliably than others. For instance, base rates for histrionic personality disorder varied from a low of 2 percent on the MMPI-2 to a high of 63 percent on the MCMI-II; in contrast, the rates for schizoid personality disorder varied between 7 percent on the SIDP-R and 14 percent on the MCMI-II. Overall, clinicians did an exceedingly poor job of identifying Axis II disorders. For example, for antisocial personality disorder (ASP), the base rate for detection was 2 percent for clinicians compared with 33 to 73 percent by test; indeed, the rate of "confirmed" diagnosis (two or more tests agreeing) for ASP was 57 percent. These discrepant findings highlight the importance of using supplemental tests to diagnose Axis II disorders in a clinical setting.

The lack of consistency in measurement of personality disorders is a serious problem. Because investigators continue to publish studies that use different instruments, the literature may be biased. That is, findings may reflect test-sensitivity differences rather than true population differences. Clearly, studies that use different test instruments cannot be compared with confidence. More important, the lack of agreement between test instruments designed to measure the same construct may indicate fundamental difficulties with the personality disorder constructs. Indeed, the current classification system has been criticized for having excessive overlap among personality disorders (Widiger et al. 1986) and for failing to consider the personality disorders as continuous, dimensional constructs (Cloninger 1987).

Personality Functioning

Both the MMPI-2 and the MCMI-II have been widely used to study treatment outcomes in male substance abusers and in both male and female psychiatric patients. Using these instruments allowed the authors to compare findings for substance-abusing women with these reference groups. Both instruments were recently revised; we elected to use the most recent versions (MMPI-2 and MCMI-II) so that changes in procedure would not need to be made once the study had commenced. (Note: The MCMI-III [Millon 1994] is now available.)

Both the MMPI-2 and the MCMI-II have yielded interesting data about personality style for the sample of mostly African-American, cocaine-dependent women. Overall, findings are similar to those reported in the literature for men, with a few important distinctions. Specifically, at intake these women are more psychologically disturbed than their male counterparts, as evidenced by higher test elevations and more pathological profile configurations. However, sex differences appear to diminish with treatment; that is, after these women undergo 5 months of treatment, findings on the MMPI-2 approximate those reported in the literature for male substance abusers (Haller and Dawson 1994). Test profiles for perinatal substance abusers who accept and reject treatment are also different, with refusers looking similar to accepters after 5 months of treatment (Haller et al. 1995). Both tests are useful in assessing personality change as a function of treatment. Significant decreases were evidenced for most scales on both the MMPI-2 and the MCMI-II, and the frequency of certain personality configurations or patterns generated changed considerably from intake to month 5 (Haller and Dawson 1994).

The authors were surprised to experience relatively few difficulties in the administration of lengthy personality inventories. Although many subjects had low general intelligence, all but one subject were able to read sufficiently to complete self-administered instruments (n=155 admissions). Although the test content is probing, the vast majority of subjects were tolerant of this. Approximately 85 percent of self-administered objective personality test protocols were valid at first administration; the remainder were deemed invalid because of subjects' adopting a deviant response set (15 percent of MMPI-2s and 17 percent of MCMI-IIs). It was standard procedure, whenever possible, to retest subjects who generated invalid protocols. At the time of retest, valid MMPI-2 protocols were generated by 42 percent of subjects and valid MCMI-II protocols by 50 percent.

Addiction and Psychiatric Severity

The ASI was used to assess severity of dysfunction in multiple domains. Of particular importance to the psychopathology project were ratings of alcohol, other drugs, and psychiatric severity. Although few administration difficulties have been noted for the ASI, the test may not be as sensitive to change in this population as in others. Nevertheless, in preliminary studies, changes were found from discharge (for treatment completers) to first followup (6 months postdischarge). More specifically, there is a tendency for subjects who have completed treatment to improve in terms of medical, psychiatric, and drug severity. However, the authors are unsure about the test's ability to assess change in other domains. For example, composite scores in the employment domain tend to become higher (meaning more dysfunctional) as a function of treatment. This

suggests that unemployment may not be perceived as a problem until the treatment program identifies it as such. ASI findings to determine the overall appropriateness of this instrument with the perinatal population currently are being evaluated. On the positive side, perinatal norms now have been developed by the instrument's author (A.T. McLellan, personal communication, 1993). Normative information will greatly increase the utility of the test with the perinatal population.

Although the authors do not have personal experience with the Addiction Problem Survey (APS) (DePhilippis et al. 1994, p. 56), we are interested in evaluating this instrument. The APS is a brief, self-administered instrument that assesses severity in the same domains as the ASI. The test's authors (DePhilippis et al. 1994) compared findings for the ASI and APS; initial results have been promising. Therefore, there may be shortcuts to obtaining information similar to that provided by the ASI.

Global Psychiatric Functioning

Global psychiatric functioning was assessed by means of multiple measures, including the Symptom Checklist-90-R (SCL-90-R) (Derogatis 1993), BDI (Beck 1978), and Hamilton Depression Rating Scale (Hamilton 1967). At intake, the SCL-90-R findings reflected high levels of emotional distress on every scale; all elevations subsided with treatment. However, relatively little additional information was obtained through use of the SCL-90-R that was not available from the objective personality tests. Information obtained from the two depression scales was also redundant (identification by BDI, 64 percent; by Hamilton, 54 percent; by both, 43 percent). Clearly, only one depression scale is needed if this domain is to be independently assessed.

Family Functioning

When this work was begun, the authors had little personal experience with the assessment of families and family functioning. Many subjects continued to be involved with their families of origin and also had families of their own; clearly, family relations and interactions were salient. Therefore, instruments were sought that would characterize a woman's family of origin, current adult relationship with a significant other, and current family constellation. Because few measures were available that could address these needs, we selected what appeared to be relevant instruments, despite limited documentation as to their validity and reliability.

The Family of Origin Scale (FOS) (Hovestadt 1985) is a 40-item self-report inventory designed to assess adults' perceptions of the intimacy and

autonomy of the families in which they were reared. Subsequently, doubt was cast on the validity and multidimensionality assumptions of the FOS (Gavin and Wamboldt 1992; Fedor 1992; Manley et al. 1990; Mazer et al. 1990). The authors also have found the FOS to provide only limited information on the population. For example, there is no evidence of change in perinatal substance abusers' views of their families of origin despite treatment. This may be an accurate reflection; on the other hand, it may indicate that the instrument is insensitive to changes that do occur. Also measured were subjects' current family status in areas of competence, cohesion, leadership, and expressiveness using the Self-Report Family Inventory (SFI) (Beavers et al. 1985). This is a relatively well-researched instrument that appears to provide meaningful data to the project. For instance, on the SFI, leadership (the degree to which the parents are "in charge") decreased from intake to month 5 and then increased at followup. This may indicate that perinatal families become more chaotic initially under the stress of substance abuse treatment. However, parental leadership reemerges, and the family appears to reorganize by followup. An alternative interpretation is that of defensive posturing at intake, followed by the subject's admission of problems, and subsequent real improvements in leadership. Because no psychometrically sound instruments were available to measure subjects' primary relationships, a questionnaire was adapted from the Acquaintance Description Form (Wright and Wright 1990) to create the Relationship Assessment Form (RAF) (McCall 1991). The RAF provides information on subjects' tendencies to minimize difficulties, take responsibility, place blame, and display other behaviors relevant to their relationships with spouses or lovers. Data analyses on all family measures are under way; initial results seem promising for the SFI and RAF. However, family measurement is an area that deserves further attention in perinatal substance abuse research.

FUTURE DIRECTIONS FOR ASSESSMENT OF PERINATAL SUBSTANCE ABUSERS

During the past 5 years, the authors identified several significant gaps in our assessment program. Unfortunately, some important domains were completely neglected. These included neuropsychological functioning (including learning disabilities), experience with violence (both as victim and perpetrator), social support, and parenting experience. We also failed to include stages and processes of change data in our initial test protocol. Clearly, the concept of readiness to change behavior should be viewed as an important covariate in psychotherapy outcome research. Hindsight also suggests that information related to the substance use experience should have been collected. More specifically, data on craving and alcohol and other drug expectancies would be extremely useful when trying to

understand who enters, stays, and succeeds in treatment. The influence of other cognitive factors, such as control orientation, probably needs to be considered in a well-balanced test protocol as well. In essence, by assessing primarily psychiatric domains, the authors ignored normal cognitive processes and learning experiences.

We are attempting to fill these identified gaps by adding to the protocol carefully selected instruments that address population-specific needs. The following instruments are being used in the Center for Substance Abuse Treatment's new Center for Perinatal Addiction-Residential Program (CPA-RP) project protocol or are under consideration for use in other clinical research projects involving this population.

Neuropsychological Functioning

Subjects now complete a brief battery of tests sensitive to brain dysfunction, including the Trailmaking Test, Parts A and B (Trails: The Halstead-Reitan . . . 1985), and the Logical Memory and Visual Reproduction subtests of the Wechsler Memory Scale-Revised (WMS-R) (Wechsler 1987). Also, we have added the reading subtest of the Wide Range Achievement Test III (Jastak and Wilkinson 1993) to obtain a literacy rating. All these are well-standardized tests that have been widely used in both research studies and clinical practice and have norms available for comparison purposes.

Experience With Violence

Few instruments are available that address participation in violent acts, either as victim or perpetrator. We were interested in pursuing the four domains of violation highlighted by the Jefferson Medical College of Thomas Jefferson University in its structured clinical interview: verbal, physical, sexual, and personal freedom. However, a self-administered instrument that focused on recent participation in aggressive activities was needed. Thus, with the assistance of the perinatal group at Jefferson Medical College, we are modifying the Self-Report of Abuse (SRA) (Family Center 1991). The result is the Aggressive Acts Questionnaire (AAQ) (Haller and Woodard 1994), a 40-item instrument that addresses involvement in aggressive acts within the past 30 days only. Scales that assess the four domains of violation, from the perspective of both victim and perpetrator, are included. Norms for males and females and blacks and whites are being obtained for comparison purposes, and a profile sheet depicting extent, frequency, and response to violence is currently under development.

Social Support

Although many measures of social support exist, the majority of these have not been adequately researched. Few have norms, and even fewer have been investigated in a rigorous manner. After reviewing numerous instruments, the authors chose two that were developed by a single research group, the Arizona Social Support Interview Schedule (ASSIS) (Barrera 1981; Heitzmann and Kaplan 1988) and the Inventory of Social Supportive Behaviors (ISSB) (Barrera 1981; Heitzmann and Kaplan 1988). The ASSIS is a structured interview that assesses an individual's perceived social support, actual social support, and satisfaction with received support along with perceived and actual conflict about support. The test measures the domains of material aid, physical assistance, intimate interaction, guidance, feedback, and positive social interactions. The ISSB is a self-administered paper-and-pencil measure of social support. It consists of 40 items rated on a 5-point scale to show how often each event occurred during the preceding month (1=not at all; 5=every day). Factor studies have revealed clusters labeled Guidance, Emotional Support, Tangible Support, Cognitive, Information, and Social Support; the first three account for most of the variance.

Parenting Experience

The Parenting Stress Index (PSI) (Abidin 1992) is designed to identify parent-child systems under stress, the risk for dysfunctional parenting, and the development of emotional pathology in a child. The PSI consists of 120 items answered on a 5-point-type scale. In addition to yielding a total stress score, the PSI generates separate scores reflecting stress in both child and parent domains along with a score for stressful life events. The child domain score represents stress the parent experiences on the basis of the child's temperament and is made up of six subscales: adaptability, acceptability, demandingness, mood, distractibility/ hyperactivity, and reinforces parent. The parent domain score assesses stress caused by the parent's own characteristics and social support and is made up of the following seven subscales: depression, attachment, restriction of role, sense of competence, social isolation, relationship with spouse, and parent's health. The score for stressful life events provides a context for the other scores.

Stages and Processes of Change

Subjects' readiness for behavioral change has been studied extensively in nonpatient self-changers. These studies have resulted in the development of a transtheoretical model of change that implies that universal stages and processes of change across problem behaviors may be more important

in predicting change than either problem severity or psychiatric status (Prochaska and DiClemente 1992; Prochaska et al. 1994). Researchers are only beginning to apply the transtheoretical model to substance abusers, and to cocaine abusers in particular (J.S. Rossi, personal communication, October 1993).

Drug Use Experience

Toward the end of the NIDA-funded CPA project, the authors introduced several measures of alcohol or other drug experience, and we have continued to use these in the CPA-RP. Of special interest were subjects' self-efficacy, locus of control in relation to substance use, and craving experiences. Pilot studies are being conducted using a self-efficacy scale adapted from a smokers' scale (DiClemente 1981); the Drinking-Related Internal-External scale (Donovan and O'Leary 1978), adapted for alcohol and other drugs; and several cocaine craving scales, including the Minnesota Cocaine Craving Scale (Halikas et al. 1991) and the Cocaine Craving Questionnaire (Jaffe et al. 1989). It is planned to add instruments measuring the related area of drug expectancies in the near future. The Cocaine Effects Expectancy Questionnaire (Schafer and Brown 1991) is a promising instrument for this purpose.

SUMMARY AND RECOMMENDATIONS

A significant amount of information about perinatal substance abusers was gained through use of CPA's initial test protocol. In particular, the authors identified the need for routine cognitive assessment. Unless a woman's intellectual capacity is known, it is difficult to design programming that will adequately meet her needs. Although the initial program curriculum was too advanced for many of the program's subjects, knowledge about the level of intellectual functioning allowed for appropriate adjustments to be made. At the same time, the majority of subjects tested in the normal range of intellectual ability. For this reason, indepth intellectual assessment would have been an unnecessary expense. It seems most appropriate to adopt a screening strategy and use specialized tests as a supplement whenever findings are abnormal. In addition, the initial protocol failed to include tests for literacy and learning ability. Accordingly, the initial protocol was expanded to include measures of these domains.

Although structured clinical interviews appear to be a preferred form of instrumentation for many grant agencies, the authors have found instruments such as the SCID and SIDP-R to be of relatively little value when cost-benefit comparisons are made. Although these instruments provide high-quality information about base rates for the various psychiatric

disorders, they are not particularly useful in treatment outcome studies. Perhaps the yes-no categorization that is typical of such interviews is simply too gross a measure, especially when many women do not meet the criteria for psychiatric diagnosis. Also, diagnoses alone do not provide the wealth of clinical information that can be obtained from other types of tests.

The objective, self-report personality inventories seem more useful. Their assets include scores on every subject that can be compared at multiple test administration points, ease of administration, and cost-effectiveness. Computer-scoring and -interpreting programs are available for programmatic support as well. Although change on some psychiatric dimensions, such as depression or anxiety, can be measured easily and reliably using disorder-specific instruments such as the BDI, score declines on these psychiatric dimensions also are reflected on the depression subscales of the multidimensional inventories. Thus, depending on the breadth of a testing program, instruments like the MMPI-2 and the MCMI-II can serve multiple purposes. The authors were pleasantly surprised to learn that most subjects tolerated these tests, despite their length and relative lack of face validity relevant to substance abuse. Reading ability also was not a problem (most personality inventories are written on the seventh-grade reading level), and the ability of most subjects to produce valid test results was generally high.

The authors now believe that the almost exclusive focus on measurement of psychiatric domains was a mistake. This emphasis accounted for neglect of such important constructs as learning disability, social support, stage of change, violence, and the parenting experience in the initial test protocol. We now are taking a more multidimensional approach to the evaluation of perinatal substance abusers. It is hoped that, at some point in the future, we will be able to determine the relative importance of these domains in predicting treatment outcome.

An area that was completely overlooked in the initial test protocol pertained to alcohol or other drug response and cognitions about alcohol or other drug use. For instance, we failed to include measures of craving, alcohol or other drug expectancies, and control orientation. All these may be important mediating variables that should be considered, and the purely psychiatric approach is now seen as too simplistic. Perhaps individuals with certain personality characteristics, who are at particular stages of change and who have a given control orientation and specific responses to alcohol or other drug cues, are most likely to be successful in treatment. A multidimensional approach to evaluation is strongly advocated.

The validity of instruments that assess Axis II disorders needs to be considered further. Unfortunately, there is no accepted standard for

detection of personality disorders. *DSM-III-R* (American Psychiatric Association 1987) criteria are highly behavioral in nature, making personality disorders somewhat difficult to assess via objective inventories that do not assess behavior directly. Similarly, face-valid structured clinical interviews allow opportunity for self-report bias and may underestimate the frequency of personality problems. For these reasons, the authors arbitrarily decided that agreement between at least two instruments (as to presence or absence of a given disorder) would be required for diagnosis of personality disorder to provide a minimum standard of convergent validity. Although this procedure increased the confidence level in diagnosing Axis II disorders, it is likely overly conservative because disorders that are diagnosed on a single test are routinely discounted. This procedure is also costly and time consuming. Either researchers need to use more rigorous assessment standards, such as use of multiple tests (possibly generating false negatives), or there needs to be agreement in the field to use specific instruments so that findings can be compared across future studies.

REFERENCES

Abidin, R.R. *Parenting Stress Index, Form 6.* Charlottesville, VA: Pediatric Psychology Press, 1992.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised).* Washington, DC: American Psychiatric Press, 1987. 567 pp.

Anastasi, A. *Psychological Testing, 5th Edition.* New York: Macmillan, 1982. 784 pp.

Barrera, M. Preliminary development of a scale of social support. *Am J Community Psychol* 9:435-447, 1981.

Beavers, W.R.; Hampson, R.B.; and Hulgus, Y.F. Commentary: The Beavers system approach to family assessment. *Fam Process* 24:398-405, 1985.

Beck, A.T. *Beck Depression Inventory (BDI).* The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1978.

Clavelle, P.R. Clinicians' perceptions of the comparability of the MMPI and MMPI-2. *Psychol Assessment* 4:466-472, 1992.

Cloninger, C.R. A systematic method for clinical description and classification of personality variants. *Arch Gen Psychiatry* 44:573-588, 1987.

Cronbach, L.J. *Essentials of Psychological Testing, 4th Edition.* New York: Harper & Row, 1984. 630 pp.

DePhilippis, D.; Metzger, D.S.; and Navaline, H.A. The Addiction Problem Survey (APS): A valid, self-report measure of addiction problem severity. In: Harris, L.S., ed. *Problems of Drug Dependence, 1993: Proceedings of the 55th Annual Scientific Meeting, The College on Problems of Drug Dependence, Inc., Volume II.* National Institute on Drug Abuse Research Monograph 141. NIH Pub. No. 94-3749. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1994.

Derogatis, L.R. *Symptom Checklist-90-R (SCL-90-R)*. Minneapolis, MN: National Computer Systems, 1993.

DiClemente, C.C. Self-efficacy and smoking cessation maintenance: A preliminary report. *Cognit Ther Res* 5:175-187, 1981.

Donovan, D.M., and O'Leary, M.R. The drinking-related locus of control scale: Reliability, factor structure, and validity. *J Stud Alcohol* 39:759-784, 1978.

Family Center. "The Self-Report of Abuse (SRA)." Unpublished structured clinical interview, Jefferson Medical College, 1991.

Fedor, A.J. A validation of the Family of Origin Scale using alcoholic and mental health patients. *Dissertation Abstracts Int* 52:2411, 1992.

Gavin, L.A., and Wamboldt, F.S. A reconsideration of the Family of Origin Scale. *J Marital Fam Ther* 18:179-188, 1992.

Halikas, J.A.; Kuhn, K.L.; Crosby, R.; Carlson, G.; and Crea, F. The measurement of craving in cocaine patients using the Minnesota Cocaine Craving Scale. *Compr Psychiatry* 32:22-27, 1991.

Haller, D.L. Unpublished data, Virginia Commonwealth University, 1995.

Haller, D.L., and Dawson, K.S. "Changes in Personality as a Function of Intensive Outpatient Therapy: A Study of Perinatal Substance Abusers." Unpublished manuscript, Virginia Commonwealth University, 1994.

Haller, D.L.; Dawson, K.S.; and Ingersoll, K.S. "Treatment Acceptors vs. Rejectors: A Study of Perinatal Substance Abusers." Unpublished document, 1995.

Haller, D.L.; Knisely, J.S.; Dawson, K.S.; and Schnoll, S.H. Perinatal substance abusers: Psychological and psychosocial characteristics. *J Nerv Ment Dis* 181(8):509-513, 1993*b*.

Haller, D.L.; Motley, C.E.W.; and Schnoll, S.H. Personality and addiction, focus on women. In: Watson, R.R., ed. *Addictive Behaviors in Women.* Totowa, NJ: Humana Press, 1993*a*.

Haller, D.L., and Woodard, C. "The Aggressive Acts Questionnaire (AAQ)." Unpublished instrument, Virginia Commonwealth University, 1994.

Hamilton, M. Development of a rating scale for primary depressive illness. *Brit J Soc Psychol* 6:278-296, 1967.

Harrell, T.H.; Honaker, L.M.; and Parnell, T. Equivalence of the MMPI-2 with the MMPI in psychiatric patients. *Psychol Assessment* 4:460-465, 1992.

Hathaway, S.R., and McKinley, J.C. *The Minnesota Multiphasic Personality Inventory Manual*. Minneapolis, MN: National Computer Systems, 1967.
Hathaway, S.R., and McKinley, J.C. *Minnesota Multiphasic Personality Inventory-2 (MMPI-2)*. The Regents of the University of Minnesota. Minneapolis, MN: National Computer Systems, 1989.
Heitzmann, C.A., and Kaplan, R.M. Assessment methods for measuring social support. *Health Psychol* 7(1):75-109, 1988.
Hovestadt, A.J. A Family of Origin Scale. *J Marital Fam Ther* 11:287-297, 1985.
Jaffe, J.H.; Cascella, N.G.; Kumor, K.M.; and Sherer, M.A. Cocaine-induced cocaine craving. *Psychopharmacology* 97(1):59-64, 1989.
Jastak, S., and Wilkinson, G.S. *Wide Range Achievement Test-Revision 3*. Wilmington, DE: Jastak Associates, 1993.
Manley, C.M.; Searight, H.R.; Binder, A.F.; and Russo, R.J. The Family-of-Origin Scale: Factorial validity for adolescents. *Fam Ther* 17(1):75-82, 1990.
Mazer, G.E.; Mangrum, O.L.; Hovestadt, A.J.; and Brashear, R.L. Further validation of the Family of Origin Scale: A factor analysis. *J Marital Fam Ther* 16:423-426, 1990.
McCall, D. "The Relationship Assessment Form." Unpublished instrument, Virginia Commonwealth University, 1991.
McLellan, A.T.; Luborsky, L.; O'Brien, C.P.; and Woody, G.E. An improved evaluation instrument for substance abuse patients: The Addiction Severity Index. *J Nerv Ment Dis* 168:26-33, 1980.
Millon, T. *Millon Clinical Multiaxial Inventory-II*. Minneapolis, MN: National Computer Systems, 1987. 249 pp.
Millon, T. *MCMI-III Manual.* Minneapolis, MN: National Computer Systems, 1994.
Morey, L.C. *The MMPI Personality Disorder Scales: A Manual and Guide to Interpretation.* 1988.
Pfohl, B.; Blum, N.; Zimmerman, M.; and Stangl, D. *Structured Interview for DSM-III-R Personality-Revised (SIDP-R).* Ames, IA: Department of Psychiatry, University of Iowa, 1989.
Prochaska, J.O., and DiClemente, C.C. Stages of change in the modification of problem behaviors. In: Hersen, M.; Eisler, R.M.; and Miller, P.M., eds. *Progress in Behavioral Medicine.* Sycamore, IL: Sycamore Publishing, 1992.
Prochaska, J.O.; Velicer, W.F.; Rossi, J.S.; Goldstein, M.S.; Marcus, B.H.; Rakowski, W.; Fiore, C.; Harlow, L.; Redding, C.A.; Rosenbloom, D.; and Rossi, S.R. Stages of change and decisional balance for 12 problem behaviors. *Health Psychol* 13:39-46, 1994.
Schafer, J., and Brown, S.A. Marijuana and cocaine effect expectancies and drug use patterns. *J Consult Clin Psychol* 59:558-565, 1991.

Shipley, W.C. *Shipley Institute of Living Scale (SILS)*. Los Angeles: Western Psychological Services, 1967.
Spitzer, R.L.; Williams, J.; and Gibbon, M. *Structured Clinical Interview for DSM-III-R: Non Patient Version (SCID)*. New York: Biometrics Research Department, New York State Psychiatric Institute, 1988.
Trails: The Halstead-Reitan Neuropsychological Test Battery. Trail Making Test. Tucson, AZ: Neuropsychology Press, 1985.
Wechsler, D. *Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1981.
Wechsler, D. *Wechsler Memory Scale-Revised (WMS-R)*. The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1987.
Widiger, T.A.; Frances, A.; Warner, L.; and Bluhm, C. Diagnostic criteria for the borderline and schizotypal personality disorders. *J Abnorm Psychol* 95:43-51, 1986.
Wright, P.H., and Wright, K.D. Measuring codependents' close relationships: A preliminary study. *J Subst Abuse* 2:335-344, 1990.
Zachary, R.A. *Shipley Institute of Living Scale: Revised Manual*. Los Angeles: Western Psychological Services, 1986.

ACKNOWLEDGMENT

This research was supported by NIDA grant R18-DA-06094.

AUTHORS

Deborah L. Haller, Ph.D.
Associate Professor of Psychiatry, Internal Medicine, and Anesthesiology
Departments of Psychiatry, Internal Medicine, and Anesthesiology
(804) 828-9925 (Tel)
(804) 828-9906 (Fax)
dhaller@gems.vcu.edu (Internet)

Karen S. Ingersoll, Ph.D.
Assistant Professor of Psychiatry
Department of Psychiatry
(804) 828-7456 (Tel)
(804) 828-9906 (Fax)
kingersoll@gems.vcu.edu (Internet)

Medical College of Virginia Hospitals
Virginia Commonwealth University
Box 980109
Richmond, VA 23298-0109

Kinchart-Sociograms as a Method for Describing the Social Networks of Drug-Using Women

Anitra Pivnick

INTRODUCTION

The overall increase in the rate of human immunodeficiency virus (HIV) infection among women in the United States has been attributed to increases in rates of infection in east coast, inner-city African-American and Latino communities where there are heavy concentrations of injection drug use and high levels of inner-city poverty (Landesman 1989; Chu et al. 1990). In New York City, almost 80 percent of pediatric cases of acquired immunodeficiency syndrome (AIDS) are children of mothers who use drugs or are the sexual partners of men who use drugs (New York City Department of Health 1990).

In an effort to bring an anthropological perspective to investigations of social and cultural influences on drug use, HIV infection, and HIV transmission, in 1987 the Centers for Disease Control funded a study to be conducted in a Methadone Maintenance Treatment Program (MMTP) located in the southwest Bronx, New York. The clinic is sited in an area with high rates of injection drug use, poverty, and HIV infection.

The MMTP serves 700 patients and includes the administration of methadone (a substitute drug for persons with histories of opiate addiction), counseling services, an HIV testing and counseling program, and primary health care. Participants of the study described in this chapter were drawn from a population of 300 women enrolled in the MMTP, most of them residents of the south and west Bronx.

The research was conducted by the author through The Women's Center, a peer support program operating in, but not administered by, the MMTP. The Women's Center was designed to provide drug-using women with a confidential place in which to discuss issues related to drug use, HIV infection, and other matters in a group setting. In addition to providing professionally facilitated weekly peer support groups, the program was conceived as a field station for the investigation of high-risk behaviors associated with drug use and sexual practices. The purpose of both the peer support program and the research was the reduction of the spread of HIV. Project staff members included a nurse-psychodramatist, a health psychologist, two anthropologists, and several peer leaders. Group

participants were fully informed of the research agenda. Risk reduction associated with drug use and HIV prevention education were prominent aspects of the weekly group meetings.

During an initial 6-month observation period, the author (a medical anthropologist) noted that the methadone program had many characteristics of a small community: Many patients were long-time associates, having been enrolled concurrently in various drug treatment facilities; some were related through blood and marriage; and many were coresidents. Those who continued to use illegal drugs were associated through drug use activities.

These observations led to the development of a methodology that would describe group participants' personal social networks, noting kin and nonkin relationships, drug use patterns, and places of residence to generate information relating these factors to HIV transmission (Pivnick 1994). The methodology was derived principally from the ethnographic method of recording kin-based relationships and depended heavily on anthropological concepts of kinship and residence. Recruitment of female clinic patients began with peer support group participants and proceeded outward into the general female methadone population through referrals from peer support group participants and drug counselors and other clinic staff members.

Ultimately, the methodology described in this chapter generated data in three independent studies: (1) a description of the drug use patterns and HIV status within the personal social networks of female methadone patients, including kin (consanguineal and affinal), sexual partners, household members, children's household members (if different from the subject's), friends, persons with knowledge of the subject's serostatus, and drug associates (Pivnick et al. 1994); (2) an investigation of the personal, social, and cultural influences on HIV-positive women's reproductive decisions (Pivnick et al. 1991); and (3) the methodology used in a National Institute on Drug Abuse (NIDA)-funded Perinatal-20 study of the effect of family case management on drug-using women's child custody arrangements and family relations.

ANTHROPOLOGICAL INQUIRY—KINSHIP AND RESIDENCE

Both a description of the places in which people eat, sleep, and congregate and the ways in which these same people perceive themselves to be related are fundamental to anthropological inquiry. When anthropologists initially studied small groups of people living in remote parts of the world, they noted that kinship relationships in these communities form the basis for all aspects of social relations, including marriage, residence, and subsistence

relationships. Working in these bounded communities, anthropologists recognized that knowledge of how people construct their relations with one another is essential to an understanding of how they organize themselves and how they perceive their particular universe. More recently, kinship terminologies have been constructed by some anthropologists as systems of shared meaning exclusive of biological derivatives (Schneider 1987); however, the more general notion recognizes kinship as "the network of relationships created by genealogical connections, and by social ties (e.g., those based on adoption) modeled on the 'natural' relations of genealogical parenthood" (Keesing 1975, p. 13).

The charting of kinship and social relations that forms the basis of the methodology described in this chapter has its origin in classical anthropological studies of kinship (Morgan 1904; Evans-Pritchard 1940; Fortes 1945; Levi-Strauss 1969). In representing kin relations, anthropologists use a standard notation derived from the principles of reproduction in which biological parents are understood to be the genitors of their offspring or in which persons other than biological parents are classified in the social roles of biological parents. Using "ego" (the subject) as the first point of departure and stratifying by generation, anthropologists construct a model of a group member's kin relations from those in closest biological or social proximity, moving outward to include those more and more distantly related (collateral relatives), and including affines (those related through marriage). For anthropologists, an analysis of kinship terminology (descriptive titles and meanings relating individuals and groups), coupled with related social, economic, and political activities, aids in an account of the past, makes the present less ambiguous, and facilitates an understanding of basic behaviors and meanings expressed by members of the group.

The study of residence, including history and patterns, was the second point of departure for anthropologists. Residence, like kinship, was essential to an understanding of the small groups that anthropologists first studied. In these communities, the study of location or residence often intersected meaningfully with the study of kinship, and a consideration of these two ways of describing people's social interactions became the point of departure for ethnographic studies of the culture, social action, and social organization of a group of people.

The conceptualization of residence that most immediately influenced the development of the kinchart-sociogram methodology was predominantly that of Netting and colleagues (1984), who suggested that households (groups of persons who coreside) can be understood not only as the product of kin groups, marriage practices, and residence rules but also as units of residence that are influenced by broader economic, social, cultural, and

political forces. Through a consideration of various economic, social, and political influences on coresidential units, the notion of change is addressed. Thus, alterations in external circumstances are expected to produce changes in residential units. In addition, households (or coresidential units) are conceptualized as the product and the producers of cognitive models or collectively shared meanings related to their forms and content. The emphasis on broader influences affecting residence patterns provided an important perspective in the formulation of a research agenda that was to address the relationships among HIV infection, reproductive decisionmaking, and mother-child coresidence history.

The final contribution to the development of the charting process was that of family systems analysis in which the structure, strength, and character of ties between family members and others are notated in the form of "genograms," diagrams of family relationships and relations. This adaptation of genealogical notation in family therapy has been well documented and is a form of modeling and individual note taking used by many family therapists and family therapy theorists (Penn 1983).

This view of personal social networks, including kin relations and coresidential units viewed as localized groups influenced by broader societal forces, provided a framework for the analysis of drug use, sexual relations, reproductive decisions, and HIV infection that the research described in this chapter was to address.

THE KINCHART-SOCIOGRAM METHODOLOGY

The notation that results in a kinchart-sociogram (figure 1) is produced by a subject's responses to a progression of questions about family members, drug-using associates, and non-drug-using friends and neighbors (see appendix at the end of this chapter). A kinchart-sociogram is constructed by the investigator, who asks the subject a standard sequence of questions designed to elicit relational information in an orderly fashion about people, beginning with siblings, progressing through consanguineal and affinal relatives, and finally identifying drug-using associates and non-drug-using members of ego's personal social networks. The subject is encouraged to participate throughout the process, offering information, evaluating the chart construction, and providing changes and corrections when appropriate.

The construction of the chart is a method that elicits information in a more complete way than, for example, the construction of a list of kin or sexual partners who are uninformed by the subject of her HIV-positive serostatus.

Charting, as a method, renders persons and categories of persons reified as symbols on paper and encourages answers to questions regarding drug use behavior, sexual behavior, and HIV infection without personal identification. It is also a method that is of interest to subjects, thereby maximizing involvement in participation. During the period of data collection, study participants regularly requested copies of their charts.

In addition, the charting process has a therapeutic application. The process of chart construction is itself an intervention. Drug-using persons enrolled in a methadone maintenance program typically have siblings, cousins, and partners who use drugs. The chart organizes information about these people in a readily accessible form, making the numbers of drug-using family members and friends obvious to subjects. One woman, having identified four of six siblings as well as several cousins as drug users, remarked with astonishment about the number of drug users in her family, a fact that, until that moment, had not been apparent to her. Another subject identified her father as an alcoholic and three of her four sisters as heroin and crack users. On completion of the chart, noting the multigenerational character of substance abuse in her family, she expressed the fear that her two children would use drugs and follow the example of her father, herself, and her sisters.

Once all associated persons (such as kin, friends, drug associates, lovers, casual sexual partners, neighbors, work friends) have been identified and their sexes, ages, and places of residence noted, various attributes are assigned to these members of ego's personal social network, or "personal community," a term borrowed from Wellman and Berkowitz (1988, p. 28). Like identifying members of the personal community, seeking the attributes of the identified persons is done in an orderly fashion. Each major attribute (drug use, friend, household member, sexual partner, member of children's household, intimate friend, and knowledge of ego's serostatus) is represented graphically (see figure 1 for key). (The Women's Center kinchart-sociogram used colors rather than the various lines and symbols used in figure 1.)

The first application of this method was an investigation of the social networks of drug-using female methadone patients; the second was a study of HIV-positive drug-using women's reproductive decisions. The specific attributes of persons identified by subjects included the drug or drugs used, friendship, emotional intimacy, household membership, children's household membership, sexual partner, casual sexual partners, cause and year of death, serostatus, knowledge of ego's serostatus, and designations further defining nonkin as neighbors, childhood friends, and work-related associates. Subjects were asked to identify only those relatives known to them, with the exception of the grandparents and

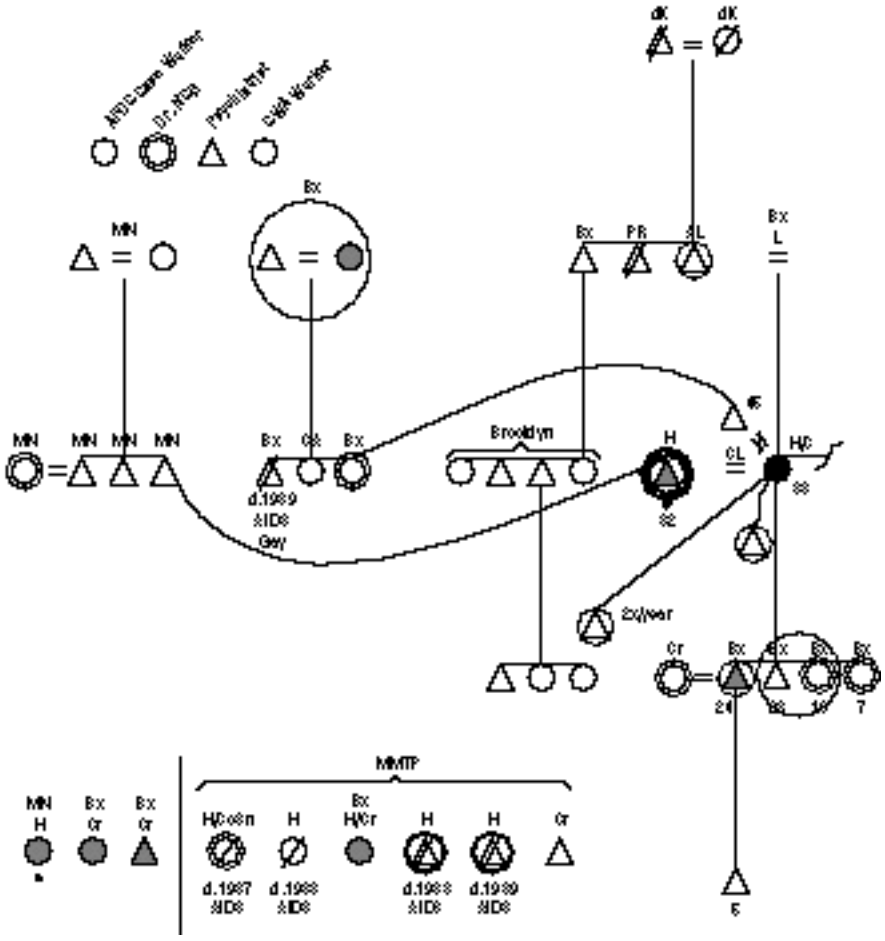
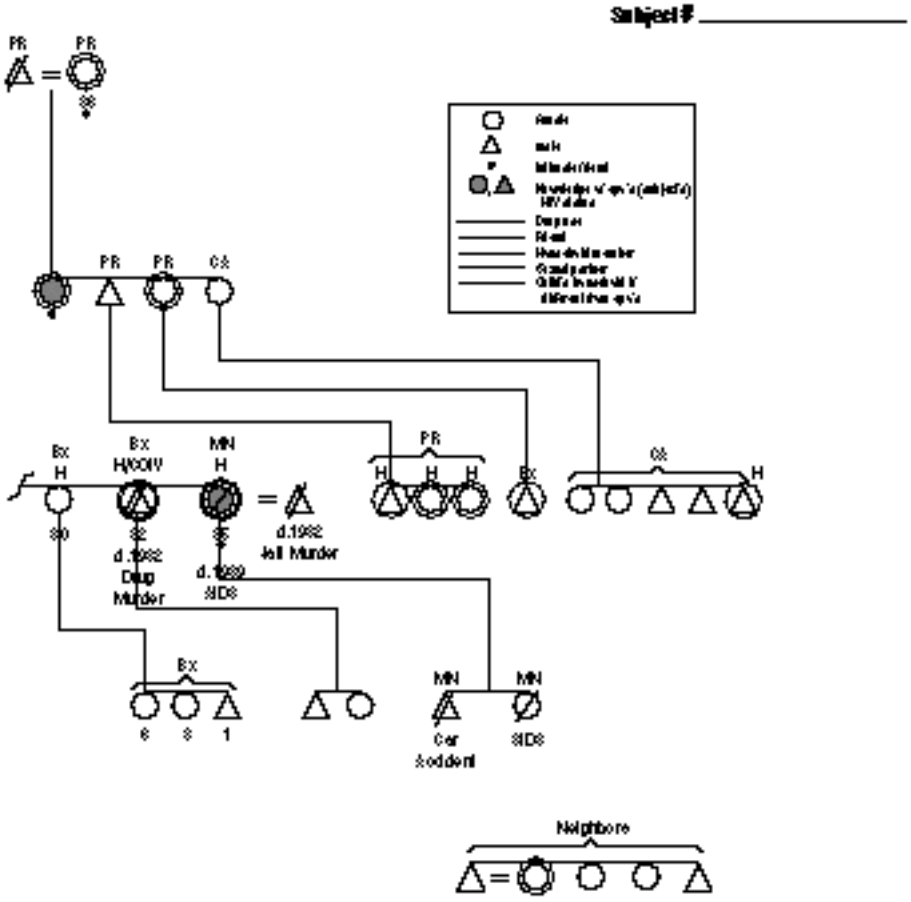


FIGURE 1. *Example of a kinchart-sociogram*

great-grandparents whose places of birth were elicited even if the individuals had not been known to the subject.

In both the study of social networks and the study of reproductive decisions of HIV-infected women, the partially completed chart was accompanied by an investigator-administered, open-ended, structured questionnaire (see appendix). Throughout the interview, the subject was asked to refer to the chart to add people or to amend information about already-identified persons. Sensitive questions about constituents' (individuals represented



KEY: AFDC=Aid to Families With Dependent Children; AIDS=acquired immunodeficiency syndrome; AL=alcohol; Bx=Bronx; CA=California; CL=common law marriage; Cr=crack; CWA=Child Welfare Administration; d.=died; dK=does not know; Dr. NCB=physician at North Central Bronx Hospital; H=heroin; H/C=heroin and cocaine; H/COIV=injected heroin and cocaine; H/COS=heroin and intranasally ingested cocaine; L=legal marriage; MMTP=Methadone Maintenance Treatment Program MN=Manhattan; PR=Puerto Rico; SIDS=sudden infant death syndrome

by symbols on figure 1) HIV infection or AIDS diagnosis were asked during the structured interview rather than during the initial chart construction because the probability of positive response increased with the trust developed during the interview and because some attributes were better asked about in contexts other than the chart construction.

At the conclusion of the interview, subjects were asked to review the kinchart-sociogram for accuracy.

EXAMPLES OF DATA ELICITED THROUGH THE KINCHART-SOCIOGRAM METHOD

Prevalence of Reported HIV, AIDS, and Violent Deaths

One product of the use of the kinchart-sociogram method is the data derived from the initial study of personal social networks describing the numbers of persons whom subjects identified as having died by violent means or from AIDS-related illnesses as well as those described as being HIV infected.

In the initial charting and interview process, 126 female methadone patients participated. During the interview, with the kinchart-sociogram as a reference, subjects were asked to identify persons they knew who had been diagnosed with or had died of AIDS and those who were HIV infected. In the course of constructing the kinchart-sociograms, subjects volunteered information about relatives, sexual partners, friends, and drug associates who had died by violent means, including accidental death by fire, drowning, electrocution, or motor vehicle accident; murder by shooting, stabbing, battering, or being pushed into or from a subway train; drug-related deaths (overdose); and suicide.

The presence of HIV infection in this population can be described in the context of these subjects' households. Of the 119 households, 54 (45.4 percent) were reported to have at least one member, possibly including the subject, who was HIV infected. Among these 54 households, 33.2 percent of the total membership (67 of 202 individuals) were HIV infected. A total of 43 households had 1 HIV-infected member; 9 households had 2 HIV-infected members; and 2 households had 3 HIV-infected members.

Of the 71 minor children who lived in 27 (50 percent) of the 54 households with HIV-positive members, 5 (7 percent) were reported to be diagnosed with AIDS.

Of the 715 total siblings (all subjects plus their siblings), 69 (9.7 percent) were reported to be HIV infected.

The 126 subjects reported 55 HIV infections, 63 AIDS deaths, and 44 deaths by violent means from among 4,928 known relatives and lovers. These HIV infections, AIDS deaths, and violent deaths were reported from among siblings, husbands, lovers, children, nieces, nephews, aunts, uncles, siblings-in-law, cousins, ex-husbands, ex-siblings-in-law, and stepparents

(individuals of all ages). These aggregated numbers of HIV infections, AIDS deaths, and AIDS diagnoses did not include identifications of deaths and infections among friends, neighbors, or drug-using associates. The 63 AIDS deaths represent 1.3 percent of the 4,928 constituents.

When these experiences with HIV infection are examined from the perspective of each woman, the impact of AIDS becomes more tangible: From among the 126 women, 47 reported having experienced AIDS deaths among family members—35 experienced 1 death, 10 experienced 2 deaths, 1 experienced 4 deaths, and 1 experienced 5 deaths.

Furthermore, when AIDS experience was stratified by generation, it was clearly demonstrated that, at this point in the epidemic, the most profound losses experienced by this low-income, urban community were suffered by members of the subjects' generation, persons between ages 20 and 45. From among the 126 women, 58 AIDS deaths were reported from among siblings, husbands, lovers, siblings-in-law, cousins, ex-husbands, and ex-siblings-in-law (those of the women's generation).

The comparison of AIDS deaths and violent deaths from among extended family members is another means derived from the chart data by which to understand the impact of the AIDS epidemic. Of the 126 women, 34 experienced the violent deaths of 44 family members; 12 women experienced the death of at least 1 sibling, and 3 women experienced the death of at least a husband. Of the 34 women who lost members of their personal communities through violent deaths, 14 also had been affected by 1 or more AIDS deaths.

The number of deaths by violent means (44), which constituted 0.9 percent of the total number of 4,928 constituents, was exceeded by the number of reported AIDS deaths (63); the AIDS deaths represented 1.3 percent of the total number of relatives. The number of persons reported to be HIV infected (55 or 1.1 percent) exceeded the number of persons reported to have died by violent means (44 or 0.9 percent).

Preliminarily, without adjusting for age or for the period of time the AIDS epidemic has been experienced by members of this community, if the numbers of reported HIV infections (55) and reported AIDS deaths (63) are considered together as a gross representation of the effect of the HIV epidemic, it is clear that AIDS already has had a far greater impact on this group of low-income, ghetto-dwelling women than even that of the high number of losses to date experienced by deaths from violent means (44).

Prevalence of Reported Drug Use Within Subjects' Social Networks

In the study of social networks, the kinchart-sociogram methodology elicited information about subjects' drug use and the drug use patterns of subjects' sexual partners, coresidents, and family members. Influential social aspects of drug use were described through categories of analysis that included the drug use patterns of subjects and their coresiding sexual partners, participants' children's and coresidents' drug use, and histories of drug use in subjects' families of origin.

The extent to which drug use and HIV and AIDS permeate the social relations of women in this methadone clinic in the Bronx is apparent from the data. Although all the subjects were enrolled in treatment for illegal drug use, 72 percent continued to use illegal drugs (mostly cocaine) at least monthly, and more than 35 percent did so daily. Of the 63 women who resided with a sexual partner, 36 reported regular use of illegal drugs. Although only a minority of the partners (fewer than 25 percent) used crack or other forms of cocaine, statistically significant positive associations were found between women's crack use and that of their partners as well as between women's use of other forms of cocaine and that of their partners. This pattern differed from that of heroin users, among whom no statistically significant association was found between women's current use of heroin and that of their coresiding sexual partners.

Drug use was common in the households with minor children. Of the 142 children who resided with their mothers, 92 (64.8 percent) lived with mothers who reported regular use of illegal drugs; 117 (82.4 percent) of the children lived with at least one parent who used drugs; and 123 (86.6 percent) lived in households where at least one adult household member was a drug user.

Within subjects' families of origin, almost one-third of the subjects' siblings were drug users. Among the 240 known parents, 17 (7.1 percent) were reported to have abused alcohol, and 19 (7.9 percent) were reported to have histories of using illegal drugs. One parent was reported to have abused alcohol and illegal drugs.

Only 13 (10.3 percent) of the 126 subjects reported that their mothers had histories of substance abuse. Seven (5.6 percent) of the mothers were identified as heroin users, four (3.2 percent) as alcoholics, two (1.6 percent) as cocaine users, and one as a marijuana user.

Of the 114 women who knew their fathers, 21 (18.4 percent) reported that their fathers had histories of substance abuse. Thirteen (11.4 percent)

fathers were identified as alcoholics, seven (6.1 percent) as heroin users, two (1.8 percent) as cocaine users, and one as a marijuana user. Two fathers were reported to have abused multiple substances.

Use of illegal drugs was widespread among subjects' siblings: 126 women in the study reported a total of 589 siblings (mean=4.7; range 0 to 20). Of the 119 women who had siblings, 84 (70.6 percent) reported that at least one of their siblings had a history of illegal drug use.

Of the 589 siblings, 180 (30.6 percent) were reported to have histories of abusing drugs other than alcohol. Of these 180 drug-using siblings, 153 (85 percent) were reported to have used heroin or some form of cocaine. Of the 126 women in the study, 8 (6.3 percent) reported having 1 sibling who abused only alcohol.

Considering the subjects and their siblings together, among these families, almost half the subjects' generation were drug users. The relatively low numbers of parental drug users (less than 8 percent) compared with the high numbers of drug-using siblings (50 percent) suggests a marked increase in drug use between generations.

The social character of drug use is further evident in the numbers of households in which one or more member is a drug user and the numbers of children living in these households. A child's residence with a drug-using adult may not necessarily lead to the child's drug use and does not necessarily imply deficient parenting, but the widespread presence of drug-using adults in households with minor children supports a perception of drug use and related behavior as normative (Bauman and Dougherty 1983; Deren 1986).

NIDA's Perinatal-20 Treatment Research Demonstration Program

The kinchart-sociogram methodology recently was amended for use in a NIDA-funded study of postpartum drug-using mothers enrolled in the Family Services Project (FSP), a program intended to improve the health outcomes of offspring born with positive urine toxicologies and their siblings. For this study, additional categories of constituency were included in the standard progression of questions designed to elicit the members of ego's personal social networks. The added categories included health care providers, mental health providers, public assistance workers, and treatment program personnel. The inclusion of additional categories was intended as a means of evaluating the family case management intervention. It was hypothesized that once maternal drug use has been addressed and treated, the membership of ego's personal

social networks would change to include fewer drug users in ego's household and among ego's friends. Non-drug-using persons (in this case, support group members, family therapists, and other program staff) and health care providers were expected to appear on the chart on readministration. The chart also identified the residences of ego's children born prior to the index child, whose birth brought the family into the research treatment, and was expected to document the reunification of egos and those children who previously resided apart from their mothers. The study protocol specified the readministration of the chart every 6 months for 2 years, beginning at enrollment, which was prior to participation in the program. These data will be forthcoming on the readministration of the charts at 6-month intervals.

In addition to eliciting data about the social networks of 8 subjects enrolled in the FSP, kinchart-sociograms have been administered to 2 comparison groups: 30 women participating in Narcotics Anonymous (NA), the well-known but little-studied self-help program for drug users, and 61 women enrolled in the MMTP where this methodology was first developed. Preliminary data from these three study groups describe drug use of subjects' families of origin and drug use of subjects' coresidents.

Substance Abuse Among Subjects' Parents

Among the FSP subjects, 50 percent of the fathers (four of eight persons) and 12.5 percent of the mothers (one of eight persons) were reported to use alcohol. Among NA subjects, 50 percent of the fathers (15 of 30) and 36.7 percent of the mothers (11 of 30) were reported to use alcohol. Among the MMTP subjects, 49.2 percent of the fathers (30 of 61) and 34.4 percent of the mothers (21 of 61) were reported to use alcohol.

None of the fathers and mothers of the FSP subjects were reported to use illegal drugs. Among NA subjects' fathers, 26.7 percent (eight) were reported to use one or more illegal drugs; 23.3 percent (seven) of the NA subjects' mothers were reported to use one or more illegal drugs. Among MMTP subjects, 9.8 percent (six) of the fathers and 14.8 percent (nine) of the mothers were reported to use one or more illegal drugs.

One possible explanation for the different trend in percentages of reported parental drug use between FSP subjects (who reported no use of illegal drugs by any parent) and NA and MMTP subjects may be issues of child custody. Many FSP mothers do not have custody of their children; many of these children are in the custody of the subjects' mothers. The FSP mothers may not volunteer histories of parental drug use because of fears of custodial consequences.

Reported Substance Abuse Among Subjects

Three (37.5 percent) of the eight FSP subjects reported using illegal drugs. Of the 61 MMTP subjects, 30 (49.2 percent) reported currently using 1 or more illegal drugs. None of the NA subjects reported using illegal drugs.

Reported Substance Abuse Among Subjects' Siblings

Among the siblings of seven FSP subjects, two (28.6 percent) were reported to have histories of alcohol abuse. Among the siblings of 30 NA subjects, 16 (53.3 percent) were reported to have histories of alcohol abuse. Among the siblings of 56 MMTP subjects, 28 (50 percent) were reported to have histories of alcohol abuse.

Of the seven FSP subjects with siblings, six (85.7 percent) reported at least one sibling with a history of using illegal drugs. Of the 30 NA subjects with siblings, 24 (80 percent) reported at least 1 sibling with a history of illegal drug use. Among the 56 MMTP subjects with siblings, 37 (66.1 percent) reported histories of drug use among these siblings.

The 99 subjects reported 181 sisters and 143 brothers. Sixty-nine (38.1 percent) of the 181 sisters were reported to have a history of illegal drug use, 47 (26 percent) were reported to have a history of alcohol abuse, and 91 (50.3 percent) were reported to have a history of substance abuse (alcohol and illegal drugs).

The 99 subjects reported that 69 (48.2 percent) of their 143 brothers had a history of illegal drug use, 43 (30.1 percent) had a history of alcohol abuse, and 81 (56.6 percent) had a history of substance abuse (alcohol and illegal drugs).

Of the total 423 subjects and siblings, 166 (39.2 percent) were reported to have histories of using illegal drugs.

Illegal Drug Use Among Subjects and Subjects' Household Members

Two (25 percent) of the eight FSP subjects reported at least one drug-using coresident. None of the 30 NA subjects reported a drug-using coresident. Of the 61 MMTP subjects, 22 (36.1 percent) reported residing with at least 1 drug-using coresident.

These preliminary data support the findings reported in the study of social networks among female methadone patients in which the social nature of drug use was documented (Pivnick et al. 1994). The association

of subjects' abstinence with households in which there were no drug users suggests the importance of residential environment and social relations in the treatment of drug abuse.

CONCLUSION

The kinchart-sociogram method elicits detailed information about the personal communities of subjects, including information vital to assessing social support for HIV illness, the frequency of drug use among family members and friends, and sexual contact patterns relevant to HIV transmission risk. Personal communities are also the framework through which people share knowledge about AIDS risk and prevention, especially in AIDS-endemic areas such as the Bronx. Moreover, these communities represent natural systems of personal support during illness, for child care and custody, and for housing.

Kinchart-sociograms represent networks of a subject's relationships and interactions with other persons. The object is to define these networks in terms of concrete social relations located in specific sites. In this analysis, relations are described between ego and individual persons among ego's kin group, network of friends, sexual partners, drug associates, and household members and are characterized by ego's needs, behaviors, and exchanges of tangible goods and support. Ego is central. The chart is a map of ego's interpersonal relations, represented by self-reported dyads.

However, the identification of an individual's personal social relationships does not necessarily imply interactions between constituent members as in traditional social network analysis. Interaction with ego is the only requirement for inclusion. Nevertheless, it is not assumed that these dyads, comprising ego and another person, are the only level of analysis possible; in reality, relationships exist among members of ego's personal social networks such as kin, persons related through drug use (e.g., persons enrolled in the same drug treatment program and needle-sharing partners), persons who are HIV positive who see the same doctor, and persons who live close to one another. A second level of analysis is therefore implied, that of the members of ego's personal community who interact with each other and whose interactions affect ego's behavior and other relations. Although implementation of the kinchart-sociogram method does not address this second level of social network analysis, the opportunity to recast the methodology to include relations between network members is readily apparent.

The kinchart-sociogram method permits a graphic and quantifiable description of complex social realities. As an intervention, it is of marked

interest to subjects and, by requiring the subject's active participation in the process of inquiry, encourages the subject's development of awareness about the nature of AIDS risk and the extent of drug use in sexual and residential relations. Furthermore, by repeating the charting at regular intervals, the method permits prospective study of changes in social networks associated with changes in drug use patterns and, in the case of HIV infection, disease progression.

An account of the numbers of HIV-infected adult household members generated by the methodology provides a perspective on the numbers of children at risk of losing one or both parents and potentially becoming dependent on an already overtaxed foster care system. The identification of drug use and specific drugs used among household members is a useful means of describing drug use patterns and residence patterns among drug users.

Finally, the kinchart-sociogram method elicits data that demonstrate the social nature of drug use and its expression in influential social contexts such as families, sexual relations, and households. In turn, the description of these social contexts suggests the development of treatment modalities that include sexual partners, children, and household members.

REFERENCES

Bauman, P.S., and Dougherty, F. Drug-addicted mothers' parenting and their children's development. *Int J Addict* 18(3):291-302, 1983.

Chu, S.Y.; Buehler, J.W.; and Berkelman, R.L. Impact of the human immunodeficiency virus epidemic on mortality in women of reproductive age, United States. *JAMA* 264:225-229, 1990.

Deren, S. Children of substance abusers: A review of the literature. *J Subst Abuse Treat* 3:77-94, 1986.

Evans-Pritchard, E.E. *The Nuer*. Oxford: Clarendon Press, 1940.

Fortes, M. *The Dynamics of Clanship Among the Tallensi*. London: Oxford University Press, 1945.

Keesing, R.M. *Kin Groups and Social Structure.* New York: Holt, Reinhart & Winston, 1975. 177 pp.

Landesman, S. Human immunodeficiency virus infection in women: An overview. *Semin Perinatol* 13:2-6, 1989.

Levi-Strauss, C. *Les Structures Elementaires de la Parente*. Paris: Plon (translation by Bell, J.; Von Stumer, J.; and Needham, R., as *Elementary Structures of Kinship*); Boston: Beacon Press, 1969.

Morgan, M.L. *The League of the Ho-De-No-Sau-Nee or Iroquois.* Reprint. New York: Dodd, Mead, 1904.

Netting, R.M.; Wilk, R.R.; and Arnould, E.J., eds. *Households, Comparative and Historical Studies of the Domestic Groups.* Berkeley, CA: University of California Press, 1984.
New York City Department of Health. *City Health Information No. 9*. 1990. 4 pp.
Penn, M. Coalitions and binding interactions in families with chronic illness. *Fam Systems Med* 1:16-25, 1983.
Pivnick, A. "Kincharts-Sociograms: An Anthropological Method for Describing the Personal Social Networks of Drug-Using Persons." Unpublished document, 1994. 17 pp.
Pivnick, A.; Jacobson, A.; Eric, K.; Doll, L.; and Drucker, E. The prevalence of AIDS, HIV infection, and illicit drug use in inner city families and social networks. *Am J Public Health* 84:271-274, 1994.
Pivnick, A.; Jacobson, A.; Eric, K.; Mulvihill, M.; Hsu, M.A.; and Drucker, E. Reproductive decisions among HIV-infected drug-using women: The importance of mother-child coresidence. *Med Anthropol Q* 2:153-169, 1991.
Schneider, D.M. *A Critique of the Study of Kinship*. Ann Arbor, MI: University of Michigan Press, 1987.
Wellman, B., and Berkowitz, S.D. *Social Structures: A Network Approach.* Cambridge, MA: Cambridge University Press, 1988. 513 pp.

ACKNOWLEDGMENTS

The development of the kinchart-sociogram methodology was accomplished through the support of the Centers for Disease Control (H64/CCH20071408) and further explored through NIDA grant DA-06925-04. Audrey Jacobson provided editorial comments and analyzed data.

AUTHOR

Anitra Pivnick, Ph.D.
Director
The Women's Center
Department of Epidemiology and Social Medicine
Montefiore Medical Center
3320 Rochambeau Avenue
Bronx, NY 10467
(718) 920-4280 (Tel)
(718) 920-2745 (Fax)
apivnick@.aecom.yu.edu (Internet)

APPENDIX. Questionnaire used with the kinchart-sociogram

Note: If death of any person is reported, record year and cause of death.

Protocol for the Construction of the Kinchart-Sociogram

How old are you?

How many brothers and sisters do you have?

What are their ages?

Where do they live?

Are your parents living?

Where do they live?

Do you know your mother's brothers and sisters?

Where do they live?

Do you know your father's brothers and sisters?

Where do they live?

Do you know your mother's mother and father?

Where do they live?

Do you know your father's mother and father?

Where do they live?

Do you or did you know your mother's and father's grandparents?

Where do or did they live?

Do you know the children of your mother's and father's sisters and brothers (your cousins)?

Where do these people live?

Do you have children?

How many children do you have?

What are their sexes and ages?

Where and with whom do they live?

Do you know your brothers' and sisters' children?

What are their ages and sexes?

Are you married?

Is your marriage common law or legal?

How old is your partner?

Do you know the brothers and sisters of your current partner?

Where do they live?

Do you know your current partner's parents?

Where do they live?

Have you been married before (number of husbands)?

Were these marriages common law or legal?

Do you know the sisters and brothers of your former husbands?

Where do they live?

Do you know the parents of your former husbands?

Where do they live?

Is your current husband the father of your children?

Who are the fathers of your children?

Do you know the brothers and sisters of your children's father or fathers?

Where do they live?

Do you know the parents of your children's father(s)?

Where do they live?

Are there any other relatives you know that I haven't asked about?

Have you had a lover or lovers during the past five years other than your spouse?

Is this lover(s) still your lover?

Do you know your lover's family?

Do you have friends who don't use drugs?
(Prompt if necessary: neighbors, old school friends, old family friends, work friends.)

Where do these friends live?

Do you have any friends (and/or associates) who do use drugs?

Do you have any friends (and/or associates) who are in the Methadone Maintenance Treatment Program?
(Construct methadone patients separately on the chart.)

Which drugs do they use?

Where do these friends (and/or associates) live?

Are any of these people your neighbors?

Do any other people on the chart use drugs?
(Prompt: siblings, parents, in-laws, etc.)

Which drugs do they use?

Are any of the following, people you care about: social workers, welfare case workers, Child Welfare Administration case workers, doctors, nurses, midwives, social workers, therapists, or other health care people?

Where do you see these people?

Are any priests, ministers, rabbis, or spiritual advisers important to you?

Where do you see these people?

Which people on the chart do you consider your friends (including all categories of persons)?

Whom do you consider your closest friends?

Whom do you live with?

(Questions asked during the administration of questionnaire)

Do you have casual sexual partners?

How frequent are these sexual encounters?

Do you know your HIV status?

Which people on the chart know you are HIV positive?

Do you know anyone who is HIV infected?

Do you know anyone who has died of AIDS?

When did they die?

Factors To Consider When Using Hair as a Cocaine-Exposure Measure for Mothers or Newborns

Paul R. Marques

INTRODUCTION

Historical information about drug exposure is available from an analysis of hair. But is hair testing good enough to reliably determine *degree of exposure* and help judge the specific contribution of cocaine to infant risk? Furthermore, is hair testing a worthwhile outcome measure for tracking maternal progress in remediation programs?

Pregnant women who are identified as occasional users of illegal drugs such as cocaine often are lumped together with seriously dependent women into an undifferentiated category of "maternal drug abuser." This occurs even though the teratologic consequences of cocaine use appear to follow the same general pattern found for legal drugs such as alcohol and tobacco. That is, the nature and degree of infant risk appear to bear some relationship to dose level, time of fetal exposure, and general maternal health and nutrition (Kopp and Kaler 1989). Risk factors are interactive; mothers often use multiple drugs; and the relative importance of different risk factors is often difficult to specify in human studies, particularly without *reliable* exposure information (Zuckerman 1991, pp. 352-362; Suomi and Higley 1991, pp. 291-302; Sokol et al. 1986). Despite such uncertainties, some service providers adopt the simplistic assumption that any degree of gestational cocaine exposure will produce defective children and accordingly find justification for heroic early intervention, a practice earlier decried by Coles (1992, pp. 248-258), Coles and Platzman (1993), and others. A similar sense of urgency rarely follows when a woman reports occasional use of alcohol or tobacco while pregnant.

A child exposed to large doses of drugs during early gestation and later raised in an environment of significant stress is thought to be at greatest risk for postnatal problems (Kosofsky 1991, pp. 128-143; Kopp and Kaler 1989). Realistic estimates of environmental stresses can sometimes be made by providers and outreach workers, but knowledge of the degree of gestational exposure is more problematic. Because maternal hair testing allows the possibility of a quantitative and historical estimate of degree of fetal cocaine exposure, it could serve as an objective screen for gauging the degree of alarm warranted on delivery. Estimates based on the ponderal index and infant head circumference suggest that intrauterine

growth retardation (IUGR) occurs in association with cocaine early in gestation (Frank et al. 1990), at a time that antedates postnatal detection through either urine or meconium sampling.

Analysis of a pregnant woman's hair, segmented by trimester, could aid in honing the presumption of drug-related risk and serve as a useful correlate of postnatal outcomes such as IUGR; however, infant cocaine hair levels probably would not be as good a diagnostic of teratologic effects as maternal hair. Callahan and colleagues (1992) segmented maternal hair samples that were presumed to represent 9 months into three trimester segments. They found the maternal cocaine hair levels of only the last trimester to significantly correlate with the infant hair cocaine levels, which confirms the conventional wisdom that cocaine found in *infant hair* reflects only the last months of gestation. However, that mothers' and infants' hair cocaine levels cross-correlate helps support the argument that the technique is reliable, a finding supported by Marques and colleagues (1993). Therefore, maternal hair, not infant hair, becomes the key specimen for assessing the degree of fetal exposure during the first trimester when drugs can retard growth or impair normal organogenesis.

In a separate domain, adult drug treatment programs often have to choose between the costs and inconveniences of frequent urine screening or the legendary inaccuracy of self-report as a means of documenting change in drug use status during long-term followup. Here, too, hair analysis holds promise for helping to objectify these estimates. However, few studies report cocaine in the hair as a repeated measures outcome variable to document program effectiveness. If the hair of the mother has a reliable record of exposure and if the technique is valid, then hair testing can prove to be more than just a surveillance tool: It can be a diagnostic tool and an outcome variable of considerable value to clinicians and researchers.

Nevertheless, before any health-related uses of hair testing can be fully realized, more rigorously analyzed information is needed to help document the factors that influence hair testing results.

ISSUES OF VALIDITY AND RELIABILITY

Some drug researchers are troubled by the technology of hair analysis for drugs of abuse because there are unanswered questions about variables that affect drug sequestration in hair, the comparability or validity of results based on different drug extraction techniques, and the degree of reliability and quantifiability of the procedure (Keegan 1991-92; Taylor 1990; Holden 1990). Such expressions of caution, articulated in the scientific literature, are partly scientific and partly derivative of ethical concerns over the many

inappropriate ways this technology could be used to exclude or punish people who disobey the law. It first should be determined that there will be a health-related benefit from hair analysis. Such concern is especially valid in view of racial differences in hair composition, differences that could well be reflected in hair's affinity for drug uptake. For example, the ratio of fibrous protein to matrix substance in hair across three major racial groups can vary by a factor of 2.5, where Asian > Caucasian > African (Dekio and Jidoi 1990). Thus, hair is not a homogeneous substrate, and until more is known about the variables that affect drug uptake and levels found in hair, it is premature to assume comparable exposure based on comparable levels in hair. This latter problem is not at issue when studying repeated measures within subjects and evaluating change over time.

Uptake of cocaine and its metabolites into hair is *not* a strict function of blood concentration. Several different research teams, including Henderson and colleagues (1992) and Möller and colleagues (1992), using the hair of South American coca chewers, and Nakahara and coworkers (1992), using rat hair, have reported that the ratio of cocaine metabolites to cocaine is different in hair from that in blood. The parent compound, untransformed cocaine, is found in much higher concentrations (4 to 9 times higher) in hair than are its main metabolites, benzoylecgonine (BZE) and ecgonine methyl ester, despite the higher blood concentrations and longer circulating half-lives of the metabolites. Therefore, although levels of cocaine and its products found in hair may reflect the amount consumed by an individual, there are factors influencing uptake or measurement that reverse the blood concentration ratios of these compounds. Reasons for these differences are not clear, but such findings render confident quantitative comparisons across individuals still slightly beyond reach at this time.

Even so, it should be acknowledged that such uncertainties are inevitable in a young technology, and they warrant caution in, but not dismissal of, the technique. Because of the uncertainties, health researchers who use hair testing results should avoid blurring the boundary between the health implications and the legal implications of cocaine use. It is neither wise nor unusual for a pregnant woman to consume alcohol or tobacco; the same holds for cocaine. But just as there is no defensible health motive to recommend drug treatment intervention for the occasional pregnant alcohol user, there is no health basis for advocating treatment of the occasional cocaine user. If the aggregated research record can show hair to have *quantitatively* valid and reliable exposure information—gleaned either from the mother's or the newborn's hair—then it may help to target therapeutic interventions prenatally or postnatally on the mothers and infants at highest risk. By contrast, law enforcement, which is enjoying somewhat of a romance with hair testing, is less concerned with exposure

levels than it is with evidence of illicit activity, an inherently more dichotomous problem and one that is less dependent on the quantitation that is critical to the health sciences.

Broadly speaking, science appears to be midway through a two-part examination of the validity and reliability of hair testing. Two criteria that should be satisfied are (1) dichotomous or qualitative accuracy of hair testing—Does it detect drugs of abuse in hair?—and (2) quantitative accuracy of hair testing—Does the level of a drug found in the hair of an individual reflect the amount consumed over time? There is general agreement that the question about qualitative accuracy has been answered in the affirmative through research conducted over the past decade. Hair does provide a record of drug exposure of adults and newborns (Cone 1990; Cone et al. 1991; Graham et al. 1989; Welch et al. 1993; Forman et al. 1992; Marques et al. 1993; Callahan et al. 1992; Baumgartner et al. 1989). Moreover, research groups that have compared hair, meconium, urine, and self-report have found hair to be the most sensitive (e.g., maximum true positives and minimum false negatives) at detecting the presence of cocaine in newborn infants (Ostrea 1992, pp. 61-79; Callahan et al. 1992). However, there is still some concern that despite the sensitivity of detection via hair, the specificity (minimum false positives and maximum true negatives) of maternal hair (not necessarily infant hair) may be low in some circumstances. This concern grows from the debate among researchers about whether drugs carried by smoke, and which then adhere to hair, can be fully washed off the surface during sample preparation. Cone and colleagues (1991) report that intentional high-dosage environmental exposure with cocaine produces false positives that cannot be fully washed out in the laboratory. The implications of this are important for both legal and health actions, because if there are to be sanctions imposed on someone (or an exposure risk assumed) when drugs are found, it is important to know that drugs found in hair have arrived there from the circulation of blood internally and not from the circulation of air externally. This is obviously not a problem when sampling hair from a newborn infant, but it becomes progressively more of a concern as the infant grows older and is exposed to crack cocaine smoke in the air.

False Negatives

When a pregnant or postpartum woman's hair is found to contain cocaine, it is useful to know how much meaning this has for the degree of exposure of her infant or for her own treatment needs. Although infant hair analysis for cocaine should have high specificity, infant hair has no direct information about first trimester exposure. Therefore, the best estimate of early gestational exposure of infants depends on the valid and reliable

analysis of maternal hair. Without full confidence in the specificity of maternal hair analysis, it cannot be assumed that levels in a mother's hair reflect consumption. Therefore, it is worthwhile identifying circumstances under which levels found in maternal hair *can* be believed. Currently, the question of fetal exposure turns more on maternal self-report and whether drugs have been found at delivery than on any historical objective basis.

Because the circulatory half-lives of cocaine and its metabolites are short and because mothers at highest risk often are not available for testing earlier than at delivery (because they receive little prenatal care), the commonly employed estimate of infant exposure based on urine samples is only an indirect approximation of infant risk and reflects only the mother's use within days of delivery. Such reliance on urine is known to have a high false-negative rate. Ostrea (1992, pp. 61-79) cites estimates of the false-negative rates that range from 32 to 63 percent of cases when relying only on urine collected at delivery. These estimates match well with the author's false-negative results based on a known sample of 136 cocaine-using women. Using maternal hair as an index criterion for a woman who is cocaine-positive, the rates of false negatives are 72 percent for urine cocaine and 27 percent for urine BZE (when both hair and new urine samples were taken on the same day an average of 2 months after delivery). Fully 98 percent of all mothers in the project sample were cocaine positive in the hair. Among all infants of hair-cocaine-positive mothers who were measured, 95 percent were hair cocaine positive.

Pair Correlations

One difficulty of doing good validational studies of hair analysis in humans is dosage control. There have been few (if any) laboratory studies of controlled, long-term, high-dosage drug use that compared hair analysis with known consumption. Such studies are difficult because dosing subjects as frequently and as lavishly as addicts dose themselves poses both ethical and liability risks that most research centers cannot take. Conversely, the purity and amount of a drug voluntarily self-administered in the field are difficult to determine from interviews; therefore, comparing hair levels with real consumption over time is loaded with uncontrollable error. Long-term frequent urine sampling over many months of self-administered use by street addicts would likely be the most accurate way to establish quantitative accuracy of hair testing in humans.

One alternative method for *estimating* reliability is to determine the degree of concordance between mother-infant pairs. This is a reasonable estimate because mother and fetus are exposed to shared blood products. Earlier data from the author's studies (Marques et al. 1993) that compared mother

and infant based on the first 62 pairs showed that a large proportion (more than 50 percent) of all maternal hair was reported by laboratory technicians to be in poor condition; such damaged hair weakens the strength of the overall correlation between mother and infant. Correlations that select on the basis of only good condition maternal hair had explained 40 percent of the variability found in infant hair cocaine (correlation coefficient [*r*]=0.63, number of cases [n]=38, probability [*p*]<0.0005) (Marques et al. 1993). Cases added since that earlier report have raised the total bivariate sample correlation (unselected for hair condition) to *r*=0.52 for 111 pairs (94 percent African-American), representing a strong degree of concordance between mothers and infants. The scatter plots with regression line and 95-percent confidence intervals are shown in figure 1. The values are normalized through square root transformations because the distribution of raw values was highly skewed. A large proportion of these hair samples were reported to be in poor condition by the laboratory, presumably because of the use of hair treatments. Washing and preparing damaged maternal hair may result in finding less cocaine in the sample (relative to infant), which causes the overall regression line for all paired cases to intersect the Y axis at a positive point. If separate confidence

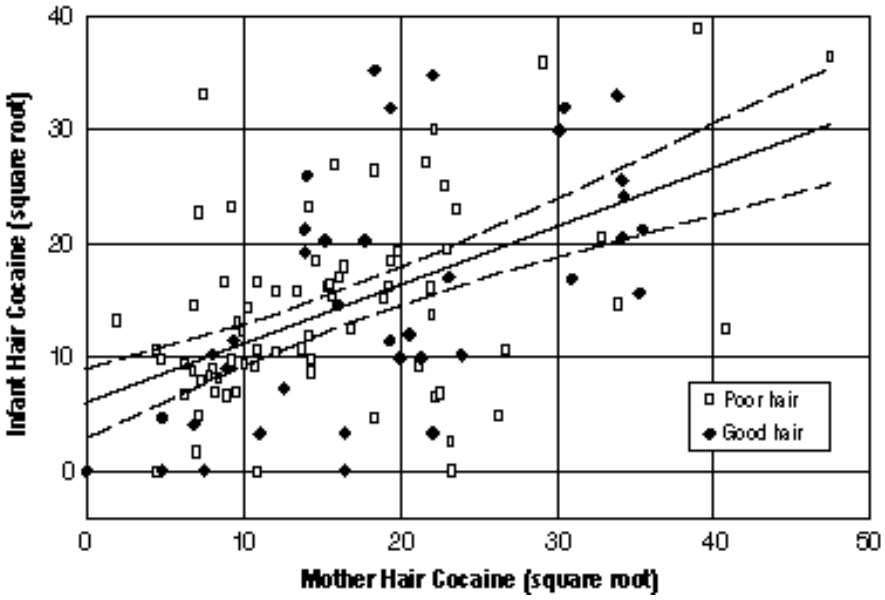


FIGURE 1. Infant/mother pairs (n=111) cocaine in hair

NOTE: r (correlation coefficient)=0.52; t (t statistic)=6.33; p (probability)<0.001. Solid line indicates regression; broken lines indicates ±95-percent confidence intervals.

SOURCE: National Public Services Research Institute

intervals and regression line for only the good maternal hair (black circles in figure 1) were plotted (not shown here), then the 0,0 point origin of the plot would be captured within the confidence intervals.

Callahan and colleagues (1992), with a sample of 82 percent Caucasian women, report an even stronger correlation between mothers and infants (r=0.72, n=52). A big difference in the Callahan data set and the author's is the proportion of (0,0) pairs (40 percent of all Callahan's cases were [0,0] pairs). The inclusion of unexposed cases is completely valid and appropriate, but it tends to constrain the regression line through the origin and raise the correlation. The data shown in figure 1 were more uniformly exposed, containing only one (0,0) pair. Artificially adding 40 percent null (0,0) pairs to the author's data (artificially raising n to 155) raises the full-scale correlation from 0.52 to 0.78, a correlation more in line with Callahan's; a similar adjustment to the cases selected for good condition maternal hair raises the correlation to 0.85. These two studies are among the first to report full-scale correlations between mothers and infants. Taken together, they lend some measure of credence to the quantitative accuracy of hair testing as a reflection of circulating blood levels of cocaine.

Correlation With Other Measures

Another reason for confidence in the quantitative accuracy of hair cocaine analysis is its correlation with other analytic methods. In the author's studies, the correlation between BZE levels in maternal urine (point-in-time sampling) and maternal hair was found to be significant for all cases (r=0.32, n=136, p<0.0005) and more strongly correlated when selecting only for hair in good condition (r=0.45, n=45, p<0.001); maternal urine BZE also was correlated with infant hair (r=0.35, n=108, p<0.0005). Even the short circulatory half-life of the parent compound cocaine was significantly correlated with hair but less strongly so (r=0.23 for both maternal and infant hair; p<0.003, n=136 and p<0.008, n=108, respectively). Although the window of time reflected by each measure is different (urine—short term, at best a few days; hair—long term, earliest detection at 1 week), just finding the relationship (and the correlation between mothers and infants) strengthens the basis for believing that hair can be quantitatively reliable. This is so because, over a large sample of users, those who use the most cocaine in the long term are also likely to show the greatest use in the short term. This principle of human nature, not overlapping windows of detectability, is the likely basis for the correlation. The presence of a correlation is more telling than the absence of one would be, especially because these are different matrices measured by different laboratories using different analytic techniques (gas chromatography/mass spectrometry and radioimmunoassay [RIA]).

Therefore, if two analytic procedures are well correlated, it would be helpful if the self-reported cocaine use also were correlated. However, the correlations found here between self-report and urine or hair endorse the position taken by McLellan and colleagues (1992) on the Addiction Severity Index (ASI), who limit questions about historical use on the ASI to the past 30 days because they have found longer term retrospective reporting of drug use to be unreliable. Of three self-report measures studied in this study's data set, only self-report in the past 30 days was correlated with analytic measures of use/exposure; both maternal urine cocaine metabolite (r=0.27, n=131, p<0.001) and maternal hair cocaine (r=0.17, n=136, p=0.022) showed weak positive correlations with 30-day self-report. Table 1 shows the cross-correlation matrix of cocaine measures relative to mother's use.

USING HAIR TESTING

If all the cautions about the analysis of drugs in hair can be accommodated or controlled to the satisfaction of the researcher, then hair testing has many advantages. From a pragmatic angle, the primary advantages are the savings of project time spent in sample handling and the potential reduction of laboratory costs (one sample every 3 or 4 months instead of twice-a-week samples). From a scientific angle, hair testing may permit the establishment of more specific relationships between exposure and outcome (mother treatment or infant health). Hair testing also brings a much less demeaning collection procedure for the client. On the down side, it is a new violation of privacy that warrants serious concern. Some women are resolutely opposed to hair sampling; some women maintain short hair; and an occasional client may be suspicious (e.g., the author's Caribbean subject who suspected that samples of her hair would be used in a clandestine Santeria ritual). Overall, about 8 percent of women have refused to provide hair for the author's study; approximately 12 percent have refused to allow their infants to be sampled.

Mothers

The greatest value of maternal hair testing may be that it allows the researcher or program evaluator to collect an integrated pool of drug-use outcome measures that represent an extended posttreatment period. Several types of drugs can be measured in a sample via a screening RIA, some (e.g., cocaine) far more accurately than others (e.g., marijuana). An example for cocaine, based on cases with repeated measures at three time points, is shown in figure 2. The ordinate is scaled in raw (untransformed) hair cocaine values; the left group shows all cases, and the right group is selected for the cases with hair judged to be in better condition. The figure

TABLE 1. *Correlation coefficients among cocaine-exposure estimates*

	Infant Hair Cocaine			Mother Urine Cocaine			Mother Urine Metabolite			Self-Report 30-Day			Self-Report History		
Drug-Exposure Measures	*r*	n	*p*	*r*	n	*p*	*r*	n	*p*	*r*	n	*p*	*r*	n	*p*
Mother hair cocaine	.52	111	.000	.24	136	.003	.32	136	.000	.17	136	.022	–.14	125	.065
Infant hair cocaine				.23	108	.008	.35	108	.000	.14	110	.069	–.14	102	.084
Mother urine cocaine							.59	136	.000	.21	131	.001	.04	120	.370
Mother urine metabolite										.27	131	.001	–.13	120	.078
Self-report 30-day													–.03	125	.359
Self-report history															

KEY: *r*=correlation coefficient; n=number of cases; *p*=probability (of chance occurrence)

SOURCE: Adapted and reprinted from Marques, P.R.; Tippetts, A.S.; and Branch, D.G., *American Journal of Drug and Alcohol Abuse*, 19(2), 1993, pp. 159-175, by courtesy of Marcel Dekker, Inc.

makes clear the possible *group* outcome monitoring that is possible with repeated hair samples.

Having access to a quantifiable index of exposure allows the researcher to ask at least two different categories of questions about the mother: (1) What are the baseline characteristics of higher vs. lower users (knowing that hair type is a possible but unproven confounder)? and (2) Which maternal characteristics are associated with reduced use from baseline over the duration of the project period?

Baseline Characteristics

None of the correlations between baseline hair cocaine levels and mother characteristics were found to be particularly strong, despite a broad variety of measures. However, among the strongest, most elements of the Caldwell HOME (Home Observation for Measurement of the Environment) scale (Caldwell and Bradley 1978) were negatively correlated with cocaine, especially the maternal involvement scale (r=-0.29, n=96, p<0.004); when 36 cases of "good hair" were selected, the correlation between maternal involvement and hair cocaine was -0.40. The Nursing Child Assessment Satellite Training (NCAST) feeding scale (Barnard 1980) also was strongly negatively correlated with good condition hair for cocaine (r=-0.50, n=42, p<0.0005), but using all 107 cases, the correlation was less interesting at r=-0.17, albeit still significant. In addition, the educational level of 142 cases was positively and significantly (but weakly) correlated to the

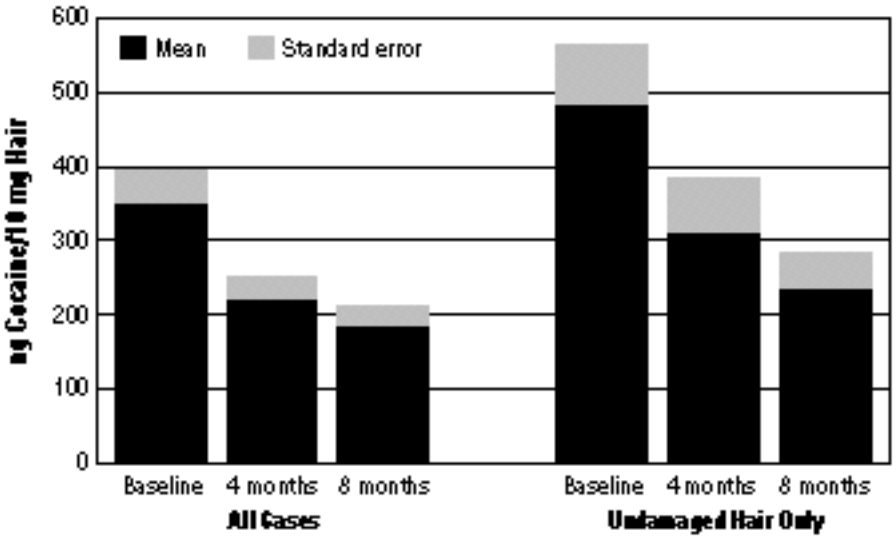


FIGURE 2. Measurement of hair cocaine in mothers (change from baseline to 4 and 8 months)

level of cocaine: The higher the educational level (within this sample of women having an average of 11 years of education), the more cocaine was found at baseline.

Subject Characteristics Associated With Decreased Use During the Intervention

Women who showed a trend toward reduced use during the project period had entry characteristics different from those who showed little decrease in cocaine use as measured by hair testing. Those who stayed in the sample (65 percent of all clients were still accessible after the first year and 50 percent after the second year) appeared to temper their drug use over time. This can be seen in figure 2, which shows cocaine levels in the hair of the women across 8 months. The decrease is significant as a linear effect and is even more accurately modeled by a second-order polynomial regression that accounts for the slowing of the magnitude of change.

The variables correlated with reduction in use over an 8-month period differ depending on how change is calculated (as a log-change-ratio score, a percent-change score, or an absolute difference; see Tippetts and Marques [this volume] for further information). There are no good rules for deciding how to calculate change. However, across the different potential change variables, the most persistent correlates of reduction in hair cocaine levels over time, for this sample, were reduction in Beck Depression Inventory scores (Beck 1978) and increased Rosenberg self-esteem scale scores (Rosenberg 1965). Those highest on the Millon passive-aggressive scale (Millon 1987) and antisocial personality scales (Millon 1987) were more apt to show an increase in hair cocaine levels across the 8-month period ($p<0.001$). These findings are preliminary and are cited only as examples of the use to which hair measures can be put in treatment outcome research.

In summary, the advantages of hair testing to researchers and clinicians come both from the convenience and the potential for greater accuracy. The risks are both scientific and societal and can derive from premature advocacy of an incompletely validated technology. These factors should be weighed by anyone planning to use hair testing in clinical research.

COMMENTS AND QUESTIONS

Many critics of hair analysis seem to believe that its value should be judged relative to the proven accuracy of urine testing, and for forensic purposes that is true. However, health outcome research has different needs. Hair testing may answer questions about long-term exposure where serial urine sampling is not a practical alternative because of its cost and

inconvenience. The arena in which hair testing should be evaluated is relative to other historical exposure estimates, such as meconium analysis and self-report. Behavioral researchers long ago learned to live with measurement error and for that reason require much larger sample sizes before confidence in a finding is warranted. In pharmacological studies a higher standard of accuracy is expected, and sample sizes are typically much smaller. Nevertheless, all researchers require better answers to some of the large questions of hair analysis that still need clarification. One is the effect of passive crack-smoke exposure on cocaine levels found in hair, and there are other important questions as well.

Hair testing advocates presume that hair grows at a rate of 1.3 cm per month (Baumgartner et al. 1989). The variables influencing hair growth rate are not clear, and it strains credibility to assume that growth is uniform across sexes, races, ages, hair types and compositions, and nutritional states. The effects of these factors should be easy questions to answer, but they do not appear to have been reported yet. Also, given those subject variables, what effect do such variables have on drug uptake and sequestration, and why is there a discrepancy between the ratio of metabolites to cocaine in hair relative to blood? Furthermore, is someone who takes more showers or sweats or swims more apt to lose drug from the hair? It is always good practice to take hair samples from the approximate same site in repeated measurements, but to what extent are there differences in regions of the head in measured drug concentration? Also, if there are differences, what should be made of them? Would it be better to sample and select hair from different regions of the head to get a truer estimate of exposure? Is the issue of hair damage important? How important is it? How old must an infant be before the correlation between mother and child falls apart?

If it can be assumed that up to half the statistical variability in the hair cocaine levels found among a group of mothers is reflected in the hair cocaine found in their infants, what accounts for the other half? How much of the unaccounted variance is analytic error as opposed to experimental error? That is, what contribution to error is made by the use of RIA as opposed to more precise techniques? Is linear regression necessarily the best way to model the relationship between exposure levels in mothers and infants? Also, given the large amount of error, what responsibility does the research community have to prevent unintended advocacy or promotion of this potentially intrusive technology? Hair analysis seems useful at the group-outcome level for researchers, but the careless use of the procedure could compromise civil liberties if hair drug levels are mistakenly believed to accurately reflect a particular individual's use.

REFERENCES

Barnard, K.E. *NCAST II Teaching Manual.* Seattle, WA: University of Washington, 1980. 65 pp.

Baumgartner, W.A.; Hill, V.A.; and Blahd, W.H. Hair analysis for drugs of abuse. *J Forensic Sci* 34(6):1433-1453, 1989.

Beck, A.T. *Beck Depression Inventory.* New York: Harcourt Brace Jovanovich, 1978. 25 pp.

Caldwell, B.M., and Bradley, R.H. *Manual for the Home Observation for Measurement of the Environment.* Little Rock, AR: University of Arkansas, 1978. 57 pp.

Callahan, C.M.; Grant, T.M.; Phipps, P.; Clark, G.; Novack, A.H.; Streissguth, A.P.; and Raisys, V.A. Measurement of gestational cocaine exposure: Sensitivity of infants' hair, meconium, and urine. *J Pediatr* 120(5):763-768, 1992.

Coles, C., and Platzman, K. Behavioral development in children prenatally exposed to drugs and alcohol. *Int J Addict* 28(13):1393-1433, 1993.

Coles, C.D. Discussion: Measurement issues in the study of effects of substance abuse in pregnancy. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children*. National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

Cone, E.J. Testing human hair for drugs of abuse. I. Individual dose and time profiles of morphine and codeine in plasma, saliva, urine, and beard compared to drug-induced effects on pupils and behavior. *J Anal Toxicol* 14(1):1-7, 1990.

Cone, E.J.; Yousefnejad, D.; Darwin, W.D.; and Maguire, T. Testing human hair for drugs of abuse. II. Identification of unique cocaine metabolites in hair of drug abusers and evaluation of decontamination procedures. *J Anal Toxicol* 15(5):250-255, 1991.

Dekio, S., and Jidoi, J. Amounts of fibrous proteins and matrix substances in hairs of different races. *J Dermatol* 17(1):62-64, 1990.

Forman, R.; Schneiderman, J.; Klein, J.; Graham, K.; Greenwald, M.; and Koren, G. Accumulation of cocaine in maternal and fetal hair: The dose response curve. *Life Sci* 50(1):1333-1341, 1992.

Frank, D.A.; Baucher, H.; Parker, S.; Huber, A.M.; Kyei-Aboagye, K.; Cabral, H.; and Zuckerman, B. Neonatal body proportionality and body composition following in utero exposure to cocaine and marijuana. *J Pediatr* 117:662-666, 1990.

Graham, K.; Koren, G.; Klein, J.; Schneiderman, J.; and Greenwald, M. Determination of gestational cocaine exposure by hair analysis. *JAMA* 262(23):3328-3330, 1989.

Henderson, G.L.; Harkey, M.R.; Zhou, C.; and Jones, R.T. Cocaine and metabolite concentrations in the hair of South American coca chewers. *J Anal Toxicol* 16(3):199-201, 1992.

Holden, C. Hairy problems for a new drug testing method. *Science* 249:1099-1100, 1990.

Keegan, A. Putting hair to the test. *NIDA Notes* 6(4):10-12, 1991-92.

Kopp, C.B., and Kaler, S.R. Risk in infancy: Origins and implications. *Am Psychol* 44(2):224-230, 1989.

Kosofsky, B.E. The effect of cocaine on developing human brain. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Controlled Studies on Effects of Prenatal Exposure to Drug Abuse.* National Institute on Drug Abuse Research Monograph 114. DHHS Pub. No. (ADM)91-1837. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Marques, P.R.; Tippetts, A.S.; and Branch, D.G. Cocaine in the hair of mother-infant pairs: Quantitative analysis and correlations with urine measures and self-report. *Am J Drug Alcohol Abuse* 19(2):159-175, 1993.

McLellan, A.T.; Kushner, H.; Metzger, D.; Peters, R.; Smith, I.; Grissom, G.; Pettinati, H.; and Argeriou, M. The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat* 9:199-213, 1992.

Millon, T. *Millon Clinical Multiaxial Inventory-II*. Minneapolis, MN: National Computer Systems, 1987. 249 pp.

Möller, M.R.; Fey, P.; and Rimbach, S. Identification and quantitation of cocaine and its metabolites, benzoylecgonine and ecgonine methyl ester, in hair of Bolivian coca chewers by gas chromatography/mass spectrometry. *J Anal Toxicol* 16(5):291-296, 1992.

Nakahara, Y.; Ochiai, T.; and Kikura, R. Hair analysis for drugs of abuse. V. The facility in incorporation of cocaine into hair over its major metabolites, benzoylecgonine and ecgonine methyl ester. *Arch Toxicol* 66(6):446-449, 1992.

Ostrea, E.M., Jr. Detection of prenatal drug exposure in the pregnant woman and her newborn infant. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Epidemiological, Prevention, and Treatment Research on Drug-Exposed Women and Their Children*. National Institute on Drug Abuse Research Monograph 117. DHHS Pub. No. (ADM)92-1881. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992.

Rosenberg, M. *Society and the Adolescent Self-Image.* Princeton, NJ: Princeton University Press, 1965.

Sokol, R.J.; Ager, J.; Martier, S.; Debanne, S.; Ernhart, C.; Kuzma, J.; and Miller, S.I. Significant determinants of susceptibility to alcohol teratogenicity. *Ann N Y Acad Sci* 477:87-102, 1986.

Suomi, S.J., and Higley, J.D. Rationale and methodologies for developing nonhuman primate models of prenatal drug exposure. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Controlled Studies on Effects of Prenatal Exposure to Drug Abuse*. National Institute on Drug Abuse Research Monograph 114. DHHS Pub. No. (ADM)91-1837. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Taylor, R. The telltale part: Combing for drugs in hair samples. *J NIH Res* 2(5):28,30, 1990.

Welch, M.J.; Sniegoski, L.T.; Allgood, C.C.; and Habram, M. Hair analysis for drugs of abuse: Evaluation of analytical methods, environmental issues, and development of reference materials. *J Anal Toxicol* 17(7):389-398, 1993.

Zuckerman, B. Discussion: Drug effects—A search for mechanisms. In: Kilbey, M.M., and Asghar, K., eds. *Methodological Issues in Controlled Studies on Effects of Prenatal Exposure to Drug Abuse.* National Institute on Drug Abuse Research Monograph 114. DHHS Pub. No. (ADM)91-1837. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

AUTHOR

Paul R. Marques, Ph.D.
Director of Drug Research
National Public Services Research Institute
8201 Corporate Drive, Suite 220
Landover, MD 20785
(301) 731-9891, ext. 102 (Tel)
(301) 731-6649 (Fax)
marques@pire.org (Internet)

The NICU Network Neurobehavioral Scale: A Comprehensive Instrument To Assess Substance-Exposed and High-Risk Infants

Edward Z. Tronick and Barry M. Lester

INTRODUCTION

This chapter summarizes the authors' work in the development of a comprehensive neurobehavioral instrument, the NICU (neonatal intensive care unit) Network Neurobehavioral Scale (NNNS) (Lester and Tronick 1993). The NNNS was designed to evaluate the neurobehavioral and neurological profiles, adaptation to stress, and withdrawal status of neonates exposed to illegal drugs in utero.

Work on this instrument grew out of research on newborn assessment and training for the Perinatal-20 project, a set of 20 National Institute on Drug Abuse (NIDA)-sponsored intervention studies of cocaine- and other substance-abusing mothers, and the specific requirements of research on the Maternal Lifestyles Study, a National Institute of Child Health and Human Development-sponsored multisite study of the development of cocaine-exposed infants from four NICUs. Investigators on these projects, as well as on several other projects related to prenatal cocaine exposure on which the authors are working, were confronted with a similar set of assessment problems that were unresolved by extant biobehavioral instruments.

These problems included a need for an instrument that would detect and describe the possibly unique neurobehavioral characteristics of infants exposed to cocaine and other substances in utero (Tronick et al., in press). Specialized instruments have already been developed for different populations of newborns: full-term healthy infants (Brazelton 1984; Tronick and Brazelton 1975, pp. 137-156), prematurely born infants (Als et al. 1982, pp. 85-132), infants with intrauterine growth retardation (IUGR) (Als et al. 1976), opiate-exposed infants (Finnegan 1986, pp. 122-146), cross-cultural samples (Dixon et al. 1982; Tronick and Winn 1992), and other groups (Tronick 1987). These instruments were developed because investigators found that, although these heterogeneous populations shared many neurobehavioral features, they also had unique neurobehavioral characteristics that are not adequately captured on existing measurement tools. Similarly, the specific instruments commonly

used to assess these populations might not adequately describe the neurobehavioral features of in utero drug-exposed infants. In the authors' experience, available instruments did not appear to be adequate to describe the neurobehavioral organization of cocaine-exposed infants (Beeghly and Tronick 1994).

A second problem was the need for a single instrument that would assess the neurobehavioral and neurological profiles, adaptation to stress, and the withdrawal status of the neonate. Such an instrument must meet standard criteria for reliability of administration, scoring, and validity, which are not necessarily characteristics of existing instruments.

INSTRUMENT SELECTION

To meet these needs, the authors first drew on extant instruments and then from state-of-the-art instruments that are well established and have been used extensively. The authors wanted the data from a new instrument to be comparable to the data generated by its parent instrument, insofar as the new instrument incorporated the approaches and techniques of its parent. The following core instruments were used: (1) Neonatal Behavioral Assessment Scale (NBAS) (Brazelton 1984), (2) Neurological Examination of the Full-Term Newborn Infant (Prechtl 1977), (3) Neurological Examination of the Maturity of Newborn Infants (Amiel-Tison 1968), (4) Neurobehavioral Assessment of the Preterm Infant (Korner and Thom 1990), and (5) the authors' scale for the Assessment of Preterm Infants' Behavior (APIB) (Als et al. 1982, pp. 85-132). These scales have different goals. The NBAS and APIB were developed to document the behavioral repertoire of the full-term and preterm infant, respectively. Prechtl and Amiel-Tison wanted to assess the neurological status of the newborn, whereas Korner and Thom wanted to evaluate the relative maturity of the infant. Because the authors and other researchers on the ongoing projects also wanted to document any withdrawal symptoms or stress associated with in utero cocaine exposure, Finnegan's (1986) Neonatal Abstinence Score was drawn on. The scales developed by Dubowitz and coworkers (1970) and Ballard and colleagues (1991) also provided guidance. Together, these instruments provide a detailed description of many characteristics that need to be assessed in in utero-exposed infants that is more comprehensive than that provided by any one instrument alone.

Second, it was necessary to limit the number of items because administering all the instruments individually would place a tremendous burden on the newborn and the examiner. The approach was to use the structure of a core examination, add items if necessary, and develop multiple, domain-specific

coding schemes. Fortunately, limiting the number of items administered was facilitated by the fact that many items in different assessments are similar, if not identical. In addition, many items that are scored in one scale (e.g., abnormal posture from the Prechtl examination [Prechtl 1977]) can be observed during the administration of another (e.g., the NBAS), although the item may not be formally elicited or scored during the administration of the other examination. Observation of these items often can be made without having to make significant changes in the scale being administered. Thus, it was possible to conceive of a "generic" examination that used many overlapping items from various extant scales and also permitted the elicitation, observation, and scoring of a wide range of features of the infant's behavior.

The NBAS was chosen as the core examination: It incorporates behavioral, neurological, and stress measures and is the widely used benchmark in the field. The NBAS had been used as the core examination for the APIB, and the authors had experience in adapting it for studies of other populations (e.g., infants with IUGR [Als et al. 1976]) and cross-cultural samples (Dixon et al. 1982). Moreover, many of its items were drawn from the other scales. (For example, the concept and scoring of behavioral state by the NBAS are based on Prechtl's [1977] work.) The authors valued its concept of eliciting best performance, flexibility, and semistructured administration.

SCORING REFINEMENTS

Once the NBAS had been selected as the core examination, several tasks remained. The first was to add items that were not part of the NBAS (e.g., items assessing withdrawal and behavioral, neurological, and stress measures). The second was to develop scoring systems for different domains of assessment—neurological status, stress, and withdrawal. These domain-specific scoring systems were based on items that were elicited and observed during the administration of the examination (e.g., a separate neurological score based on all infant reflexes and postures and a stress scale based on observation of the infant's stress signs). For some domain-specific scoring systems, the scoring was based on added items; in other cases, some items were used and scored for more than one domain of functioning. For example, tremors are scored in the motor maturity items of the NBAS as an item for both the stress and neurological scales. Scoring of these multiple-domain items is often different for each scale because of differences in focus among the scales. (For example, tremors are related to state in the NBAS but not in the stress scale.) This multiple coding of items is burdensome for the examiner, but it allows for maximal use of the information generated during the administration

of the examination. The third task was to add scores to items to describe either hypothesized features of drug-exposed infants or characteristics suggested by research that were not captured by the original scoring system. This approach was based on work with the NBAS in cross-cultural studies in which scores were added to specific items to capture a heretofore undescribed characteristic (e.g., the additional score on the standard NBAS motor maturity scale developed by Tronick and colleagues [Dixon et al. 1982; Keefer et al. 1982] to describe the unique balance and strength of motor behavior among Gusii infants of Kenya). An advantage of this approach is that the original scoring remains intact, and scores still can be compared across studies. The prevalence of the additional score specifies a unique feature of the population. For example, several scores were added to the habituation scales from the NBAS to capture the lack of response observed in exposed infants.

ADMINISTRATION PROCEDURES

It first was decided that the examination would be administered in a semistructured manner. In an unstructured examination, different examiners may conduct the examination differently and elicit different behavioral qualities in the infant, whose scores can be affected by these stylistic differences, reflecting an examiner-infant interaction rather than the infant's performance when faced with a standard challenge. However, the solution is not a rigidly standard examination because an inflexible examination does not elicit the infant's best performance. The authors' solution was to define and limit when items could be administered based on the infant's state. Second, we decided to have a relatively invariant sequence. However, required deviations from the standard sequence are scored. Variations on the sequence thus become data rather than errors among examiners.

CONCLUSIONS

These are the major features of the NNNS. It is designed to provide a comprehensive examination of normal full-term infants, preterm infants, and especially infants at risk because of prenatal substance exposure. Infants must be medically stable, and although a precise lower gestational age limit cannot be set, the examination is appropriate for infants as young as 34 weeks. The upper limit may be 44 weeks, but it may prove to be useful with older infants who were or are medically compromised. The scale's reliability has been demonstrated, and its concurrent and predictive validity is being evaluated in the Maternal Lifestyles Study

with 1,000 infants (also see Napiorkowski et al., in press). A complete manual, film, and training are available from the authors.

REFERENCES

Als, H.; Lester, B.M.; Tronick, E.Z.; and Brazelton, T.B. Towards a research instrument for the assessment of preterm infants' behavior (A.P.I.B.). In: Fitzgerald, H.E.; Lester, B.M.; and Yogman, M.W., eds. *Theory and Research in Behavioral Pediatrics.* Vol. 1. New York: Plenum, 1982.

Als, H.; Tronick, E.; Adamson, L.; and Brazelton, T.B. The behavior of the full-term yet underweight newborn infant. *Dev Med Child Neurol* 1:590-602, 1976.

Amiel-Tison, C. Neurological evaluation of the maturity of newborn infants. *Arch Dis Child* 43:89-93, 1968.

Ballard, J.L.; Khoury, J.C.; Wedig, K.; Wang, L.; Eilers-Walsman, B.L.; and Lipp, R. The new Ballard score, expanded to include extremely premature infants. *J Pediatr* 119:417-423, 1991.

Beeghly, M., and Tronick, E. Effects of prenatal exposure to cocaine in early infancy: Toxic effects on the dyadic regulatory infant-caregiver system. *Infant Ment Health J* 15:281-286, 1994.

Brazelton, T.B. *Neonatal Behavioral Assessment Scale*. 2d ed. Clinics in Developmental Medicine. No. 88. London: Spastics International Medical Publications; Philadelphia: Lippincott, 1984.

Dixon, S.; Keefer, C.; Tronick, E.; and Brazelton, T.B. Perinatal circumstances and newborn outcome among the Gusii of Kenya: Assessment of risk. *Infant Behav Dev* 5:11-32, 1982.

Dubowitz, L.M.; Dubowitz, V.; and Goldberg, C. Clinical assessment of gestational age in the newborn infant. *J Pediatr* 77(1):1-10, 1970.

Finnegan, L.P. Neonatal abstinence syndrome: Assessment and pharmacotherapy. In: Rubaltelli, F.F., and Granati, B., eds. *Neonatal Therapy: An Update*. New York: Excerpta Medica, 1986.

Keefer, C.H.; Tronick, E.; Dixon, S.; and Brazelton, T.B. Specific differences in motor performance between Gusii and American newborns and a modification of the Neonatal Behavioral Assessment Scale. *Child Dev* 53(3):754-759, 1982.

Korner, A.F., and Thom, V.A. *Neurobehavioral Assessment of the Preterm Infant*. New York: The Psychological Corporation, 1990.

Lester, B.M., and Tronick, E.Z. "The NICU Network Neurobehavioral Scale." National Institute of Child Health and Human Development contract N01-HD-2-3159. Unpublished manuscript, Brown University School of Medicine, 1993.

Napiorkowski, B.; Lester, B.M.; Frier, C.; Brunner, B.A.; Dietz, B.A.; Nadra, M.S.; and Oh, W. Effects of in utero substance exposure on infant neurobehavior. *Pediatrics*, in press.

Prechtl, H.F.R. *The Neurological Examination of the Full Term Newborn Infant*. 2d ed. *Clinics in Developmental Medicine*. No. 63. London: Spastics International Medical Publications with Heinemann; Philadelphia: Lippincott, 1977.

Tronick, E. The Neonatal Behavioral Assessment Scale as a biomarker of the effects of environmental agents on the newborn. *Environ Health Perspect* 74:185-189, 1987.

Tronick, E., and Brazelton, T.B. Clinical uses of the Brazelton Neonatal Assessment. In: Friedlander, B.Z., ed. *The Exceptional Child*. Vol. 3. New York: Bruner/Mazel, 1975.

Tronick, E., and Winn, S.A. The neurobehavioral organization of Efe (pygmy) infants. *J Behav Dev Pediatr* 13(6):421-424, 1992.

Tronick, E.Z.; Frank, D.A.; Cabral, H.; Mirochnick, M.; and Zuckerman, B. Late dose response effects of prenatal cocaine exposure on newborn neurobehavioral performance. *Pediatrics*, in press.

ACKNOWLEDGMENTS

Work on this project was supported by NIDA grants R01-DA-06532 to Deborah Frank, R01-DA-06882 to Edward Z. Tronick, and N01-HD-2-3159 to Barry M. Lester. The authors also acknowledge the observational and practical contributions of T. Berry Brazelton, Zach Boukydis, and Kevin Nugent to the development of the NNNS and the conceptual and editorial work of Marjorie Beeghly on the manuscript.

AUTHORS

Edward Z. Tronick, Ph.D.
Chief
Child Development Unit
Children's Hospital
Associate Professor of Pediatrics
Harvard Medical School
300 Longwood Avenue
Boston, MA 02115
(617) 355-6948 (Tel)
(617) 859-7215 (Fax)
tronick@a1.tch.harvard.edu (Internet)

Barry M. Lester, Ph.D.
Professor of Pediatrics
Department of Pediatrics
Women and Infants' Hospital
Brown University School of Medicine
101 Dudley Street
Providence, RI 02905
(401) 453-7640 (Tel)
(401) 453-7646 (Fax)
barry_lester@brown.edu (Internet)

Measurement of the Early Rearing Environment: Caregiver-Child Interaction

Leila Beckwith

INTRODUCTION

This chapter identifies measures that can be used to capture status and change in the quality and quantity of early caregiver-child interactions. Because maternal drug abuse often necessitates the use of alternative caregivers for the child, through either legal or informal arrangements, the measures discussed herein are suitable for the biological mother, father, or grandparent; professional foster parent; or any other designated primary caregiver. The discussion considers the significance of such measures in evaluating the efficacy of maternal intervention programs and in understanding the sequelae of prenatal drug exposure for children. The chapter also reviews the conceptual dimensions of caregiving, assessment procedures used in the home and laboratory, and level of analysis, whether by behavioral counts or rating scales. Comments on the strengths and pitfalls of specific measures also are included.

Caregiving interaction between adult and child is manifested in a variety of different contexts and can be measured using a variety of data collection strategies; however, this chapter is not exhaustive. Deliberately excluded are discussions of measures of parental beliefs and attitudes as well as parental self-reports of caregiving behavior, which were obtained through questionnaires or interviews. Rather, the chapter focuses specifically on observed transactions as appraised by objective researchers. Selected measures are highlighted because they have been used in longitudinal, normative studies or in investigations of children of drug-abusing mothers or of other at-risk groups, for example, preterm infants, children from low socioeconomic status (SES) families, and children of depressed mothers.

EVALUATING INTERVENTION

The study of caregiver-child interaction has three goals: to (1) describe how the caregiving is done, (2) understand the etiology of individual differences among caregivers, and (3) determine the consequences of caregiving for the developing child (Belsky 1984). Each goal is significant in evaluating intervention with drug-abusing mothers.

Although little systematic, empirical study has been done of deficits in caregiver ability associated with drug abuse, clinical studies (Escamilla-Mondanaro 1977; Rosenbaum 1979) and research with other populations with psychopathology (e.g., maternal depression) (Cohn et al. 1986, pp. 31-46; Field et al. 1985; Lyons-Ruth et al. 1986, pp. 61-82; Tronick and Field 1986) suggest that maladaptive parenting is likely in several domains. Therefore, to decide what will be addressed by an intervention, it is essential to identify the domains and to understand the internal vulnerabilities and external stresses that influence individual differences within those domains. Lack of clarity in defining what will be changed by intervention is one of the serious impediments to determining effectiveness of intervention.

Recent research indicates that mediating factors that affect individual differences in caregiving reside in the caregiver's personal history in the family of origin, personal psychological resources, contextual sources of stress and support (Belsky 1984), and attitudes and beliefs about children (Sameroff and Fiese 1990, pp. 119-149). Thus, a broad range of foci for intervention is possible, with feasibility and implementation differing among studies.

A comprehensive evaluation of the effectiveness of an intervention whose goal is to generate scientific as well as clinical information should target each of the three goals stated in the beginning of this section. That is, the evaluation should assess changes in dimensions of caregiving, changes in the presumed mediating factors, and changes in the hypothesized sequelae for the developing child.

DEVELOPMENTAL CONSEQUENCES OF PRENATAL DRUG EXPOSURE

Although prenatal drug exposure is associated with a higher occurrence of preterm births, intrauterine growth retardation (Frank et al. 1988; Hadeed and Siegel 1989; Hawley and Disney 1992; Zuckerman et al. 1989), and alterations in neonatal neurobehaviors such as crying (Lester et al. 1991), much controversy still exists as to whether prenatal drug exposure is a causal agent in adverse cognitive, emotional, and social development.

Two issues are in contention. One is long-term outcome, and the other is the exact nature of the causal agent. First, findings conflict as to the later development of children who were exposed in utero to illicit drugs. Some studies report a significant increase in the percentage of such children who function poorly in cognitive abilities (Chasnoff et al. 1992; Lifschitz et al. 1985), whereas the same studies and others report no differences in mean

scores between the target and comparison groups (Neuspiel and Hamel 1991).

Second, some investigators question whether prenatal drug exposure is the main causal agent or whether co-occurring factors during the prenatal period, such as poor maternal nutrition and lack of prenatal care, and interrelated factors occurring during the postnatal period, such as poverty, maternal depression, unstable home environments, and emotional and physical neglect, are the true causes (Hawley and Disney 1992; Lifschitz et al. 1985; Myers et al. 1992; Zuckerman and Frank 1992).

The issues have implications for both developmental theory and social policy. To date, however, the debate remains philosophical and mainly unaddressed by empirical research. Few studies have examined "the critical issue of the quality of the home environments of children of addicted mothers" (Hawley and Disney 1992).

The premise of this chapter is that no conclusions can be drawn in advance of research. Disentangling the influence of prenatal drug exposure from multiple postnatal influences can be done only through the systematic assessment of group and individual rearing conditions for children who were and were not exposed in utero to drugs and who are and are not being brought up in drug-abusing families. Although such an aim is ambitious, multiple individual studies can begin to contribute an informed answer.

DECISION ABOUT DOMAINS OF CAREGIVING

Assessing rearing conditions requires deciding which of the multiple roles that caregivers fulfill with infants and young children will be theoretically relevant and feasible for a specific study.

Caregivers provide protection (Ainsworth 1973, pp. 1-94; Bowlby 1969); provide comfort for distress (Ainsworth et al. 1978; Belsky and Isabella 1985; Del Carmen et al. 1993; Sroufe 1985); elicit and maintain positive social-affective exchanges, including smiling and vocalizing (Blehar et al. 1977; Campos et al. 1983, pp. 783-916); encourage environmental exploration and learning (Yarrow et al. 1975); and exercise control (Maccoby and Martin 1983, pp. 1-101). Caregivers also ensure physiological homeostasis through the concrete provisions of food, cleanliness, and warmth. Thus, the caregiver functions as nurturer, protector, comforter, playmate, teacher, and governor, among other roles.

Caregivers also provide the inanimate aspects of the environment: The number and kinds of toys, novelty or familiarity and variety of settings,

and degree of crowding also influence children's affective and cognitive development (Wachs and Gruen 1982).

Which domains should be measured? Most assessment procedures concentrate on one or more dimensions, but no procedure covers all. The choice of measures depends on which functions are of specific interest for a particular study. The selection might be guided by a focus on functions that are believed to be altered by drug abuse (e.g., affect management, protection). The selection also might be determined by considering functions that influence specific developmental outcomes of interest (e.g., facilitation or inhibition of environmental exploration as it affects a child's intelligence quotient [IQ]) or more general functions important to every child, such as those that underlie the quality of attachment between infant and caregiver (e.g., caregiver sensitivity).

There are wide individual differences within groups in how caregivers fulfill each role. Moreover, it is important to know similarities and differences among groups for children who were and were not prenatally drug exposed and who are and are not growing up in drug-abusing families. Similarities may exist in some domains of caregiving experience; differences may exist in others. Only when multiple domains are examined can there be a precise understanding of the relationship between drug abuse and caregiving.

FIELD-LABORATORY DECISION

Research contexts for the assessment of the rearing environment vary in the degree to which they approximate the child's naturalistic physical and social environment (Parke 1979, pp. 15-36). Multiple variations can be produced through alterations in either or both the physical and social domains. A variety of strategies are possible, and the choices are dictated by the experimental questions, the skill of the investigator, and the study's resources in terms of time and money.

The procedure that most approximates the natural environment is a field study in the home in which the caregiver and other members of the household, including siblings, are asked to go about their usual activities with the child just as if a stranger were not present. No experimental manipulation is introduced, and thoughtful consideration is given to diminish the influence of the observer (and, often, the video camera). The significant advantage of such a study is that it most closely reveals the animate and inanimate events within the natural environment that shape the development of the child. This procedure has been used to marked advantage to describe individual and group differences in rearing

environments in studies of children born preterm (e.g., Beckwith and Cohen 1984, pp. 235-271; Thoman et al. 1981). To date, few such studies have been published that used samples of drug-abusing mothers or infants who were prenatally exposed to illicit drugs. Much more work needs to be done.

Other assessments in the home, while maintaining the familial context, may alter everyday routines by instructing the caregiver to carry out a particular activity such as feeding, diapering, playing, or teaching. Such a procedure has the advantage of standardizing the observational context for each subject and ensuring that important activities are observed equally for each subject. Whereas the activities are usually chosen to duplicate common routines, some caution is necessary because ecological validity may be altered by the artificiality introduced in the timing. Timing, whether sensitive or insensitive to the child's cues, and predictability or unpredictability of the routine may be as important to the child's experience as the quality of the activity. Moreover, the focus on predetermined interactions obscures observation of other events in the environment that are perhaps equally important.

Partial approximations of the social environment also can be conducted in the laboratory. Procedures can vary from (1) unstructured "free play," to (2) specific feeding or teaching interactions, to (3) situations in which the caregiver is instructed to behave in an unusual manner. The "face-to-face" procedure as well as the "strange situation" (described below) are examples of the third scenario.

Paradigms that simultaneously impose unfamiliar contexts and unusual caregiver behaviors have a goal different from approximating the natural social environment. Their aim is to perturb the ordinary routines in the belief that adapting to perturbations is more revealing of individual differences than reacting to familiarity. The principle, as demonstrated in more than a decade of research with the "strange situation" (Ainsworth et al. 1978), has generated a powerful tool. However, as pointed out by Ainsworth, only naturalistic observations in the home can demonstrate the ultimate validity of the laboratory procedure.

Thus, to more accurately determine the rearing environments associated with drug abuse, naturalistic studies in the home and experimental perturbations are required. Practical considerations also guide the choices. Home observations may involve safety issues for the staff as well as the families. Some neighborhoods may be too dangerous for a single staff member to enter or make observations because of increased gang activity and drug-dealing in the late afternoon and evening. Home observations

may fail because subjects who abuse drugs may forget appointments, may not be awake and ready, or may be preoccupied with activities of drug use.

Laboratory visits may be equally difficult. Transportation to the laboratory as well as child care for other children may have to be provided. The staff members may have to help ensure the subjects' readiness by waking the mothers, organizing their preparations, and ensuring the children's dressing and feeding. The laboratory staff members must provide diapers, formula, snacks, and meals. Furthermore, subjects may seek the researchers' attention for themselves or fail to soothe and help their children feel comfortable in the laboratory situation. The decision whether to use the home or laboratory depends on both theoretical and practical considerations, which in turn will affect the knowledge gained.

METRIC DECISION: BEHAVIORAL COUNTS, RATING SCALES, OR Q-SORTS

The choice of measures depends not only on decisions about the domains of caregiving to be investigated, the theoretical rationale and feasibility of the setting, and the structure of the procedure but also on the method of quantification. Observational techniques can be quantified by behavioral counts, rating scales, or Q-sorts (Block 1978).

The use of behavioral counts as the method of quantification can reveal how frequently specific behaviors or events occur, the frequency of affects accompanying the behaviors, and the sequence of events. The advantage of behavioral counts is that they produce a precise, albeit incomplete, record of what occurred during an observation. The disadvantage is that they are affected by transient conditions and may be less stable over time and situation than rating scales. Behavioral counts also require complex analysis procedures to meaningfully integrate specific, discrete behaviors.

In contrast, rating scales automatically summarize discrete behaviors over time as well as suppress or ignore deviations from a subject's central tendency (Barnard and Kelly 1990, pp. 278-302). Statistical analysis becomes simpler because the observer becomes a filtering and integrating tool (Block 1978). However, to the extent that observers synthesize events, ratings tend to be subject to differences among observers in how much weight they give to specific behaviors; thus, the bases of the ratings may remain idiosyncratic or unspecified to some degree.

Q-sorts (Block 1978) retain the focus on the observer as the means of synthesis. However, rather than being asked to measure aspects of a subject's behavior against a presumed or specified norm, the observer

measures an array of specified behaviors against each other for a particular subject. Q-sorts include many more enumerated behaviors than do rating scales, thus making the differences among subjects and groups more specifiable. Q-sorts also diminish some biases that trouble rating scales, including halo effects and differential preferences among observers for the middle or extremes of the scale. However, Q-sorts are complex to develop and difficult to score. Although no Q-sort measures are described in this chapter (the ones in common use have different aims), interested readers should consult Block and Block (1980, pp. 39-101) and Waters and Deane (1985).

Whether behavioral counts, rating scales, or Q-sorts are the measure chosen, their usefulness ultimately depends on the care and insight of the observer. To reach a high level of skill and to establish interobserver and test-retest reliability, observers must be trained for each study.

HOME ASSESSMENTS

There are several naturalistic observation procedures that differ in duration, time of day, behaviors encoded, and coding complexity (Beckwith and Cohen 1984, pp. 235-271; Bakeman and Brown 1977; Bornstein and Tamis-LaMonda 1989, pp. 49-62; Clarke-Stewart 1973; Del Carmen et al. 1993; Lewis and Coates 1980; Thoman et al. 1981; Yarrow et al. 1975). The procedures also differ in the use of preestablished behavior categories or narrative reports, time sampling (e.g., observer views occurrence or nonoccurrence of specified behaviors for 10 seconds and records during the next 20 seconds) or event sampling (occurrence of behaviors is recorded regardless of when they occur), and electronic recording devices, audiotape recordings, stenographer's books, or paper checklists.

Regardless of procedure, the following domains can be measured: involvement (i.e., amount of caregiver-child interaction); amount and quality of physical, verbal, and visual contact; amount of responsivity of caregiver to child; degree of caregiver restrictiveness of exploration; degree of caregiver intrusiveness; kinds of control strategies; and caregiver affect. The child also can be assessed as to amount and kind of social bids, positive and negative affects, responsiveness to caregiver, amount of environmental exploration, and regularity of sleeping-awake states.

The procedures tend to be time consuming: Observations are precise and intense; statistical analyses often are laborious; and data reduction is challenging. Yet, naturalistic observations hold the promise of detailing objectively, minute by minute, caregiver-child behavioral sequences as

they differ in children who were and were not exposed in utero to illegal drugs and who are and are not being reared in drug-abusing households.

System for Rating Maternal-Care Behavior

This system consists of scales (Ainsworth et al. 1978) (available in microfiche form from the Educational Testing Service, Princeton, NJ 08540) that rate maternal behavior during the infant's first year. They were derived from careful examination of detailed narrative reports of repeated home observations of a normative sample and then were used to predict later quality of attachment and social behavior in that sample (Ainsworth et al. 1978) as well as in longitudinal studies of preterm infants (Goldberg et al. 1986), low-SES full-term infants (Egeland and Farber 1984; Sroufe 1983, pp. 41-84), and children exposed in utero to phencyclidine (PCP) and cocaine (Rodning et al. 1991).

These rating scales contain 26 nine-point subscales, with points 1, 3, 5, 7, and 9 anchored in detailed behavioral descriptions. The areas rated include general attitude of caregiver toward baby and the caregiver role, feeding, availability and interaction, physical contact, response to crying, social contact, facilitation of sensorimotor development, sensitivity-insensitivity to child signals, acceptance-rejection, cooperation-interference, and accessibility-ignoring. Many of the subscales tend to be highly intercorrelated, which allows the use of a few selected scales.

Home Observation for Measurement of the Environment

This inventory (Bradley and Caldwell 1984; Caldwell and Bradley 1978) is a combination observation and interview lasting about 1 to 2 hours. It is administered in a child's home when the child is present and awake, during which time the child's primary caregiver is interviewed. There are two versions of the Home Observation for Measurement of the Environment (HOME), one for use with families of infants from birth to 3 years of age and one for use with families of preschoolers from 3 to 6 years of age. The HOME is a rating scale that scores items in a binary (yes-no) manner. Because the items are well specified and concrete, the task is made easy for the rater, and rater bias is reduced.

The HOME measures some inanimate aspects of the environment as well as some facets of the caregiver-child relationship. The items are clustered into six subscales: emotional and verbal responsivity of caregiver, avoidance of restriction and punishment, organization of the physical and temporal environment, provision of appropriate play materials, caregiver involvement with child, and opportunities for variety in daily stimulation.

Since its initial development, the HOME has become one of the most widely used measures in child development research for evaluating the quality of a child's early physical and social environment (Gottfried 1984). A wealth of information is available about internal consistency, test-retest reliability, and the influence on HOME scores of family demographics, including SES, ethnic group, birth order, and sex. In addition, multiple studies with diverse samples have effectively demonstrated predictive and concurrent associations between HOME scores and cognitive performance from infancy to school age.

The HOME is reliable, valid, and easy to use and has the advantage of measuring physical as well as social dimensions of the caregiving environment. One caution: For some studies or samples, it may not be sufficiently sensitive to individual differences. The one published study that used the HOME to assess the early caregiving environments of infants who were prenatally exposed to drugs found equivalent scores for drug-abusing mothers, comparison mothers who did not abuse drugs, and professional foster parents, although the Ainsworth rating scales did differentiate (Rodning et al. 1991).

Purdue Home Stimulation Inventory

Wachs (1984, pp. 273-328) has been a strong exponent of a theory of environmental specificity, which contrasts with global models and postulates that different aspects of the environment influence different facets of later development. The effects are dependent on the age of the child and are mediated by the individual characteristics of the child. Thus, the use of multidimensional predictors and criteria, across distinct subgroups of children and across ages, becomes the recommended research design.

Derived from that model, the Purdue Home Stimulation Inventory (PHSI) is designed to assess specific animate and inanimate features of the environment. During the observations (approximately 45 minutes long), the observer dictates into a tape recorder all relevant child-environment, child-person interactions. Also, every 15 minutes the observer uses a precoded checklist to check off appropriate categories for ongoing aspects of the child's physical environment. Whereas the sections of the scale that assess the social features of the environment are complicated to observe and analyze, the section that assesses the physical aspects is much simpler and can be used independently. Items measured are availability and variety of stimulus material, responsivity of the physical environment, noise-confusion, crowding, regularity of temporal and spatial scheduling, and physical restriction of exploration.

Because the inanimate aspects of the environment are an overlooked domain of influence in children's development and because few measures exist for such assessment, the PHSI could be particularly useful in researching the effects of maternal drug abuse on children, particularly because noise-confusion and regularity of temporal and spatial scheduling are probably altered by drug abuse.

ASSESSMENTS IN EITHER HOME OR LABORATORY

Nursing Child Assessment Satellite Training Teaching and Feeding Scales

The Nursing Child Assessment Satellite Training (NCAST) scales (Barnard 1979) assess the quality of the interaction between caregiver and child, ages 1 to 36 months, in the familiar situation of feeding and in a more novel situation in which the caregiver is asked to teach the child two tasks, one at the child's age level and one somewhat beyond the expected ability of the child. The feeding scale comprises 73 items that the observer judges did or did not occur during the feeding situation, and the teaching scale comprises 76 yes-no items that did or did not occur during the teaching tasks. The items for each scale are clustered and generate either a total score (a score for the caregiver and a score for the child) or scores for six subscales, four of which describe caregiver behavior (including sensitivity to child cues, responsiveness to child distress, fostering social-emotional growth, and fostering cognitive growth) and two of which describe child behavior (clarity of cues and responsiveness to caregiver) (Barnard and Kelly 1990, pp. 278-302). One additional score, contingent responsiveness, can be generated for the teaching scale.

These rating scales have been widely used in research with diverse groups of children, including preterms and full terms; longitudinal as well as normative data exist. The scales have the advantage of being highly structured; the behaviors to be noted are clearly specified; and judgments are simplified because they depend only on occurrence or nonoccurrence rather than matching against an implied norm. The use of the scales requires direct training by an NCAST-certified instructor.

Parent/Caregiver Involvement Scale

When this scale (Farran et al. 1986) is used, interactive play of at least 20 minutes is either videotaped in the laboratory or observed at home, and 11 domains of the caregiver's behavior with an infant or young child up to age 5 are assessed: physical involvement, verbal involvement, responsiveness, play, teaching, control, directives, relationship among

activities, positive emotions, negative emotions, and goal setting. Each domain is rated on 5-point scales as to amount, quality, and appropriateness. Amount notes quantity of the behavior; quality judges smoothness, pacing, and affect; and appropriateness rates the match of the caregiver's behavior to the developmental skill and interest levels of the child. Summary scores may be created for amount, quality, and appropriateness ratings collapsed across the 11 behaviors, or profiles of individual scores can be analyzed for each of the 11 domains of adult behavior for each of the three qualities.

The rating scales are "behaviorally anchored" with descriptions at three of the five points of each scale. However, the numerous items make scoring by memory, without videotape, difficult and subject to distortion. Also, there is no assessment of child characteristics or individual differences in the children's contribution to the interactions.

This scale has been used in research studies in the development of diverse groups of children, including those from low-SES families and those with handicaps (Farran et al. 1986; Farran et al. 1987, pp. 299-312). Reliability and validity data are available. A videotape and workbook provide an introduction to the scale and practice sessions in scoring.

Mother-Child Rating Scales

These scales (Crawley and Spiker 1983) were developed originally to detect individual differences in mother-child interactions with 2-year-olds with Down syndrome. Since then, selected scales have been used in the large, multisite Infant Health and Development Program to examine the efficacy of early intervention with preterm children (Spiker et al. 1993) as well as in a study of mothers who abused alcohol during pregnancy (O'Connor et al. 1993). The inclusion of child qualities in the scale proved to be particularly revealing in the latter study in that exposure to alcohol during the prenatal period was linked to increased irritability in the child, which in turn led to diminished maternal elaboration and stimulation.

The scales have been used only with videotaped interactions in either home or laboratory settings. The situations have included free play with standard sets of toys, cleanup, and problemsolving tasks from a paradigm previously used by Matas and colleagues (1978).

The scales consist of 10 child and 6 caregiver characteristics rated on 5-point scales. In addition, four other caregiver behaviors are rated dichotomously, and there is one rating of the dyadic quality of the interactions. Among the child qualities that can be measured are social initiative, social responsivity, interest, object initiative, positive affect,

negative affect, and animation. Caregiver behaviors include directiveness, elaborativeness, sensitivity, stimulation, mood, pacing, developmental appropriateness, and intrusiveness. No specific training is required to use this measure.

LABORATORY ASSESSMENTS

Face to Face

This approach (Brazelton et al. 1974, pp. 49-76; Tronick 1989; Tronick and Weinberg 1990) assesses the vocal/affective communication system between infants 3 to 9 months of age and their caregivers. The procedure has been used effectively, with different scoring systems, to predict later cognitive ability (Roe et al. 1982), to determine the social capacities of preterms (Field 1980, pp. 113-132), and to understand the relationship between depressed mothers and their infants (Cohn et al. 1986, pp. 31-46; Field et al. 1985; Tronick and Field 1986).

As stated by Tronick (1989), the regulation of emotions, self and other; degree of interactive success; and affective and communicative reparation of interactive errors are major influences on the emotions the infant experiences, how well the infant accomplishes its goals, and the infant's developmental outcome.

The infant is placed in an infant seat, and the caregiver is seated facing the infant, within touching distance. No toys are allowed nor are functional activities such as feeding. Both adult and infant are videotaped simultaneously using two cameras. There are typically three episodes 2 to 3 minutes in length: The caregiver is instructed to talk or play with the baby; the interaction is perturbed by either having the caregiver leave and a stranger play with the baby or instructing the caregiver to maintain a "still face" and not to interact; then the caregiver resumes talking and playing with the baby.

The first face-to-face episode highlights the caregiver's and infant's interactive capacities and examines the ability of each to regulate and maintain a mutually satisfying interaction as well as their ability to repair interactive errors. The second episode is a mild stressor and permits the infant's capacity to regulate state and affect and to cope to be discriminated from the caregiver's competency. The third episode, the reunion, examines the infant's ability to use the caregiver as a resource for self-regulation and the caregiver's ability to soothe the infant and to reengage the infant in satisfying interaction.

Scoring the videotapes is done typically on a time-sampling or second-by-second basis. Several coding manuals exist: The most recent are the Maternal Regulatory Scoring System and Infant Regulatory Scoring System (Tronick and Weinberg 1990). The caregiver is scored as to proximity; degree of social, object, and visual engagement; vocalization; touch; and effect on infant state. The infant is scored as to social, object, and visual engagement; vocalization; gesture; self-comfort; distance (avoidance); inhibition; and distress.

Data reduction is complicated, and several approaches have been used, depending on the research questions asked. Analyses have involved all episodes or just the first or second. Multivariate analyses of variance have been used to compare groups on individual codes. Also, the degree of matched states between caregiver and infant has been assessed by deriving the proportion of the interaction that the caregiver and infant spent simultaneously in attending or social play with each other as well as the proportion of time that the caregiver and infant spent simultaneously in jointly attending to an object (Tronick 1989). The rate of change from matched to mismatched states and the rate of repair from mismatched to matched states also have been analyzed.

The procedure holds promise for insights into the early relationships of drug-abusing mothers and other caregivers with infants who have been exposed in utero to alcohol and other drugs. However, the procedure is difficult: Scoring the tapes is time consuming, and data reduction is challenging.

Strange Situation

This is a reliable and valid measure for assessing the quality of the caregiver-child attachment relationship when a child is between ages 12 and 48 months (Ainsworth et al. 1978). The classifications derived have shown associations with antecedent caregiving behavior in the home (Blehar et al. 1977; Goldberg et al. 1986; Grossmann et al. 1985, pp. 233-256), predictive validity to later affective and peer behavior (Matas et al. 1978; Sroufe 1983, pp. 41-84; 1985), and coherence with parental representations of their own relationships to their family of origin (Main and Hesse 1990, pp. 161-182).

The procedure must take place in an unfamiliar locale, such as a laboratory room, and consists of seven 3-minute phases given in invariant order. The sequences are (1) caregiver and child alone in a room with appropriate toys, (2) stranger enters and interacts first with caregiver and then with child, (3) caregiver departs leaving child with stranger, (4) caregiver and child reunite while stranger departs, (5) caregiver departs the second time

leaving child alone, (6) stranger returns, and (7) caregiver and child reunite a second time while stranger departs. The episodes when the child is separated from the caregiver are curtailed if the child is very distressed.

The procedure assesses the quality of the caregiver-child attachment relationship by intentionally introducing a graduated series of perturbations designed to activate the child's attachment system. Although all behavior of the child in the procedure is noted—in part by scoring each phase, as appropriate, as to proximity seeking, contact maintenance, avoidance, and resistance—it is the sequence of behavior during reunions that is particularly informative.

An integrated judgment is made for each child that classifies the organization of the child's attachment behavior as to subgroups within three major groups: (1) secure (B group), (2) insecure-avoidant (A group), or (3) insecure-resistant (C group). An additional classification then is made as to the degree of disorganization and disorientation (D group) (Main and Solomon 1986, pp. 95-124).

The hallmark of children in the B group is their active use of the caregiver as a secure base from which to explore the environment. When distressed, they seek and receive comfort from the caregiver, and when not distressed, they actively greet and initiate positive social exchanges with the caregiver. Insecure-avoidant children do not seek comfort and avoid proximal contact. Insecure-resistant children manifest angry resistant behavior that fluctuates with comfort-seeking, but they are neither soothed by physical contact nor calmed by the presence of the caregiver. Children classified as disorganized and disoriented show contradictory or changing attachment strategies, fear of the caregiver, dazed facial expressions, or stereotypic and anomalous postures and use of space. See Ainsworth and colleagues (1978) for further description of the procedure, instructions about setup of the room and toys, and scoring criteria and classifications.

The procedure is labor intensive (three staff members are needed to administer the paradigm) and also requires a trained person to determine a child's classification. Training requires a thorough knowledge of the theory of attachment plus direct training by experts in the field.

CONCLUSIONS

Review of the diverse measures cited in this chapter indicates that multiple, alternative measures exist through which to better understand the development of children prenatally exposed to drugs and by which to evaluate changes in caregiver behavior associated with early intervention.

No one measure is recommended above any other because measures differ in feasibility, degree of training required, ease of administration, complexity of analysis, sensitivity, concept, and purpose. Some assess the inanimate environment; many others survey the animate environment. Some assess only the caregiver; others assess the child's contribution as well as that of the caregiver. Some assess only behavior; others measure affect as well as behavior.

The author recommends the use of multiple measures to assess child characteristics as well as those of the caregiver, physical and social domains of experience, affect as well as behavior, the contingent nature of the interactions, and continuity over time. No matter what choices are made, interpretation of results must recognize what was not measured as well as what was.

REFERENCES

Ainsworth, M.D.S. The development of infant-mother attachment. In: Caldwell, B.M., and Ricciuti, H.N., eds. *Review of Child Development Research.* Vol. 3. Chicago: University of Chicago Press, 1973.

Ainsworth, M.D.S.; Blehar, M.; Waters, S.E.; and Wall, S. *Patterns of Attachment*. Hillsdale, NJ: Erlbaum, 1978. 391 pp.

Bakeman, R., and Brown, J.V. Behavioral dialogues: An approach to the assessment of mother-infant interaction. *Child Dev* 48:195-203, 1977.

Barnard, K.E. *Instructor's Learning Resource Manual*. Seattle, WA: NCAST Publications, University of Washington, 1979.

Barnard, K.E., and Kelly, J.F. Assessment of parent-child interaction. In: Meisels, S.J., and Shonkoff, J.P., eds. *Handbook of Early Childhood Intervention*. Cambridge, MA: Cambridge University Press, 1990.

Beckwith, L., and Cohen, S.E. Home environment and cognitive competence in preterm children during the first five years. In: Gottfried, A.W., ed. *Home Environment and Early Cognitive Development.* Orlando, FL: Academic Press, 1984.

Belsky, J. The determinants of parenting: A process model. *Child Dev* 55:83-96, 1984.

Belsky, J., and Isabella, R.A. Marital and parental-child relationships in family of origin and marital change following the birth of a baby: A retrospective analysis. *Child Dev* 56:342-349, 1985.

Blehar, M.; Lieberman, A.F.; and Ainsworth, M. Early face-to-face interaction and its relation to later infant-mother attachment. *Child Dev* 48:182-194, 1977.

Block, J. *The Q-Sort Method in Personality Assessment and Psychiatric Research.* Palo Alto, CA: Consulting Psychologists Press, 1978.

Block, J.H., and Block, J. The role of ego-control and ego-resiliency in the organization of behavior. In: Collins, W.A., ed. *The Minnesota Symposia on Child Psychology.* Vol. 13. Hillsdale, NJ: Erlbaum, 1980.

Bornstein, M.H., and Tamis-LaMonda, C.S. Maternal responsiveness and cognitive development in children. In: Bornstein, M.H., ed. *Maternal Responsiveness: Characteristics and Consequences. New Directions for Child Development*. Vol. 43. San Francisco, CA: Jossey-Bass, 1989.

Bowlby, J. *Attachment and Loss: Vol. 1. Attachment*. New York: Basic Books, 1969. 428 pp.

Bradley, R.H., and Caldwell, B.M. The relation of infants' home environment to achievement test performance in first grade: A follow-up study. *Child Dev* 55:803-809, 1984.

Brazelton, T.B.; Koslowski, B.; and Main, M. The origins of reciprocity: The early mother-infant interaction. In: Lewis, M., and Rosenblum, L.R., eds. *The Effect of the Infant on Its Caregiver.* New York: Wiley-Interscience, 1974.

Caldwell, B.M., and Bradley, R.H. *Manual for the Home Observation for Measurement of the Environment.* Little Rock, AR: University of Arkansas, 1978. 107 pp.

Campos, J.J.; Barrett, K.C.; Lamb, M.E.; Goldsmith, H.H.; and Stenberg, C. Socioemotional development. In: Haith, M.W., and Campos, J.J., eds. *Handbook of Child Psychology: Vol. 2. Infancy and Developmental Psychobiology*. New York: Wiley, 1983.

Chasnoff, I.J.; Griffith, D.R.; Freier, C.; and Murray, J. Cocaine/polydrug use in pregnancy: Two-year follow-up. *Pediatrics* 89(2):284-289, 1992.

Clarke-Stewart, K.A. Interactions between mothers and their young children: Characteristics and consequences. *Monogr Soc Res Child Dev* 38(6):1-109, 1973.

Cohn, J.F.; Matias, R.; Tronick, E.Z.; Connell, D.; and Lyons-Ruth, K. Face-to-face interactions of depressed mothers and their infants. In: Tronick, E.Z., and Fields, T., eds. *Maternal Depression and Infant Disturbance. New Directions for Child Development.* San Francisco, CA: Jossey-Bass, 1986.

Crawley, S.B., and Spiker, D. Mother-child interactions involving two-year-olds with Down syndrome: A look at individual differences. *Child Dev* 54(5):1312-1323, 1983.

Del Carmen, R.; Pedersen, F.A.; Hoffman, L.C.; and Bryan, V.E. Dyadic distress management predicts subsequent security of attachment. *Infant Behav Dev* 16:131-147, 1993.

Egeland, B., and Farber, E.A. Infant-mother attachment: Factors related to its development and changes over time. *Child Dev* 55:753-771, 1984.

Escamilla-Mondanaro, J. Women: Pregnancy, children, and addiction. *J Psychedelic Drugs* 9(1):59-68, 1977.

Farran, D.; Kasari, C.; Comfort, M.; and Jay, S. *Parent/Caregiver Involvement Scale*. Greensboro, NC: Child Development-Family Relations, University of North Carolina at Greensboro, 1986. 30 pp.

Farran, D.C.; Kasari, C.; Yoder, P.; Harber, L.; Huntington, G.; and Comfort-Smith, M. Rating mother-child interaction in handicapped and at-risk infants. In: Tamir, D., ed. *Stimulation and Intervention in Infant Development*. London: Freund Publishing, 1987.

Field, T.; Sandberg, D.; Garcia, R.; Vega-Lahr, N.; Goldstein, S.; and Guy, L. Pregnancy problems, postpartum depression, and early mother-infant interactions. *Dev Psychol* 21(6):1152-1156, 1985.

Field, T.M. Interactions of preterm and term infants with their lower and middle-class teenage and adult mothers. In: Field, T.; Goldberg, S.; Stern, D.; and Sostek, A.M., eds. *High-Risk Infants and Children: Adult and Peer Interactions*. New York: Academic Press, 1980.

Frank, D.A.; Zuckerman, B.S.; Amaro, H.; Aboagye, K.; Bauchner, H.; Cabral, H.; Fried, L.; Hingson, R.; Kayne, H.; Levenson, S.M.; Parker, S.; Reece, H.; and Vinci, R. Cocaine use during pregnancy: Prevalence and correlates. *Pediatrics* 82:888-895, 1988.

Goldberg, S.; Perotta, M.; Minde, K.; and Corter, C. Maternal behaviors and attachment in low-birth-weight twins and singletons. *Child Dev* 57:34-46, 1986.

Gottfried, A.W., ed. *Home Environment and Early Cognitive Development.* Orlando, FL: Academic Press, 1984. 383 pp.

Grossmann, K.; Grossmann, K.E.; Spangler, G.; Suess, G.; and Unzner, L. Maternal sensitivity and newborns' orientation responses as related to quality of attachment in northern Germany. *Monogr Soc Res Child Dev* 50(1-2):233-256, 1985.

Hadeed, A., and Siegel, S. Maternal cocaine use during pregnancy: Effect on the newborn infant. *Pediatrics* 84:205-210, 1989.

Hawley, T.L., and Disney, E.R. Crack's children: The consequences of maternal cocaine abuse. Social policy report. *Soc Res Child Dev* IV(4):1-23, 1992.

Lester, B.M.; Corwin, M.J.; Sepkowski, C.; Seifer, R.; Peucker, M.; McLaughlin, S.; and Golub, H.L. Neurobehavioral syndromes in cocaine-exposed newborn infants. *Child Dev* 62:694-705, 1991.

Lewis, M., and Coates, D.L. Mother-infant interaction and cognitive development in twelve-week-old infants. *Infant Behav Dev* 3:95-105, 1980.

Lifschitz, M.H.; Wilson, G.S.; O'Brian Smith, E.; and Desmond, M.M. Factors affecting head growth and intellectual function in children of drug addicts. *Pediatrics* 75(2):269-274, 1985.

Lyons-Ruth, K.; Zoll, D.; Connel, D.; and Grunebaum, H. The depressed mother and her one-year-old infant: Environmental context, mother-infant interaction and attachment, and infant development. In: Tronick, E.Z., and Field, T., eds. *Maternal Depression and Infant Disturbance. New Directions for Child Development.* San Francisco, CA: Jossey-Bass, 1986.

Maccoby, E.E., and Martin, J.A. Socialization in the context of the family: Parent-child interaction. In: Hetherington, E.M., ed. *Handbook of Child Psychology: Vol. 4. Socialization, Personality, and Social Development.* New York: Wiley, 1983.

Main, M., and Hesse, E. Parents' unresolved traumatic experiences are related to infant disorganized attachment status: Is frightening and/or frightened parental behavior the linking mechanism? In: Greenberg, M.; Cichetti, D.; and Cummings, M., eds. *Attachment in the Preschool Years.* Chicago: University of Chicago Press, 1990.

Main, M., and Solomon, J. Discovery of an insecure-disorganized/ disoriented pattern. In: Brazelton, T.B., and Yogman, M., eds. *Affective Development in Infancy.* Norwood, NJ: Ablex Publishing, 1986.

Matas, L.; Arend, R.; and Sroufe, L.A. Continuity of adaptation in the second year: The relationship between quality of attachment and later competence. *Child Dev* 49:547-556, 1978.

Myers, B.J.; Olson, H.C.; and Kaltenbach, K. Cocaine-exposed infants: Myths and misunderstandings. *Zero to Three* 13(1):1-5, 1992.

Neuspiel, D.R., and Hamel, S.C. Cocaine and infant behavior. *J Dev Behav Pediatr* 12:55-64, 1991.

O'Connor, M.J.; Sigman, M.; and Kasari, C. Interactional model for the association among maternal alcohol use, mother-infant interaction, and infant cognitive development. *Infant Behav Dev* 16:177-192, 1993.

Parke, R.D. Interactional designs. In: Cairns, R.B., ed. *The Analysis of Social Interactions*. Hillsdale, NJ: Erlbaum, 1979.

Rodning, C.; Beckwith, L.; and Howard, J. Quality of attachment and home environments in children prenatally exposed to PCP and cocaine. *Dev Psychopathol* 3:351-366, 1991.

Roe, K.V.; McClure, S.; and Roe, A. Vocal interaction at 3-months and cognitive skills at 12 years. *Dev Psychol* 18:15-16, 1982.

Rosenbaum, M. Difficulties in taking care of business: Women addicts as mothers. *Am J Drug Alcohol Abuse* 6(4):431-446, 1979.

Sameroff, A.J., and Fiese, B.H. Transactional regulation and early intervention. In: Shonkoff, J.P., and Meisels, S.J., eds. *Handbook of Early Childhood Intervention*. Cambridge, MA: University of Cambridge Press, 1990.

Spiker, D.; Ferguson, J.; and Brooks-Gunn, J. Enhancing maternal interactive behavior and child social competence in low birth weight, premature infants. *Child Dev* 64(3):754-768, 1993.

Sroufe, L.A. Infant-caregiver attachment and patterns of adaptation in preschool: The roots of maladaption and competence. In: Perlmutter, M., ed. *The Minnesota Symposia on Child Psychology.* Vol. 16. Hillsdale, NJ: Erlbaum, 1983.

Sroufe, L.A. Attachment classification from the perspective of infant-caregiver relationships and infant temperament. *Child Dev* 56(1):1-14, 1985.

Thoman, E.; Denenberg, V.; Sievel, J.; Zeidner, L.; and Becker, P. State organization in neonates: Developmental inconsistency indicates risk for developmental dysfunction. *Neuropediatrics* 12:45-54, 1981.

Tronick, E. Emotions and emotional communication in infants. *Am Psychol* 44:112-119, 1989.

Tronick, E.Z., and Field, T., eds. *Maternal Depression and Infant Disturbance. New Directions for Child Development.* San Francisco, CA: Jossey-Bass, 1986.

Tronick, E., and Weinberg, K. "Mother-Infant Interactive Paradigms for Evaluating the Socio-Emotional Development of Infants." Paper presented at the National Institute on Drug Abuse, Washington, DC, November 8-9, 1990.

Wachs, T.D. Proximal experience and early cognitive-intellectual development: The social environment. In: Gottfried, A., ed. *Home Environment and Early Cognitive Development.* New York: Academic Press, 1984.

Wachs, T.D., and Gruen, G.E. *Early Experience and Human Development.* New York: Plenum, 1982. 299 pp.

Waters, E., and Deane, K.E. Defining and assessing individual differences in attachment relationships: Q-methodology and the organization of behavior in infancy and early childhood. *Monogr Soc Res Child Dev* 50(1-2):41-65, 1985.

Yarrow, L.J.; Rubenstein, J.L.; and Pederson, F.A. *Infant and Environment: Early Cognitive and Motivational Development.* New York: Wiley, 1975. 255 pp.

Zuckerman, B., and Frank, D.A. "Crack kids": Not broken. *Pediatrics* 89(2):337-339, 1992.

Zuckerman, B.; Frank, D.A.; Hingson, R.; Amaro, H.; Levenson, S.M.; Kayne, H.; Parker, S.; Vinci, R.; Aboagye, K.; Fried, L.E.; Cabral, H.; Timperi, R.; and Bauchner, H. Effects of maternal marijuana and cocaine use on fetal growth. *N Engl J Med* 320(12):762-768, 1989.

ACKNOWLEDGMENT

Preparation of this chapter was supported by National Institute on Drug Abuse grant R18-DA-06380.

AUTHOR

Leila Beckwith, Ph.D.
Professor Emeritus
Child Development Studies
Department of Pediatrics
University of California, Los Angeles School of Medicine
300 UCLA Medical Plaza, Suite 3300
Los Angeles, CA 90095-7033
(310) 794-1456 (Tel)
(310) 206-4215 (Fax)
lbeckwit@ucla.edu (Internet)

Measures of Service Utilization

Mary E. McCaul and Dace S. Svikis

INTRODUCTION

There has been considerable progress in the development of quantitative and qualitative measures of substance abuse treatment over the past two to three decades. Early treatment outcome research often approached treatment as a "black box." Reports typically indicated only that study subjects were enrolled in a particular service modality, such as inpatient, residential, therapeutic community, or outpatient care, and there was little or no description of the content or components of the treatment modality. This black box approach to treatment research at times was exacerbated by the failure to characterize the extent to which subjects received the intended services. Dichotomous measures often were used to characterize service delivery; that is, subjects either did or did not participate in the experimental treatment program. It was not uncommon for all the emphasis to be placed on assessment of a subject at the time of treatment entry, discharge, and posttreatment followup rather than on the nature and extent of the intervening services that were delivered. Not surprisingly, much of this earlier research found no differences between alternative types of services and suggested that subject characteristics were better predictors of treatment outcome than was treatment type.

Fortunately, measures of treatment content and delivery became more sophisticated over time. Researchers adopted categorical and continuous measures of treatment exposure. For example, some investigators examined whether the number of treatment sessions that subjects received was an important predictor of postdischarge outcome (Polich et al. 1980), and others focused on subjects' time in treatment (i.e., treatment retention) (Hubbard et al. 1989). In these studies, treatment measures generally were used as independent variables or predictors of posttreatment outcome, and there was little analysis of the effects of various subject and treatment factors on treatment retention or utilization (e.g., the percentage of scheduled treatment visits that were kept). For example, a robust positive relationship has been demonstrated between treatment retention and posttreatment abstinence rates for both alcohol- and opiate-dependent clients across a variety of treatment modalities (Hubbard et al. 1989; Polich et al. 1980; Simpson and Sells 1982).

Recently, psychotherapeutic treatment researchers have adopted the more rigorous clinical trials methodology that has become the standard in medication development research. In this research, it is critical to describe

fully the medication type, dosage regimen, and subject compliance with dosing procedures. Similarly, in psychotherapeutic treatment research, it is critical to systematically characterize type, delivery regimen, and subject utilization of treatment services. At a minimum, investigators need to describe fully the theoretical orientation (e.g., cognitive-behavioral, psychoanalytic, rational-emotive) and therapeutic content (e.g., anger management, communication skills, drug refusal skills) of planned interventions. It is equally important to characterize both intended and actual service utilization patterns, including frequency of subject participation, session duration, and length of retention in treatment.

This chapter examines types of utilization measures employed across the range of treatment modalities studied by the National Institute on Drug Abuse (NIDA) Perinatal-20 projects. The chapter discusses the utility of these measures for better understanding treatment effects on pregnant substance-abusing women while considering treatment utilization measures as both dependent and independent variables. Such measures offer the potential for increased understanding of the complex interactions of subject, therapist, theoretical orientation, therapeutic content, and service utilization variables in treatment outcome research with drug-addicted women.

UTILIZATION MEASURES FOR DIFFERENT TREATMENT COMPONENTS

Intake and Assessment

Many potential subjects are lost to treatment between the initial program contact and the startup of services. Research suggests that as many as one-third of persons seeking treatment fail to complete the steps to successful program admission (Olkin and Lemle 1984; Wanberg and Jones 1973). Admission rates are strongly influenced by several program characteristics. On the one hand, programs may establish intake criteria that significantly restrict the volume or prognostic characteristics of subjects entering treatment. For example, some Perinatal-20 projects targeted women early in their pregnancies, whereas others admitted women with children without regard to their pregnancy status. On the other hand, some programs specifically introduced interventions to improve capture rate into care, including telephone calls and reminder letters prior to scheduled appointments or following missed appointments. It is important to accurately measure and describe program characteristics that influence the rate of treatment entry because they may ultimately influence rates of retention and program completion as well.

Of particular concern in treatment outcome research is the effect of the randomization process on subjects' willingness to participate in treatment and thus in the research project. Subjects often present for study enrollment with strong opinions as to the type of treatment services that they need or want. When randomization results in assignment to a different treatment condition, some women may drop out even before starting the assigned services. For example, one Perinatal-20 project that examined the effectiveness of residential vs. outpatient treatment reported considerably elevated early dropout rates for women randomized into the outpatient condition. It is important to obtain information on subjects' reasons for early treatment termination whenever possible to examine the contribution of the random assignment process to treatment outcomes.

Several measures are readily available to most studies to characterize their success in engaging subjects in treatment. Often, this measurement is accomplished somewhat after the fact by examining overall program utilization rates based on established static capacity. That is, programs track the percentage of subjects who were active in treatment within the reporting period as a function of an established program capacity. For outpatient programs, static capacity typically reflects staffing levels and the recommended client-to-staff ratio based on the intensity of program services. For an inpatient or residential facility, static capacity is simply a reflection of the number of beds available within the program. Timeframes for measuring static capacity can range from daily for bed utilization, to monthly for counselor caseloads, to annually for overall program utilization levels. It is important to consider the impact of static capacity on the total number of subjects who can be enrolled in the project and, therefore, on the statistical power available to test study hypotheses. For example, power was often of particular concern for the Perinatal-20 projects that studied residential treatment services in which limited numbers of beds or relatively long stays considerably restricted the number of study subjects. Although utilization rate provides a gross measure of subject flow into and out of treatment, it fails to clarify the potential loss of subjects prior to formal admission into the treatment program, because these subjects are not counted in such a statistic.

The simplest strategy for measuring admission rate into treatment may be to track the total number of subjects who are formally admitted into the program as a function of the total number of women who request admission. This can be done on the basis of telephone calls, walk-in visits, scheduled appointments, or kept appointments. Such a continuum of measures provides information on a narrowing pool of subjects. For example, the volume of telephone calls inquiring about program availability may reflect an overall level of demand in the geographic area served by the program. In contrast, the ratio of scheduled-to-kept

appointments may be a more specific indicator of program success at capturing into treatment those women who were evaluated as appropriate for program admission during the initial program contact. Such study loss prior to the treatment intervention is illustrated by one of the Perinatal-20 programs that focused on outpatient couples therapy for drug-abusing women. Of the pool of 113 women who consented to project enrollment, 76 percent completed the initial intake interviews, 43 percent completed the couples pretests, but only 34 percent entered a treatment condition.

It is well established that the longer a person has to wait from initial program contact to treatment entry, the more likely it is that the individual will drop out prior to receiving care (Horgan et al. 1991, pp. 123-144). Thus, it may be informative to monitor the number of days or weeks that elapse from the time that a subject first telephones or visits the program to seek admission to the time that she completes formal admission requirements into the program. It also is important to characterize the number of separate visits and different staff contacts that are required to complete the admission process. If there is an additional delay from time of assessment completion to therapeutic service delivery, this should be measured. Finally, selected subject characteristics that can be expected to affect a woman's likelihood of completing the admission process should be monitored. Such factors include referral source (e.g., child protective services, criminal justice agency, social services agency, family member), payment status (e.g., medicaid or medicare, private insurance with deductibles and copayments, self-pay), work or child care arrangements, and readiness for change (Prochaska et al. 1992). It is certainly reasonable to expect that the more demands placed on subjects prior to their receipt of therapeutic services, the higher the rates of treatment dropout.

Traditional Outpatient Treatment

Outpatient, abstinence-oriented drug treatment is the most common and least expensive service modality available in this country (Butynski 1991, pp. 20-52; National Institute on Drug Abuse and National Institute on Alcohol Abuse and Alcoholism 1993). It creates relatively little disruption in the daily activities of the subject (e.g., work, family obligations) and also permits the individual to directly apply and practice what she has learned in treatment. Unfortunately, in this regard, outpatient treatment also increases the risk for relapse because alcohol and other drugs are freely available to subjects who return to their communities following each treatment contact.

For methadone maintenance programs, Federal regulations define minimally acceptable levels of pharmacotherapy and counseling services. However, there are no similar national service standards that define a drug-free outpatient substance abuse treatment program. Although most

programs offer some combination of individual and group counseling sessions, theoretical orientation, therapeutic content, and quality of services can vary considerably across treatment sites (Price et al. 1991, pp. 63-92). Sessions can include didactic material (e.g., drug education, relapse prevention strategies, skills building) as well as more traditional psychotherapy.

Intended length of stay and quantity and frequency of counseling sessions also vary across treatment sites. Planned treatment retention can vary from a relatively brief, fixed-length intervention to a more long-term or even open-ended stay. Some programs require participation in only 1 hour per week of individual or group therapy, whereas others require multiple visits for 3 or more hours of counseling per week. For example, treatment services in one Perinatal-20 project consisted of a 16-week, one-session-per-week skills training program. In several others, treatment duration was open-ended and was determined by subjects' achievement of identified treatment goals, such as abstinence, residential stability, and parenting skill development.

Program requirements also may change as a woman progresses through treatment. Early recovery may require more intensive participation, followed by a gradual decrease in quantity and frequency of attendance as subjects achieve longer periods of abstinence and other therapeutic goals. In the authors' Perinatal-20 project, women graduated from intensive-day to partial-day to standard outpatient care as a function of treatment goal achievement, number of drug-free urinalyses, and regularity of attendance.

In considering treatment utilization rates, it is important to specify the length and intensity of prescribed treatment so that a meaningful rate of dropout can be reported and compared with other programs with shorter or longer prescribed terms of treatment. A variety of measures can be used to characterize attendance and participation in standard outpatient treatment settings. The number of individual and group counseling sessions can be counted. Furthermore, to adjust for length of time in treatment, mean sessions per month or the percentage of scheduled appointments kept by the subject also should be examined. Treatment retention or length of stay (typically defined as the time from first to last face-to-face contact with the program) can be measured easily. In addition, if individual or group counseling sessions vary in duration, total treatment hours may provide useful information about the quantity of treatment received by the subject.

More refined analyses permit the study of trends or patterns in subject attendance over time. Rather than simply looking at total days in attendance, it is important to consider the distribution of visit days over

the length of stay in treatment. For example, two subjects who participate in an equal number of individual counseling sessions may have dramatically different attendance patterns and outcomes. One subject attends a particular program as prescribed several times a week for the first month or two of treatment with no evidence of alcohol or other drug use but then relapses and abruptly discontinues treatment with no further program contact. In contrast, another subject continues to use drugs intermittently early in treatment and, as a result, misses many of the required clinic visits; however, she eventually stops her illicit drug use and stabilizes her attendance. Although both subjects may have attended an equal number of sessions, their attendance patterns are different and are associated with different treatment outcomes. Such participation patterns provide valuable data for treatment planning and assist in matching individuals to the appropriate therapeutic interventions. To date, patterns of attendance have not been widely studied as a predictor of discharge status or posttreatment outcomes.

In addition to monitoring quantity, frequency, and pattern of attendance, more qualitative analyses of service delivery ideally should be used. At the grossest level, types of services (e.g., individual counseling, group therapy, educational groups) can be examined. Additional measures can include staff ratings of subject participation in sessions (e.g., five-point scale ratings of subjects' alertness, talkativeness, self-disclosure). In recent years, there has been increased emphasis on subjects' estimates of the number of services they received in the designated period and their ratings of the usefulness of treatment services received. The Treatment Services Review (McLellan et al. 1992) is being widely adopted in this area (see the section titled "Program-Level Utilization Data" for a more detailed description). Finally, subject and staff ratings or presession and postsession evaluations can assist in measuring how well a participant understood the presented material or appeared to benefit from the session. These qualitative measures can enhance understanding and interpretation of quantitative utilization measures. For example, subjects are not likely to continue to attend services that they perceive as irrelevant or poorly delivered.

Intensive Outpatient Treatment

In the past 10 years, intensive outpatient programs have become more popular for the treatment of drug addiction (Center for Substance Abuse Treatment 1994). Generally, such programs consist of 3 or more hours per day of individual and group counseling, with health care, supportive services, and meals available at the program. Subjects typically remain at this intensive level of care for several weeks to several months, depending

on demographic (e.g., residential stability) and drug use (e.g., prior treatment history) characteristics and treatment progress. Several Perinatal-20 projects examined the effectiveness of intensive outpatient programs delivered in a "one-stop shopping" model of care compared with that of standard outpatient services.

Many of these standard outpatient treatment measures also are relevant for characterizing intensive service programs. However, in addition to recording the number of days of attendance, intensive outpatient programs may find it important to report the subject's attendance behaviors within the treatment day. Subjects may arrive late, leave early, or miss sessions for excused and unexcused reasons. Treatment curricula are often cumulative in design, with information in one group session building on or applying information from a preceding session. Thus, absenteeism can significantly affect the overall quantity and quality of a subject's treatment experience. Furthermore, if a subject is showing consistent patterns of missed groups (e.g., first or last daily group), this may alert the program clinicians to external conflicts that may be interfering with her ability to attend the program. It is essential that rates and patterns of attendance be recorded so that causes for poor attendance (i.e., underutilization) can be determined. For example, in the authors' Perinatal-20 project, approximately one-quarter of the patients had chronic problems with late arrival to the daily treatment program; however, this stemmed from such diverse causes as a day care center that opened at the same time as the treatment program, ongoing prostitution as the only source of financial support, or simply the inability to organize morning routines in the household. Individualized functional analyses of this common utilization problem (i.e., lateness) led to different remedial goals for each woman.

At a minimum, daily treatment documentation should include the number and type of standard counseling services received, including individual counseling, group education, and group therapy. In addition, as described above, staff ratings of subject attention and participation can be obtained at each therapeutic session during the day. Development of a within-treatment-day profile of subject activity also is useful for quantifying the nonroutine services that subjects receive only from time to time. Referrals to social services, vocational services, onsite or offsite medical and psychiatric services, or other ancillary appointments often are recorded only as "no shows" on routine group treatment documentation. If a detailed log of daily subject activity is maintained, then a wider range of information can be captured and quantified in a single source document. This same type of documentation system can be used successfully in residential or therapeutic community settings.

Obstetric and Pediatric Medical Services

The service components described above apply to a variety of subjects in diverse treatment settings. However, when the target population is pregnant drug-abusing women, additional services (e.g., prenatal care) are needed and should be carefully measured. As a component of the research design, such ancillary services may be provided onsite at a single comprehensive treatment program or offsite with coordination of services among various treatment sites. In either case, it is critical to monitor and evaluate subject utilization of such services in conjunction with other treatment participation and retention variables.

For obstetric care, the primary units of service include number of prenatal care visits and number and type of fetal monitoring sessions (e.g., sonograms, nonstress tests). Unfortunately, such measures are influenced by the estimated gestational age of the subject's fetus at the time of treatment entry. Thus, a woman who initiates treatment in the first trimester of pregnancy will have more opportunities for prenatal care visits than a woman who initiates treatment in the third trimester of pregnancy. To minimize such bias and more effectively examine subject compliance with prescribed obstetric care, the ratio of kept vs. scheduled obstetric appointments may provide a sensitive measure of service utilization.

In addition to these measures of routine obstetric care, other pregnancy-related measures of medical service utilization may include the number of emergency room visits, number of hospitalizations and length of stay, and number of specialty consultations (e.g., cardiology).

At the time of delivery, new measures are introduced. For the mother, important variables include length of hospitalization, number of postnatal medical followup visits, and utilization of family planning services. For the infant, a variety of pediatric measures can be monitored and evaluated. Specifically, an infant's length of stay in the hospital is an important variable. Examination of the need for and length of stay of infants in the neonatal intensive care unit compared with the standard care nursery also provides valuable information. Following the infant's discharge from the hospital, utilization measures include kept vs. scheduled well-baby pediatric visits, adherence to the recommended immunization schedule, number and pattern of sick visits, infant need for and utilization of specialty services, and frequency and duration of hospitalizations.

Other Ancillary Services

It has been well established that individuals with substance use disorders present with myriad other problems, including psychiatric comorbidity

(e.g., depression), medical problems (e.g., hypertension, reproductive dysfunction), legal issues (e.g., drug dealing, prostitution), and vocational deficits (e.g., poor work history, lack of marketable skills) (James et al. 1991; Lex 1991; Marsh and Miller 1985; McCaul et al. 1991). Subjects often require immediate intervention and assistance in one or more of these life areas.

Most drug treatment programs and clinical research initiatives assess subjects at intake to determine their specific needs and deficits. Such assessments should be completed using a structured intake interview, such as the Addiction Severity Index (ASI) (McLellan et al. 1980). Subjects with a variety of psychosocial or medical problems may find a substantial amount of their time devoted to resolution of these concomitant problems, with less time available for traditional drug treatment services. It is important to monitor the frequency of and compliance with ancillary, offsite appointments and to distinguish between excused (e.g., subject missed a group session because of a medical appointment) and unexcused (e.g., subject was absent from the group because she overslept) absences from drug treatment.

Accurate assessment of subject utilization and followthrough with ancillary services may be a formidable undertaking, requiring substantial staff time and effort. Typically, appointments are scheduled offsite, and intensive followup is required to confirm that subjects kept their scheduled appointments. One strategy is to monitor the most common service referral, particularly if many subjects are referred to a single site. For example, if women with psychiatric issues generally are referred to a nearby mental health clinic for evaluation and counseling, the drug treatment program may elect to establish a formal liaison with this facility. Such collaborative arrangements typically facilitate monitoring of followthrough with prescribed care. Formal mechanisms for data exchange among care sites also can be established, thereby routinizing information provision by program staff. For example, a monthly report of client services could be generated and eliminate the need for followup on individual cases. An alternative strategy is to require subjects to bring written confirmation of a kept appointment; such confirmation could include a brief note on provider stationery, discharge or followup instructions generated by the provider, or even the bill for services.

Primary measures of ancillary care utilization are the number of appointments and the percentage of scheduled appointments kept by a subject. Depending on the number and variety of service referrals, ancillary care utilization can be examined as a single variable summed across all domains of care (e.g., medical + social services + vocational), or it can be examined separately for each domain. The total time spent

in ancillary care also may be examined, although this requires more precise monitoring of subject participation in offsite services and introduces greater variability into the data.

Applications of Utilization Data

Treatment programs already spend a considerable amount of effort reporting on a variety of program and client characteristics and services to various monitoring agencies. Unfortunately, such measures are often aggregated in State and national reports that do not provide sufficient detail to permit meaningful program self-evaluation. However, these data can be approached from several different perspectives that give increasingly more program- and client-specific information for use in evaluating program operation. These perspectives include program-, counselor-, and client-level data. Given the substantial paperwork burden already existing in most addiction treatment programs, the primary goal should be to use existing data in more sophisticated ways rather than to generate entirely new data sets.

Federal and State Reporting Systems

Program utilization data are already routinely collected for a variety of Federal and State agencies. The majority of these surveys include descriptive information of program services, sociodemographic information on program clients, and measures of service utilization. For example, the National Drug and Alcoholism Treatment Unit Survey (NDATUS) (now known as the Uniform Facility Data Set), funded by NIDA and the National Institute on Alcohol Abuse and Alcoholism, obtains information every 2 years from public and private alcohol and other drug treatment programs in the United States. NDATUS includes questions on types of care provided, client count and capacity, client demographic characteristics, specialized programs, funding amounts and sources, staffing, and waiting lists. It obtains point-prevalence information on treatment utilization (program service data for active clients in treatment on a single predetermined calendar day each year) and annualized information on service delivery to target client populations, such as pregnant women and HIV-positive clients.

The National Treatment Improvement Evaluation Survey (NTIES) (Center for Substance Abuse Treatment 1993), funded by the Center for Substance Abuse Treatment, is more intensive and includes two utilization components conducted in treatment programs selected to be representative of the national treatment system. First, NTIES conducted a detailed examination of the scope of program services, and then programs

have been required to report utilization data on a monthly basis over a multiyear period.

Each State is required to implement a management information system to collect ongoing client characteristic and service delivery data. For example, the Maryland Substance Abuse Management Information System obtains monthly reports on each client admitted to and discharged from treatment during the reporting period. These client-level data include demographics, substance use characteristics, and psychosocial functioning at treatment entry and discharge. In addition, programs are required to report admission and discharge dates; units of service for individual, group, and family counseling sessions; and number of urinalyses conducted. This system also compiles program-level data on waiting list volume as a function of treatment modality. In recent years, some State alcohol and other drug agencies also have developed management information systems specifically to examine the number of pregnant women and pregnancy outcomes of drug-dependent women in the treatment system. Such systems can provide basic feedback to individual programs on aspects of operation, such as client retention, mean number of sessions per client, and changes in client psychosocial functioning from pretreatment to posttreatment. This information allows programs to assess compliance with internally established program standards of care as well as track outcome trends over time. Thus, most programs have access to basic information on program utilization as a function of their mandated participation in these Federal and State reporting systems.

Program-Level Utilization Data

These required program monitoring systems can serve as the foundation on which more sophisticated program operation and client care monitoring systems can be built. Total quality management programs and continuous quality improvement monitoring have been useful for focusing attention on the need for more sophisticated data collection strategies to identify specific opportunities for program enhancement. The variety of quantitative measures described above can be used to enrich the basic mandated service delivery measures. In addition, measures can be combined or examined in novel ways to increase the meaningfulness of the resultant information. For example, as described above, rather than a simple count of the number of individual counseling sessions as required by the State reporting system, analysis can be made of patterns of session attendance and the percentage of scheduled sessions that were kept by the client. Although requiring little additional effort, such strategies can yield much more meaningful information on client participation and retention in treatment.

Recently, standardized tools have been developed to capture client feedback on perceived quantity and utility of program services. For example, the Treatment Services Reveiw (McLellan et al. 1992) is a brief, semistructured interview that elicits service information in the seven domains examined in the ASI (McLellan et al. 1980) and is included in most comprehensive individualized treatment plans. These domains are medical, employment/ vocational, alcohol, drug, legal, family/social, and psychological status. Within each domain, clients report on quantity of services received since the last interview and then rate perceived helpfulness of these services. This combined quantitative and qualitative approach can provide programs with expanded feedback on service delivery levels and client responses to these services. Client ratings of service utility can provide important information for program planning of treatment expansion or in meeting the needs of special populations. For example, pregnant women might be expected to rate onsite medical care as more useful than would the population of young men traditionally served by drug treatment programs.

Other aspects of program operation that already are monitored routinely and may be readily available to incorporate into utilization analyses include program billing records, pharmacy order records, medication dispensing logs, urinalysis logs, and referral records.

Counselor-Level Utilization Data

Several studies have demonstrated that, within a given treatment program, there can be considerable variability among addiction counselors in their ability to retain clients in treatment (Rosenburg et al. 1976; Valle 1981). McLellan and colleagues (1988) also have reported variability across counselors for within-treatment client measures such as the amount of the daily methadone dose, employment status, and utilization of supportive medical services in a methadone maintenance setting. Within the authors' women's treatment clinic, there are many differences among counselors in maintaining subject participation as measured by treatment duration and discharge status. Over a 1-year period (July 1992 through June 1993), the percentage of subjects who complied with program attendance standards ranged from 40 to 100 percent across counselor caseloads. Similarly, the percentage of subjects in compliance with clinic morning arrival times ranged from 58 to 100 percent across counselor caseloads. Such divergence may reflect several factors operating at both the counselor and program levels. Counselor variables may include education and experience, theoretical orientation, and therapeutic style. Program factors may include targeted assignment of more challenging, specialized caseloads (e.g., dually diagnosed clients) and caseload volumes. Aggregating service utilization data for each counselor can begin to provide information on the

potential effect of these counselor and program variables and may point to important areas for individual training or program improvement.

Subject-Level Utilization Data

The following data will be used to illustrate how the level of analysis can dramatically affect outcome findings. The data were abstracted from subject records of an intensive outpatient drug treatment program that specializes in the care of women. The program was one of the sites participating in the NIDA Perinatal-20 Treatment Research Demonstration Program.

Subject A was a 29-year-old Caucasian female who was referred for her first treatment episode by a community health center. The subject was unemployed, lived with her significant other and their children, and reported no significant psychiatric problems. Subject A denied any arrests in the 2 years prior to treatment enrollment. Substance use diagnoses included alcohol, heroin, and cocaine dependence. The subject remained in treatment for approximately 4 1/2 months and was discharged prior to completing the first and most intensive phase of treatment because of lack of attendance.

Subject B was a 30-year-old African-American female who was referred for her third treatment episode by the attending physician who had treated her during a recent hospitalization. She was unemployed and living with her children. She reported a prior history of psychiatric treatment and admitted to three arrests in the 2 years prior to treatment enrollment. Substance use diagnoses included cocaine and alcohol dependence. This subject remained in the first phase of treatment for approximately 5 1/2 months and then was transferred to less intensive treatment programming.

Treatment utilization data for these two subjects are summarized in table 1. If treatment utilization is examined at the macrolevel (e.g., length of stay), the two women appear similar, with Subject B remaining in treatment only 1 month longer than Subject A. However, if treatment utilization is examined at the microlevel in conjunction with other dependent variables, a somewhat different picture emerges. For example, status at discharge shows that Subject A terminated treatment unsuccessfully, whereas Subject B transferred successfully to the second and less intensive phase of treatment.

Additional information is obtained by a more detailed examination of treatment attendance data. Subject A had 58 days with face-to-face treatment contact, whereas Subject B had 87 days of program contact.

TABLE 1. *Treatment utilization data for two representative subjects enrolled in Perinatal-20 treatment services*

Measure	Subject A	Subject B
Length of stay (months)	4.5	5.5
Number of treatment days (face-to-face contact)	58	87
Number of days with ontime clinic arrival	25	48
Number of individual counseling sessions	31	25
Number of group counseling sessions	270	98
Percent positive drug toxicologies	75	48
Discharge status	Dropout	Service transfer

Subject B also was somewhat more likely to arrive on time at the treatment program (55 percent ontime attendance) compared with Subject A (43 percent ontime attendance).

When treatment contact (i.e., number of individual and group counseling sessions attended) is examined, a different pattern emerges. Specifically, despite fewer days in attendance, Subject A participated in more individual counseling sessions than Subject B. For group counseling, the difference is more dramatic, with Subject A attending nearly three times as many groups as Subject B.

These data may seem somewhat contradictory. Subject A received more treatment than Subject B, yet Subject A was unsuccessfully discharged from the program and Subject B was transferred to less intensive treatment programming after successfully completing the first phase of treatment. Nevertheless, they typify the diversity of outcomes that are seen in substance abuse treatment programs. The example also illustrates the difficulty in selecting the most appropriate measure to use in evaluating treatment outcomes because different measures can suggest different results.

CONCLUSION

Clearly, there has been substantial progress in the number, scope, and quality of substance abuse treatment outcome evaluations over the past several decades. Treatment researchers are developing highly sophisticated scales to describe and measure the components, content, and integrity of therapeutic interventions. It is likewise critical that researchers develop more sophisticated strategies for measuring treatment participation and retention. As this chapter illustrates, utilization measures can be inexpensive to collect and often are readily available within mandated treatment reporting systems. These measures can be examined

differently or supplemented in small but meaningful ways to yield far more precise and informative measures of program effectiveness. Such detailed measurement of service utilization will become increasingly important with the growing emphasis on cost-effectiveness of care.

It also is important to remember the utility of these measures as both independent and dependent variables. For example, numerous earlier studies demonstrated a positive relationship between the number of treatment visits and postdischarge treatment outcome. Because few treatment programs are able to afford extensive posttreatment followup of program clients, the number of treatment visits and length of retention can become an inexpensive surrogate dependent measure for determining the effects of a change in program service delivery on client outcomes. Clients often communicate important information on the quality and utility of treatment services through the extent of their attendance and participation in these services. In the face of increasing demands from regulatory agencies, insurance payers, and treatment clients, it will become even more important to have well-established, routine strategies for utilization monitoring at the program, counselor, and client levels.

REFERENCES

Butynski, W. Drug treatment services: Funding and admissions. In: Pickens, R.W.; Leukefeld, C.G.; and Schuster, C.R., eds. *Improving Drug Abuse Treatment.* National Institute on Drug Abuse Research Monograph 106. DHHS Pub. No. (ADM)91-1754. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Center for Substance Abuse Treatment. *National Treatment Improvement Evaluation Study. Project Summary*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment, 1993. 8 pp.

Center for Substance Abuse Treatment. *Intensive Outpatient Treatment for Alcohol and Other Drug Abuse*. Treatment Improvement Protocol (TIP) Series, No. 8. DHHS Pub. No. (SMA)94-2077. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment, 1994. 104 pp.

Horgan, C.; Rosenback, M.; Ostby, E.; and Butrica, B. Targeting populations with drug abuse problems: Pregnant women. In: Kaple, J.M., ed. *NIDA Drug Abuse Services Research Series: Background Papers on Drug Abuse Financing and Services Research. Volume 1*. DHHS Pub. No. (ADM) 91-1777. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Hubbard, R.L.; Marsden, M.E.; Rachal, J.V.; Harwood, H.J.; Cavanaugh, E.R.; and Ginzburg, H.M. *Drug Abuse Treatment: A National Study of Effectiveness*. Chapel Hill, NC: University of North Carolina Press, 1989. 213 pp.

James, M.E.; Rubin, C.P.; and Willis, S.E. Drug abuse and psychiatric findings in HIV-seropositive pregnant patients. *Gen Hosp Psychiatry* 13:4-8, 1991.

Lex, B. Some gender differences in alcohol and polysubstance users. *Health Psychol* 10:121-132, 1991.

Marsh, J.C., and Miller, N.A. Female clients in substance abuse treatment. *Int J Addict* 20:995-1019, 1985.

McCaul, M.E.; Svikis, D.S.; and Feng, T. Pregnancy and addiction: Outcomes and interventions. *Md Med J* 40:995-1001, 1991.

McLellan, A.T.; Alterman, A.I.; Cacciola, J.; Metzger, D.; and O'Brien, C.P. A new measure of substance abuse treatment: Initial studies of the Treatment Services Review. *J Nerv Ment Dis* 180(2):101-110, 1992.

McLellan, A.T.; Luborsky, L.; Woody, G.E.; and O'Brien, C.P. An improved diagnostic evaluation instrument for substance abuse patients: The Addiction Severity Index. *J Nerv Ment Dis* 168(1):26-33, 1980.

McLellan, A.T.; Woody, G.E.; Luborsky, L.; and Goehl, L. Is the counselor an "active ingredient" in substance abuse rehabilitation? An examination of treatment success among four counselors. *J Nerv Ment Dis* 176:423-430, 1988.

National Institute on Drug Abuse and National Institute on Alcohol Abuse and Alcoholism. *National Drug and Alcoholism Treatment Unit Survey (NDATUS) 1991: Main Findings Report.* DHHS Pub. No. (SMA)93-2007. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Substance Abuse and Mental Health Services Administration, 1993. 99 pp.

Olkin, R., and Lemle, R. Increasing attendance in an outpatient alcoholism clinic: A comparison of two intake procedures. *J Stud Alcohol* 45:465-468, 1984.

Polich, J.M.; Armor, D.J.; and Braiker, H.B. *The Course of Alcoholism: Four Years After Treatment.* Santa Monica, CA: The Rand Corporation, 1980. 361 pp.

Price, R.H.; Burke, A.C.; D'Aunno, T.A.; Klingel, D.M.; McCaughrin, W.C.; Rafferty, J.A.; and Vaughn, T.E. Outpatient drug abuse treatment services, 1988: Results of a national survey. In: Pickens, R.W.; Leukefeld, C.G.; and Schuster, C.R., eds. *Improving Drug Abuse Treatment.* National Institute on Drug Abuse Research Monograph 106. DHHS Pub. No. (ADM)91-1754. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Prochaska, J.O.; DiClemente, C.C.; and Norcross, J.C. In search of how people change: Applications to addictive behaviors. *Am Psychol* 47:1102-1114, 1992.

Rosenburg, C.N.; Gerrein, J.R.; Mandhar, V.; and Leetick, J. Evaluation of training of alcoholism counselors. *J Stud Alcohol* 37:1236-1246, 1976.

Simpson, D.D., and Sells, S.B. Effectiveness of treatment for drug abuse: An overview of the DARP Research Program. *Adv Alcohol Subst Abuse* 2:7-29, 1982.

Valle, S.K. Interpersonal functioning of alcoholism counselors and treatment outcome. *J Stud Alcohol* 42:783-790, 1981.

Wanberg, K.W., and Jones, E. Initial contact and admission of persons requesting treatment for alcohol problems. *Br J Addict* 68:281-285, 1973.

AUTHORS

Mary E. McCaul, Ph.D.
Associate Professor
Johns Hopkins University School of Medicine
Director
Comprehensive Women's Center
911 North Broadway
Baltimore, MD 21205
(410) 955-5439 (Tel)
(410) 955-4769 (Fax)
betsymac@welchlink.welch.jhu.edu (Internet)

Dace S. Svikis, Ph.D.
Assistant Professor
Johns Hopkins University School of Medicine
Director
Center for Addiction and Pregnancy
Johns Hopkins Bayview Medical Center
4940 Eastern Avenue
Baltimore, MD 21224
(410) 550-3023 (Tel)
(410) 550-3027 (Fax)

A Database Model for Studies of Cocaine-Dependent Pregnant Women and Their Families

Peter A. Charpentier and Richard S. Schottenfeld

INTRODUCTION

Collecting, organizing, and analyzing data for clinical research projects that involve families of cocaine-dependent mothers present significant data management challenges. Depending on the specific administrative task or analysis, a "record" or unit of analysis might be a mother, a pregnancy, or a child. Data must be organized to permit such multiple views while preserving the family structure, and controlling access to personal identifying information is a higher priority for these data than for other clinical research database systems. This chapter describes the authors' efforts to design and implement a database system for the Mothers Project, a Perinatal-20 project established in New Haven, Connecticut. A brief description of the Mothers Project follows.

THE MOTHERS PROJECT

The Mothers Project is both a clinical epidemiological study of the correlates of cocaine and other drug abuse during pregnancy and a controlled clinical trial in which cocaine-dependent women are randomized into one of two treatment programs: an enhanced primary care treatment program or a comprehensive day treatment program that provides family support services and child care.

Routine drug abuse screening in a hospital-based prenatal clinic uses structured interviews and urine testing to identify cocaine users and enroll them in the clinical trial. Assessments of the mother and child are made at 3, 6, 9, 12, 18, and 24 months. It is permissible for women to be enrolled twice for two pregnancies.

SPECIFICATIONS OF THE DATABASE SYSTEM

The database system described in this chapter addresses the data management needs of the clinical trial component of the Mothers Project.

The design of a database system can be described in terms of *functional* and *performance* specifications (Stevens 1987, pp. 35-36); researchers interested in the fundamentals of relational database theory should see Date (1990, pp. 3-24). Functional specifications state what a database system must do, whereas performance specifications describe how often and how fast specific functions must be carried out, as well as such "real world" issues as constraints on the computing environment and the skill levels of the system operators. The performance and functional specifications of the Mothers Project database are shown in the following lists.

Performance Specifications

1. All demographic, identifying, and tracking data will be available for inspection or modification on an as-needed, per-record basis.

2. The number of data elements associated with demographic, identifying, and tracking data will be fairly small (fewer than 50 fields for a given mother, pregnancy, or child).

3. Demographic, identifying, and tracking data will be continually updated.

4. All subject tracking and contact scheduling functions will be performed by operators who have good computer skills but no programming training.

5. Interview data will be processed in "batches"; interactive record retrieval systems are not necessary.

6. The volume of the interview data, in terms of record length, will be large (e.g., several hundred variables per record).

7. The system must run on inexpensive desktop computers in either a stand-alone or networked, multiuser setting.

8. The Mothers Project staff will include data analysts and managers who are proficient in a statistical package with an excellent flat-file data management programming language (e.g., SAS) but will not need to include staff members who are proficient in relational database methods and software.

9. Although much of the database system will be developed by consultants, all software systems will be *maintainable* by project staff members who can make minor changes or enhancements to the database system.

Functional Specifications

1. Provide for the entry, storage, and modification of basic demographic, clinical, and personal identifying data for mothers, pregnancies, and children.

2. Provide a high level of security for subject identifiers such as names, addresses, and hospital unit numbers. Make it possible for all approved staff members to access the database and encrypt certain sensitive data to all but a few selected operators.

3. Track the current participation status of each subject.

4. Maintain the contact schedule of each subject.

5. Generate lists, cover sheets, and mail/merge data files for contacts falling within a specified date range. These lists assist field staff members in arranging for interviews; cover sheets are used by interviewers as "snapshots" of subjects' current status and addresses; and mail/merge data files are used for labels and followup letters.

6. Track basic information, such as dates and outcomes, for all contacts with subjects.

7. Keep track of and report missed contacts with subjects to arrange for special retrospective interviews that attempt to recover selected data.

8. Provide a framework or basis for additional planned and unplanned subsystems, such as collecting health care utilization information.

9. Provide "read only" access for ad hoc queries and for user-designed reports.

10. Provide for the entry and storage of data from baseline and followup interview data collection forms.

11. Check interview data for errors and allow operators to make appropriate changes to the data files while maintaining a complete audit of all changes.

12. Prepare files for analysis using a standard statistical package.

13. Facilitate data analysis by providing detailed data documentation.

Functional specifications fall into two major groups that might be called the *administrative* and *analytic* function groups. In practical terms,

different database and statistical software and file handling procedures, as well as staff skills, were optimally suited to each function group. Therefore, separate administrative and analytic database systems were developed to meet the functional specifications of the Mothers Project. These two database systems are described below.

Administrative Database System. The administrative database system was associated with the necessity to monitor and contact subjects. Administrative services provided by the Mothers Project database system included rapid access to records (e.g., mothers, pregnancies, children) to ascertain or change current status, addresses, and other information; periodic conduct of study reports (e.g., those concerning screening, enrollment, and participation rates); contact scheduling; and mail/merge applications for contact and followup letters.

Key requirements of the administrative database system were that data be retrievable on an individual record basis and through a variety of views. Furthermore, as the performance specifications below indicate, the administrative database system was operated by field staff members who were competent in the use of computers but were not computer programmers. As in most clinical research projects, there were not inexhaustible resources for equipment or for exotic database implementations. Thus, these factors led to considering a commercial relational database package designed for inexpensive office computers. Any of several modern, business-oriented database products would have been suitable for the Mothers Project administrative database system. A quasi-relational database package was chosen that featured sophisticated application development tools and a reasonably "user-friendly" interface for ad hoc queries and reports.

Structure of the Administrative Database. The administrative data for the Mothers Project were stored in three data sets, or tables, called the *mother*, *pregnancy*, and *child* tables. The mother table includes mother and family descriptors, such as mother's name and address and a contact person's name and address. The pregnancy table includes not only information relevant to each pregnancy (e.g., date, outcome) but also enrollment, treatment assignment, and selected followup information. Because enrollments were based on pregnancies rather than mothers, it was possible for the same mother to be included twice in the enrollment list, but for different pregnancies. The child table contains descriptors for each child, including name, birth statistics, and vital status.

The three tables that make up the foundation of the administrative database were designed to store data in an efficient and "normalized" manner. Each piece of information about a mother, pregnancy, or child is stored in only

one place (i.e., in a record within the appropriate table) so that updating information is straightforward. This structure can accommodate any number of pregnancies per mother, or children per pregnancy, without the loss of data storage efficiency. To illustrate, consider a family consisting of a mother and three children, two of whom are twins. The record structure for such a family (one mother record, two pregnancy records, and three child records) is diagramed in figure 1. Mother records are identified uniquely by a *mother ID code*, a number assigned to each new mother who is enrolled. Pregnancy records are identified uniquely by the mother ID code and a *pregnancy number*. Finally, child records are identified by the mother ID code, the pregnancy number, and a *child number*. The codes used to identify records are in a sense arbitrary and are used only for that purpose. Assembling a family's records, or those associated with a pregnancy or child, is a matter of arranging the "relationships" among the tables to construct the proper view or "relation."

Views Into the Data. Three main "views" arose from the administrative database structure: the mother, or family, view; the pregnancy, or enrollment, view; and the child view. These views are diagramed in figure 2 using the same family structure illustrated in figure 1. For most applications, the pregnancy view is used because it was the basis for enrollment. However, ancillary studies that focus on mothers or children make use of the other views.

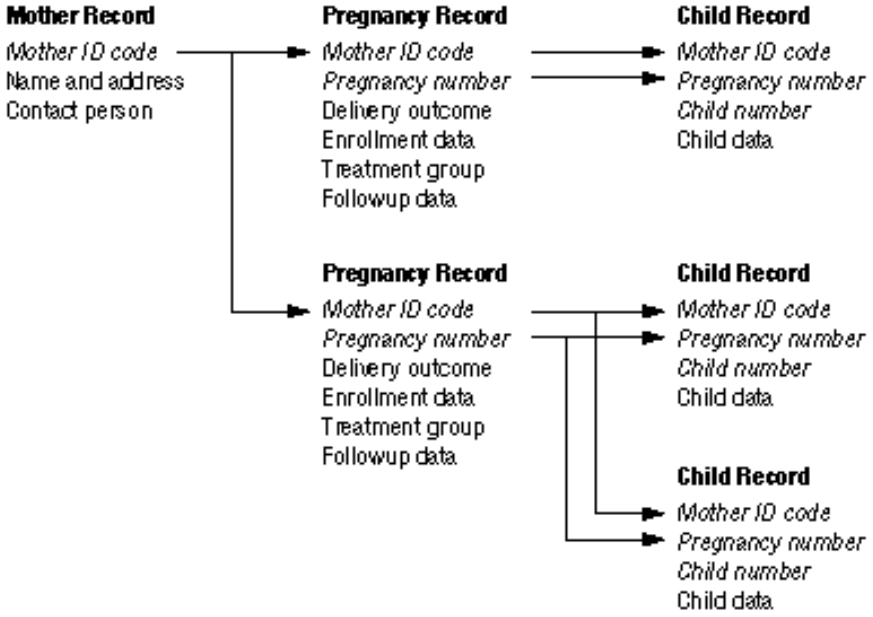


FIGURE 1. Administrative data structure for a family

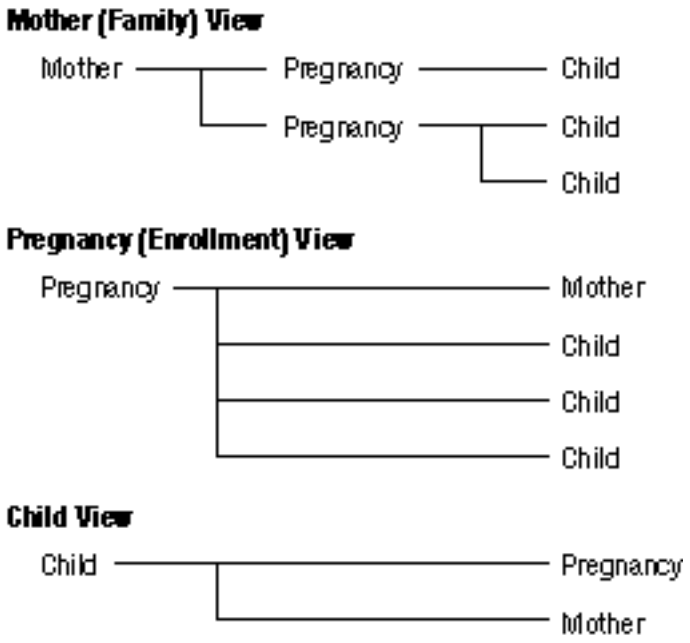


FIGURE 2. The three administrative views

Function and Implementation. As indicated above, the administrative database system supports critical subject management functions and is operated by staff members who are not computer specialists, which makes the interface a major consideration. The software application is designed to isolate inexpert users from the underlying complexity of the database and the command language by controlling access to functions and data through menus and control panels.

Figure 3 illustrates a "screen shot" of the main data access screen. A control panel at the left provides the basic record location functions, and the data are available (in read-only mode) in three windows corresponding to the three main tables. What is shown in these windows depends on the view, which is selectable on the control panel. To obtain more detailed information on a mother, pregnancy, or child record, the user activates the appropriate window by "clicking" it on with a mouse, moves the cursor to the desired record, and presses the "Enter" key. The view is then expanded to a full screen showing all the information in that record, and the data are available for inspection or modification.

Schedules and Tracking Reports. Once the enrollment information is entered for a pregnancy, the administrative database system generates an optimal contact schedule. Several reports are implemented that use the

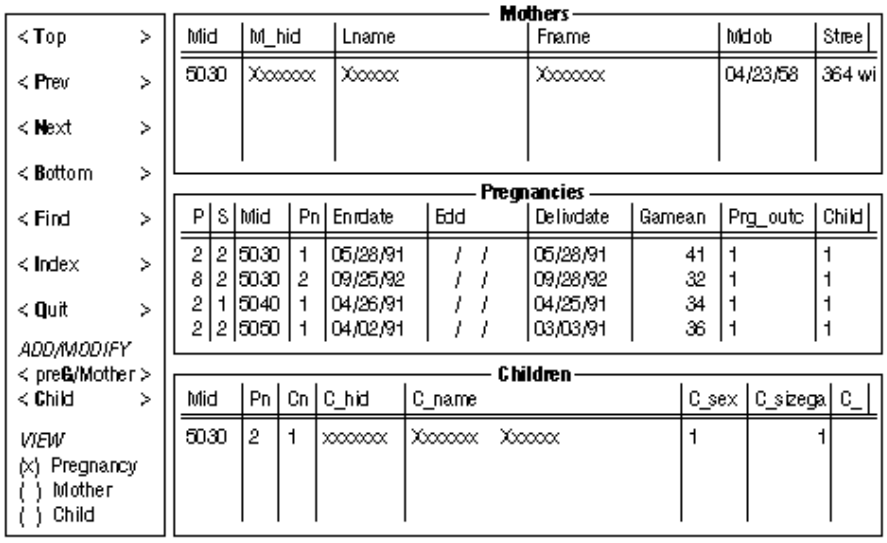


FIGURE 3. The main view screen

SOURCE: APT Foundation (New Haven, CT). Copyright 1991.

optimal and actual contact information to provide various tracking reports. For example, a sweep of the database is made periodically to identify pending and overdue contacts for all participants. Two reports are then printed: a summary list of contacts for that week and a set of "cover sheets" that report detailed information helpful in arranging the contact. The cover sheet includes the name and address of the mother and a contact person, the optimal contact schedule, and dates and outcomes of actual contacts with a special note of missed followups. (Interviewers recover some data from missed contacts through specifically designed questions.) A mail/merge data file is automatically generated when any report is printed and can be used with any word processing program to generate contact letters or other correspondence.

Security. Protecting the identities of cocaine-dependent women and their families is a requirement of the administrative database system. Therefore, a considerable effort is made to make these data secure. All information that might be used to identify any individuals, including names, addresses, Social Security numbers, and hospital unit numbers, is stored in encrypted form. The data are displayed or reported in decrypted form only for users who enter the correct encryption key when the application is loaded. The encryption key can be changed periodically by a single staff member who has access to the software. It is also possible to use the database system without providing the encryption key; the sensitive data appear garbled, but the application functions otherwise.

Additions to the Administrative Database System. The administrative tasks have grown in complexity, particularly as the Mothers Project expanded its coordination with other drug abuse studies based in New Haven. There are ongoing studies, for example, on children of cocaine-dependent mothers. The flexible database structure makes it possible to switch among mothers, pregnancies, and children as units of analysis for the purposes of new projects and substudies that become attached to the Mothers Project. It is also a simple matter to add other tables to the database and incorporate them into any of the views. More detailed information on health care utilization by mothers and children has been added to the tracking system in this manner and is immediately accessible through the interface.

Because the underlying software package features strong reporting and querying tools and has sophisticated tools that can attach such modules to existing applications, advanced users can add significant functions to the database system without additional programming.

Analytic Database System. In contrast to the administrative functions that require record-based processing, questionnaire and clinical data intended for analysis are processed at the *data set* level, that is, in *batches*. Easy access to individual records is not a priority. Furthermore, whereas administrative records include at most a few dozen fields (i.e., variables), questionnaire records typically include several hundred. Because complex relations among questionnaire records are not required, nonrelational or flat-file data management methods are suitable, such as those provided by statistical packages.

The analytic database consists of a series of data sets, one for each kind of contact or interview, all linked by a single key field (the subject identification code) in a one-to-one relationship. Each data set contains many hundreds of fields. Reports required by even the simplest analyses include statistical functions not found in business-oriented relational database packages. All this argues for using a statistical package with strong flat-file handling and reporting capabilities to manage the analytic data.

Administrative data are always processed by project staff members, but analytic data sometimes are made available to outside collaborators who have different computing environments. Thus, the “cross-platform” capability of a statistical package is a major factor. The package used includes a complete flat-file database programming language in addition to a standard set of statistical functions and procedures. This particular package is in wide use on virtually every computing system used by research projects and has a “data transport” utility that makes it possible to transfer data between different operating systems.

The flow of analytic data is diagramed in figure 4 and explained in the sections that follow.

Data Entry. As noted, two characteristics of the Mothers Project analytic database system are that records can be processed in batches and that records from questionnaires can be fairly large. Therefore, the authors chose to implement a traditional keypunching procedure for analytic files. Data collection instruments are designed so that data entry fields are of fixed length, column specified, and located in the right margin of each page. Completed forms undergo a preliminary review where they are checked for completeness and obvious coding errors. Forms with errors are returned immediately to the field staff for correction. Approved forms are assembled in batches and transferred to an in-house data entry professional. Using a keypunch emulation program, data are keypunched

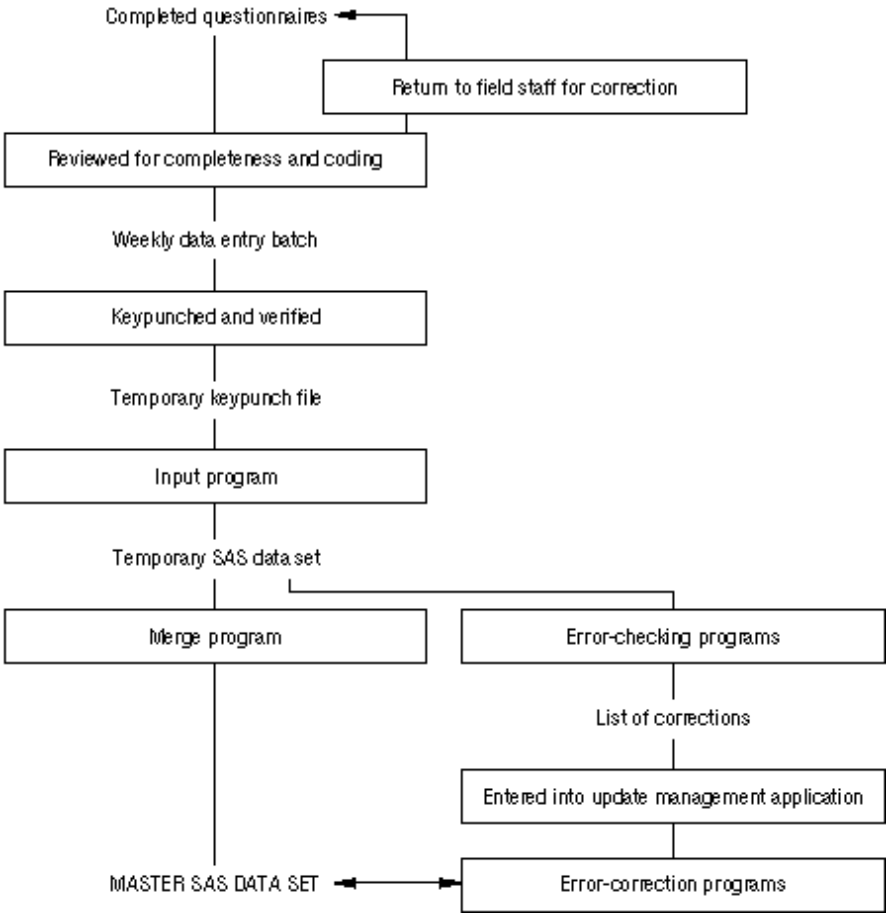


FIGURE 4. *Data flow diagram for analytic (questionnaire) data*

and verified (i.e., keyed twice). The data entry operator then returns a file of newly entered records to the data manager.

With well-designed data collection forms and appropriately trained data entry operators, data entry rates of more than 10,000 keystrokes per hour have been achieved. This style of data entry appears well suited to projects that do not have in-house data entry capability because outside data entry service bureaus usually can keypunch traditional, column-specified data.

The Case for Keypunching

As an alternative to keypunching the analytic data, the statistical package's built-in data entry facility could have been used. Using this facility, an interactive system with point-of-entry error checks and branching could have been developed. Mothers Project staff members found the keypunching procedure more efficient for three reasons. The first is data entry speed: No screen-oriented, interactive system can match the speed attained by a professional keypunch data entry operator. Second, trapping errors during data entry may sound appealing, but this practice slows down data entry. Keypunch errors, made by the data entry operator, can be detected by double-pass data entry or verification. Errors detected by range and consistency check routines are almost always *coding* errors that were made by the person who filled out the form and were not detected in the coding review. Those persons usually are not available to resolve coding errors that are detected during data entry. Thus, the data entry operator has to stop and decide whether a coding error can be resolved immediately or must be deferred; even a brief pause significantly interrupts the flow of a data entry session. The third reason for traditional keypunching of long records concerns development costs. It would have taken time to develop a responsive, interactive data entry system for the Mothers Project data collection instruments. Even simple changes to the data collection instrument would have caused delays while the data entry system was being modified. Thus a keypunch style of data entry for long records is used that permits delayed processing of batches.

Error Checking and Master File Updating. Each data collection form has an associated *input* program, one or more *error-checking* programs, and a master file *update* program. These programs are run on batches of newly keypunched records. The input program converts the keypunched ASCII data to the "native" format of the statistical package. The temporary file thus created is then passed to the error-checking program and to the update program. New records are checked for errors and simultaneously added to master data sets.

The printout from the error-checking program(s) is checked by the data manager and passed on to the field staff for resolution. Corrections always are made to master data sets (i.e., not to keypunch files), and all changes are completely documented. Auditing the changes to analytic data is considered to be so important that a specialized database application is employed solely for this purpose. As corrections are identified, they are entered into the system as "data value update" records. Each data value update record consists of a record ID, a variable name, and the correct value. The changes are made to the master data files using the error-correction program written by the specialized software. These programs generate printouts that clearly document each change that is made.

Data Documentation. The primary function of the analytic database system is to prepare files for analysis. A major part of this task is to develop and distribute comprehensive data documentation. The specialized application is also used for this purpose. In addition to managing data *value* changes, the software is designed to manage data *descriptors* for the databases. The software also provides for the development of descriptors at the database, data set, data element (i.e., variable), and scale (i.e., aggregate data element) levels and includes extensive reporting capabilities.

SUMMARY

The database management functions for the Mothers Project are arranged into administrative and analytic task groups, and separate systems are devised for each. The task groups can be distinguished not only by differences in data structure but also by interface requirements. The administrative database system uses a relational database technology, whereas the analytic database system employs more traditional flat-file methods. Although the database management systems are complex, they are based on standard database practices, used in widely available software packages, and run on inexpensive desktop computing equipment.

REFERENCES

APT Foundation. Mothers Project database system. New Haven, CT: APT Foundation, 1991.

Date, C.J. *Introduction to Database Systems, Vol. 1*. 5th ed. Reading, MA: Addison-Wesley, 1990.

Stevens, A.C. *Data Base Development*. Portland, OR: Management Information Source, 1987.

AUTHORS

Peter A. Charpentier, M.P.H.
Research Affiliate in Epidemiology
Yale University School of Public Health
60 College Street
New Haven, CT 06520-8034
(203) 785-7178 or (203) 785-6386 (Tel)
(203) 737-4260 (Fax)
charpentpa@maspo2.mas.yale.edu (Internet)

Richard S. Schottenfeld, M.D.
Associate Professor of Psychiatry
Yale University School of Medicine
Director
Substance Abuse Treatment Unit
Connecticut Mental Health Center
34 Park Street
New Haven, CT 06519
(203) 789-7079, ext. 308 (Tel)
(203) 789-7087 or (203) 789-7088 (Fax)
schottenrs@maspo3.mas.yale.edu (Internet)

Designing a Research Database Management System

Kathryn S. Dawson and Sidney H. Schnoll

INTRODUCTION

For a Perinatal-20 study, many data were to be collected over an extended period. A system for managing the data was needed so that the integrity of the data, and the quality of the research based on the data, could be ensured from the study's start. More than 60 types of data collection forms were designed. To monitor a subject's progress through the treatment program, some forms were collected at multiple time points. Data from more than 45,000 forms were entered throughout the course of this 5-year study.

Before developing the database, it was necessary to have in place the research protocol that formally described the testing instruments and data collection forms used. The time points at which the forms were to be collected also had to be well defined. The use of this information allowed the database system implementation to include methods of validating the data. Data validation was accomplished on several levels: The system checked for data entry errors on an item-by-item basis and also jointly guaranteed the validity of several items. To ensure that only authorized personnel had access to the data, security protections also were implemented.

The design of the database was based on a relational database structure that allowed data to be viewed logically in tables of information. In general, a table was designed to correspond to each form. Columns in the table represented data items on the form. Ensuring that appropriate forms shared common fields allowed the data from multiple forms to be appropriately merged later. One advantage of this type of database structure was that additions to the database could be made easily. This proved to be an important advantage; at several times throughout the course of this study, additional forms were added to the protocol.

The following discussion describes various guidelines that were considered in the design and implementation of this project's database management system. Although the implementation of any database management system depends in part on the computer software chosen, the overall design considerations are the same.

DATA ENTRY FORMS

Well-designed data entry forms are essential to collecting data in any large project. In most large projects, numerous staff members enter data. To ensure consistency in the data collection, a manual should be written that defines all data entry fields and should be read and used by all personnel who collect data. Otherwise, there is no assurance that consistent meanings will be given to the data.

In general, the flow of the research data can be described according to the diagram shown in figure 1. The data originating from the subject are recorded on the data entry form and then entered into the computerized database. A data entry form can range from a piece of paper to a computerized data entry screen. These forms can correspond to predesigned testing instruments, such as the Addiction Severity Index (McLellan et al. 1985), or to project-specific information, such as session attendance or urine toxicology results. Each data item to be completed on the form is a field. The form should be designed to ensure that the valid values for each field are well defined.

In general, forms fall into two categories. One-to-one forms are those that are collected only once for each study subject. In contrast, one-to-many forms can be collected at one or more time points.

In the illustrative one-to-one Demos form shown in figure 2, the valid entries for the race field are clearly specified. When possible, short character or numeric codes should be used to enter data into each field. Variable length fields always should be avoided because they are difficult to validate and analyze later. *DSM-IV* (American Psychiatric Association 1994) and *ICD-10* (World Health Organization 1992) diagnostic codes, rather than text fields, can be used to describe diagnoses and medical conditions. The data entry fields should be arranged in a manner that makes it easy for the person who enters the data. For example, all fields are aligned in the form shown in figure 2.

Each form should contain a single field that uniquely identifies each study subject. Although names, Social Security numbers, and other personal identifying information may be included on the paper data entry form, to guarantee patient confidentiality, the database should not include these identifiers. Ideally, a unique study-specific case number should be assigned to each new subject. The project director should maintain the list, matching the case number with names in a secure location separate from the database. For data entry purposes, the unique case number should be included on all forms.

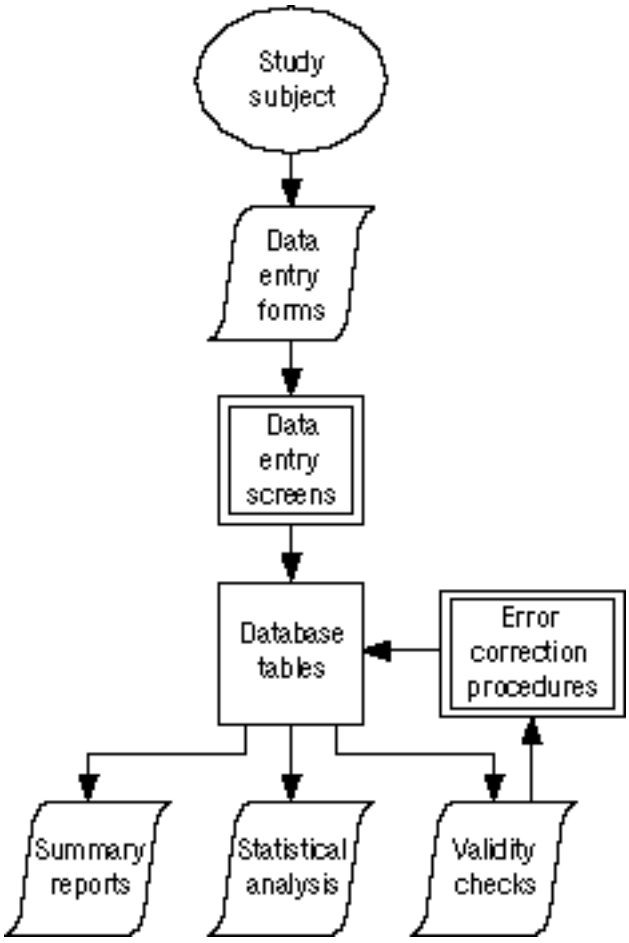


FIGURE 1. *Data flow from subject to database to application*

Figure 3 illustrates a one-to-many form called Beck (based on the Beck Depression Inventory) (Beck et al. 1988) that is collected at three time points. Within both one-to-one and one-to-many forms, a set of fields should be defined that uniquely identifies a single particular data entry form. This set of fields is used to retrieve the data for a single subject at a fixed time or to merge those data properly with data originating from other forms. The case number can serve as the unique identifier for one-to-one forms. For one-to-many forms, at least two fields must be used. For the one-to-many form in figure 3, the combination of case number and the date can be used. When more than one person collects the same data on a single subject, a tester identification code also may be needed.

Form

Intake Demographic Information
Collected at Admission

Subject Case Number: ___ ___ ___ ___ ___

Date of Admission (mm/dd/yy): ___/___/___

1. Date of Birth (mm/dd/yy): ___/___/___

2. Race: A. White B. Black ___
C. Hispanic D. Other

3. Gender: M F ___

4. 5-Digit ZIP Code of Residence: ___ ___ ___ ___ ___

Table Definition

Table Name: Demos

case	admit	dob	race	gender	ZIP

Table Documentation

Table Name: Demos

Field Description	Variable Name	Key Fields	Data Type	Validity	Null
Subject Case Number	case	key	character-5		N
Date of Admission	admit		date		N
Date of Birth	dob		date		N
Race	race		character-1	A,B,C,D	N
Gender	gender		character-1	M,F	N
5-Digit ZIP Code	ZIP		character-5		Y

FIGURE 2. *Form, table definition, and table documentation for a one-to-one form*

For one-to-many forms that are collected over time, a visit indicator, in addition to the date of the evaluation, may be helpful. For example, the Beck form in figure 3 is filled out during three specified assessments. Therefore, the visit field lists I, M2, and D as valid entries, which indicate specific data entry time points (Intake, Month 2, and Discharge) and may be more useful than the specific date on which the data are entered. For

Form

Monthly Beck Depression Inventory (BDI)*
Collected at Intake, Month 2, and Discharge

Subject Case Number: ___ ___ ___ ___ ___

Date of Evaluation (mm/dd/yy): ___/___/___

Visit (I, M2, D): ___

BDI Score (0-63): ___

Table Definition

Table Name: Beck

case	evaldate	visit	score

Table Documentation

Table Name: Beck

Field Description	Variable Name	Key Fields	Data Type	Validity	Null
Subject Case Number	case	key	character-5		N
Date of Evaluation	evaldate		date		N
Visit	visit	key	character-2	I, M2, D	N
BDI Score	score		integer	0-63	Y

FIGURE 3. *Form, table definition, and table documentation for a one-to-many form*

KEY: I=Intake; M2=Month 2; D=Discharge

*SOURCE: Beck et al. 1988

example, the forms corresponding to a given assessment could have varying dates and hence may be difficult to merge by assessment date. The use of a visit indicator also eliminates the necessity of determining the type of assessment, by comparing two date fields.

Null vs. Unknown

Answers may not be provided for all questions. In general, nonresponses can be categorized as either missing or unknown. A missing or null value is an indicator of a nonresponse, which is distinguished from an

"unknown" response where the subject has responded to the question. Careful review of each question will determine when either of these values can be considered the valid option.

Null responses are appropriate if a subject does not complete an entire form or, for some reason, is not presented with a given question. These responses can be indicated on the data entry form as blanks. The implementation of null values in the database depends on the database computer software. Some associate null values with blanks. Other database software requires the use of a dummy value to represent a null value. For example, for a multiple-choice question with valid options A, B, and C, the additional value X can be used to indicate a null value. However, when analyzing these data, it is important to ensure that these null values are handled specifically as missing data. Certain fields, such as the fields that compose the unique identifiers in a form, should never assume null values. Null values also can be restricted from data fields that should be readily available, such as gender on the demographic form or fields considered of primary importance to the research questions being addressed.

In some instances, multiple answers may be given for a single question. In figure 4a the question prompts for a list of physical complaints. Only three response spaces are given in this example, but the subject may have all the complaints. In the analysis of these data, the absence of a particular code is intended to imply that the subject did not have that complaint. However, this fact cannot be distinguished from a missing or unknown value. A preferred format for this question is given in figure 4b, in which an answer is given for each complaint and a missing value can be assumed to have been omitted deliberately.

For some multiple-choice questions, an "other" option can be included in the list of valid answers. In conjunction with an "other" response, a small text field, where the option can be listed, should be included on the form. Although this text field need not be entered into the database, it is available for later review. If the "other" category is checked many times, the data forms can be updated to reflect additional options.

For multiple-choice questions, an "unknown" option can sometimes be included in the list of valid options (figure 4b). Whenever there is even a remote possibility that the subject will not know the response to a question, this option should be listed. The unknown option may not be appropriate for all questions, such as gender on the demographic form.

a. Limited Number of Responses

List Medical Complaints: ___ ___ ___

A. Headache B. Dizziness C. Pain

D. Fatigue E. Numbness F. Other _____

b. Preferred Format

Medical Complaints:	Circle One		
A. Headache	Y	N	U
B. Dizziness	Y	N	U
C. Pain	Y	N	U
D. Fatigue	Y	N	U
E. Numbness	Y	N	U
F. Other _____	Y	N	U

FIGURE 4. *Two formats for a question with multiple responses: limited number of responses and preferred format*

RELATIONAL DATABASE

For any large research project, a relational database structure is recommended (Date 1982, 1983). Many software packages are available to implement this structure on a mainframe, stand-alone, or networked personal computer. The minimal components of relational database software should include (1) procedures for creating tables, (2) procedures for developing the data entry process, (3) validation checks at data entry, (4) built-in security procedures, and (5) query language to facilitate retrieval of the data. As discussed in the "Null vs. Unknown" section above, the implementation of null values is also helpful.

In this type of database, the data can be viewed logically in tables. Each table contains information that, typically, is related in some way. In figure 2 the columns of the Demos table are shown. This table contains the intake demographic information for each subject. The Beck table, illustrated in figure 3, contains the information concerning a subject's level of depression. Each table comprises columns and rows. A column

corresponds to a single data field, or attribute. The values in the column are drawn from a predefined set of values. For example, the gender column in the Demos table will contain only the values M and F. Each row in a table is a collection of column attributes that describe a single study entity. A row in the Demos table describes a study subject, whereas a row in the Beck table describes the subject's depression on a given visit. To avoid use of data more than once, multiple instances of an identical row should not be included in a table.

There are several advantages associated with this type of database design. Viewing the data as a series of tables clearly presents the overall organization of the project's database, which helps facilitate the manipulation and retrieval of the data for reporting and analytic purposes. This is especially useful when a particular application requires that data be merged from several tables. Another important advantage is the ability to easily accommodate a nonstatic data system. It is not uncommon for research instruments and data fields to be added to a database after it has been implemented. The relational database structure is well suited to a dynamic environment where these changes can be made with little or no modification of preexisting reports, analyses, or procedures. Last, a rigorous set of guidelines, normal form theory (Dawson and Parker 1988; Kent 1983), has been developed that is helpful in designing this type of database. The set of properties contained in the guidelines ensures that the integrity of the data is maintained and results in an overall database design that is easier to manipulate for reporting and analysis purposes.

DATABASE TABLES

A useful guideline for creating the database tables is to associate a single table with a single form. Even in cases where a form might consist of only one or two pieces of information, a separate table is recommended. In this way the absence of a row in a given table can indicate that a particular subject's form is missing or has not been entered into the database. Each field in the form can correspond to an attribute in the table. In figures 2 and 3, tables are shown that are associated with each of the illustrated forms.

Within a form, a set of fields is noted that can be used to identify a single, specified data entry form. In general, the set of table attributes corresponding to these fields is called the key of the table and therefore uniquely identifies each row of the table. As examples, the attribute case number uniquely identifies a row in the Demos table (figure 2), and the combination of attributes, case number and visit, identifies a row in the Beck table (figure 3). Because the key identifies each row in a table, it

follows that no two rows in a given table will have identical entries for these key fields. The key attributes are used to retrieve specified single rows from a table. The key values also are used to merge data from two or more tables. Hence, for programming convenience, common key fields from multiple tables should be assigned the same variable name. For example, suppose a researcher were interested in knowing whether the mean Beck Depression Inventory score at intake was related to the age of the subject. Age at intake can be determined from fields in the Demos table (figure 2). The rows of the Beck table (figure 3) with the visit "I" could then be merged with the rows of the Demos table over the common field case number. Therefore, case number is the key value that allows all the information on a given subject to be merged.

NORMAL FORMS

As mentioned earlier, the relational database structure has a theoretical basis, normal form theory, that can be used to validate the design of tables in the database. The goal is to ensure that the integrity and quality of the data will be maintained throughout the process of updating and accessing the database. This theory, which to some extent is founded in common sense, also may be helpful in designing the data entry forms and creating the corresponding tables. Although it is beyond the scope of this chapter to describe this theory in detail, the following guidelines, which define three normal forms, are useful.

First Normal Form

To ensure the validity of updates and data retrieval, each column in a table should contain single data items whose values are selected from a predefined set of possible values. To illustrate a table that would violate this constraint, consider the form shown in figure 5a. In the table construction shown in figure 5b, a single three-character text column, "answers," is defined to contain the answers to the three multiple-choice questions. If a subject answered the three questions as B, D, and C, the value entered in this column would be the three-character value "BDC." Several practical difficulties are associated with this construction. Validating this field involves comparing its value with the set of all three-letter combinations of the letters A through F. To analyze or update the values for one of the three questions implies the user must take care to ensure that only the appropriate character is modified. A more reasonable approach, which satisfies first normal form, is shown in figure 5c. Here three columns are used that correspond to each of the three questions, "condom use," "intox during sex," and "needle sharing."

a. Form

HIV Transmission Risk Behaviors

Subject Case Number: ___ ___ ___ ___ ___

Date of Evaluation (mm/dd/yy): ___ /___ /___

1. Frequency of condom use in past 6 months: ___
   A. Always B. Most of time C. Half the time
   D. Seldom E. Never F. Not applicable/abstinent
2. Frequency of intoxication during sexual encounters: ___
   A. Always B. Most of time C. Half the time
   D. Seldom E. Never F. Not applicable/abstinent
3. Frequency of needle sharing: ___
   A. Always B. Most of time C. Half the time
   D. Seldom E. Never F. Not applicable/non-IV user

b. Table in Violation of First Normal Form

Table Name: HIV

case	evaldate	answers

Field Description	Variable Name	Key Fields	Data Type	Validity	Null
Subject Case Number	case	key	character-5		N
Date of Evaluation	evaldate	key	date		N
Answers	answers		character-3	(A-F)(A-F)(A-F)	Y

c. Preferred Table Construction

Table Name: HIV

case	evaldate	condom	intox	needle

Field Description	Variable Name	Key Fields	Data Type	Validity	Null
Subject Case Number	case	key	character-5		N
Date of Evaluation	evaldate	key	date		N
Freq Condom Use	condom		character-1	A-F	Y
Freq Intox During Sex	intox		character-1	A-F	Y
Freq Needle Sharing	needle		character-1	A-F	Y

FIGURE 5. *Illustration that shows violation of first normal form: form, table in violation of first normal form, and preferred table construction*

KEY: HIV=human immunodeficiency virus

Second normal form and third normal form depend on the identification of dependences that exist between or among data fields within a form. In this context, a dependence is defined as follows: Field X is dependent on field Y; if given the value of field Y, a single value for field X is retrieved. For example, in the Demos form (figure 2), date of birth is dependent on case number; that is, given a subject's case number, a single date of birth can be retrieved. Note that the converse is not necessarily true; that is, case number is not dependent on date of birth because, theoretically, several cases could be retrieved with the same date of birth. Dependences can be defined in terms of set of fields. In the Beck form (figure 3), a particular subject's monthly score is dependent on the combination of the case number and visit. This is because, given just the case number, several rows of Beck data theoretically could be retrieved because this form is collected repeatedly over time. However, given the combination of case number and visit, at most one row of data and hence one score will be retrieved.

Second Normal Form

Each field in the table should depend on the entire key and not a subset of fields in the key. Consider the form shown in figure 6, which was designed to collect weekly information concerning subject participation in certain self-help groups; a table associated with the form is also shown. The key for this table is the combination of case number and visit. The attributes associated with questions 1, 2, and 3 are dependent on both these key fields. In contrast, the last question or attribute is dependent solely on the case number because the answer to this question is constant over time. This table design therefore violates second normal form. If a coding error for the last question is made at one of the visits, these data are inconsistent for that subject and may later be reported incorrectly. In general, any field that is constant over time should not be collected in a one-to-many form. Collecting it more than once presents the chance of introducing an error in that field each time the data are collected and modified. The last question should be removed from the form shown in figure 6 and placed more appropriately in a form such as the one-to-one demographic form (figure 2).

Third Normal Form

All data fields that constitute a row in a table should be dependent on the key and on no other field. Consider now the form and table shown in figure 7. This form, collected at patient admission, contains five true/ false questions as well as the total number of questions answered as true. Because this is a one-to-one form, the case number is the unique identifier of the form and key of the corresponding table. However, note that the "total" field is dependent on the combination of the case number and the set of fields associated with all five true/false questions. Therefore,

Form

Weekly Participation in Self-Help Groups

Subject Case Number: ___ ___ ___ ___ ___

Date of Evaluation (mm/dd/yy): ___/___/___

Week Number (1-12): ___

1. Currently has a sponsor (Y/N): ___

2. Attendance in past week: ___
   A. Refuses to attend B. Less than required C. As required
   D. More than required E. Far exceeds required

3. Current attitude toward self-help group: ___
   A. Positive B. Neutral C. Negative

4. Age at attending first AA/NA meeting: ___

Table Definition

Table Name: Self-help

case	evaldate	visit	sponsor	attend	attitude	agefirst

FIGURE 6. *Form and table definition that violate second normal form*

KEY: AA/NA=Alcoholics Anonymous/Narcotics Anonymous

this table violates the third normal form. If one or more of the fields associated with questions 1 through 5 were updated and the total were not also updated, the data would be inconsistent. For this example, no information would be lost if the total field were eliminated altogether. For analysis purposes, the total can easily be determined by examining the values for questions 1 through 5.

In general, when the tables are constructed in the database, dependences between or among fields must be addressed. In some situations, when dependences are determined, fields can be moved to different tables or eliminated altogether. However, in some cases it may be helpful to maintain these dependences in the database. Consider the following two-part question:

Were you employed in the past month? (Y/N) _____

If Y, enter monthly net income. _____

Form

Treatment Attitude

Subject Case Number: ___ ___ ___ ___ ___

Date of Evaluation (mm/dd/yy): ___/___/___

	Circle One	
1. I frequently disagree with the therapist/counselor.	T	F
2. I expect treatment to be difficult.	T	F
3. I could stop using alcohol/other drugs if I wanted.	T	F
4. I go to an AA meeting at least once a week.	T	F
5. I have friends who support me.	T	F

Total number of true answers (1-5): _____

Table Definition

Table Name: Attitude

case	evaldate	q1	q2	q3	q4	q5	total

FIGURE 7. *Form and table design that violate third normal form*

KEY: AA=Alcoholics Anonymous

Here income is nonnull only when the first question is answered Y. Because this field depends, in part, on the value in the employment field, this is a violation of third normal form. However, it is inconvenient in this case to move the second field to a separate table. The documentation of the table that lists these data should include a note of this dependence. The correctness of the combination of the fields also can be reviewed in periodic validation checks.

DOCUMENTATION

The properties of each table must be clearly documented. Examples of table documentation are shown in figures 2 and 3. The fields that constitute the key should be clearly noted. Possible values that a field can assume should be enumerated. Fields that cannot assume null values should be noted. Validity checks between or among tables also can be documented.

Two Necessary Forms and Tables

For research projects that study subjects over time, forms associated with each subject's start and completion of treatment should be included in the protocol. Minimally, the intake form should include the subject's case number and date of admission. Other information collected at intake can be included (figure 2). Similarly, a discharge form must include the case number and date of termination but also can contain information pertaining to the type of discharge. When these two tables are merged over their common field (the case number) the current status of all subjects can be determined: Those listed in both tables are no longer in treatment; those listed in the intake table only are still active. This concept can be extended to other phases of treatment as well. For example, a form can be incorporated to indicate the termination of the followup phase of treatment.

At any time, the intake form is the minimal set of information included in the database about a given subject. When the database is queried, the table associated with the intake form lists all the subjects currently admitted into the study. Hence, the case numbers in this table can serve as the links to all other tables in the database. This is especially useful if the protocol varies by subject, which implies that not all subjects will be listed in every table. For this reason, it is suggested that, before any other form can be entered for a given subject, a check be made that a row in the intake table exists for that subject.

Validation

Validation of the data can be implemented in several ways and at multiple times. During data entry, the value entered for a particular field can be compared with the list of valid entries. An error message can prompt the data entry technician to correct an invalid entry. Before an entire form's data are appended to its appropriate table, within-form validity checks can be made. For example, if a data field indicates that the subject is a nonsmoker, the field containing the packs-per-day data should be null. In some cases, fields within a form can be compared with data already included in the database. For example, as described in "Two Necessary Forms and Tables" above, a form's case number can be compared with the list of case numbers in an intake table to ensure that intake data have already been entered. If the intake data for that case are missing, an error message can be shown.

Although the data that have been entered can be considered valid, they still may differ from data on the data entry form. When personnel resources are available, double entry of the data can be implemented to

help eliminate this level of data entry error. Here, two different data entry technicians independently enter the data. In a timely fashion, the data are reconciled. An outline of an algorithm that has been used to implement this type of system is shown in figure 8. Another option involves letting a reviewer who is not involved with the data entry compare the data in the database with the corresponding data entry form. With either of these systems, entry errors are noted and promptly corrected.

Regular validity checks should be made on the entire database. This overall check can reexamine the validity of each field, within a form and between or among forms, depending on how the rules are defined for data entry purposes. This is especially important when dependences, as discussed above in "Database Tables," are maintained in the database. In addition, more elaborate checks, either between or among forms, can be made. For example, if a subject at admission indicated she was in the third trimester of pregnancy, forms associated with the birth of her infant should be included in the database within a reasonable amount of time. A check also can be made to ensure that all forms associated with the subject's protocol are entered in the database in a timely fashion; that is, within 3 months of admission, the month 2 Beck data (figure 3) should be entered. Last, when a subject completes the treatment program, there should be an overall check that data from all appropriate forms listed in the protocol, from intake to discharge, have been entered.

SECURITY AND PROTECTIONS

To maintain the integrity of the data, as well as ensure patient confidentiality, security can be incorporated at several levels. Although the implementation of these security procedures depends on the software and computer environment, the guidelines are the same. Only authorized users have permission to access the database. A user who can access the database may then have permission to access all the tables in the database or only a subset of these tables. Certain users may be restricted from accessing tables containing highly sensitive data. The type of access also can be specified for a particular table. Some users may have permission to retrieve, but not modify, data from a given table. Last, but important, a system for periodically backing up the database also must be incorporated and clearly documented. In that way, if there is a computer hardware failure, the data can be restored.

Data initially entered into table 1 and, independently, doubly entered into table 2.
The final table contains only the reconciled data.

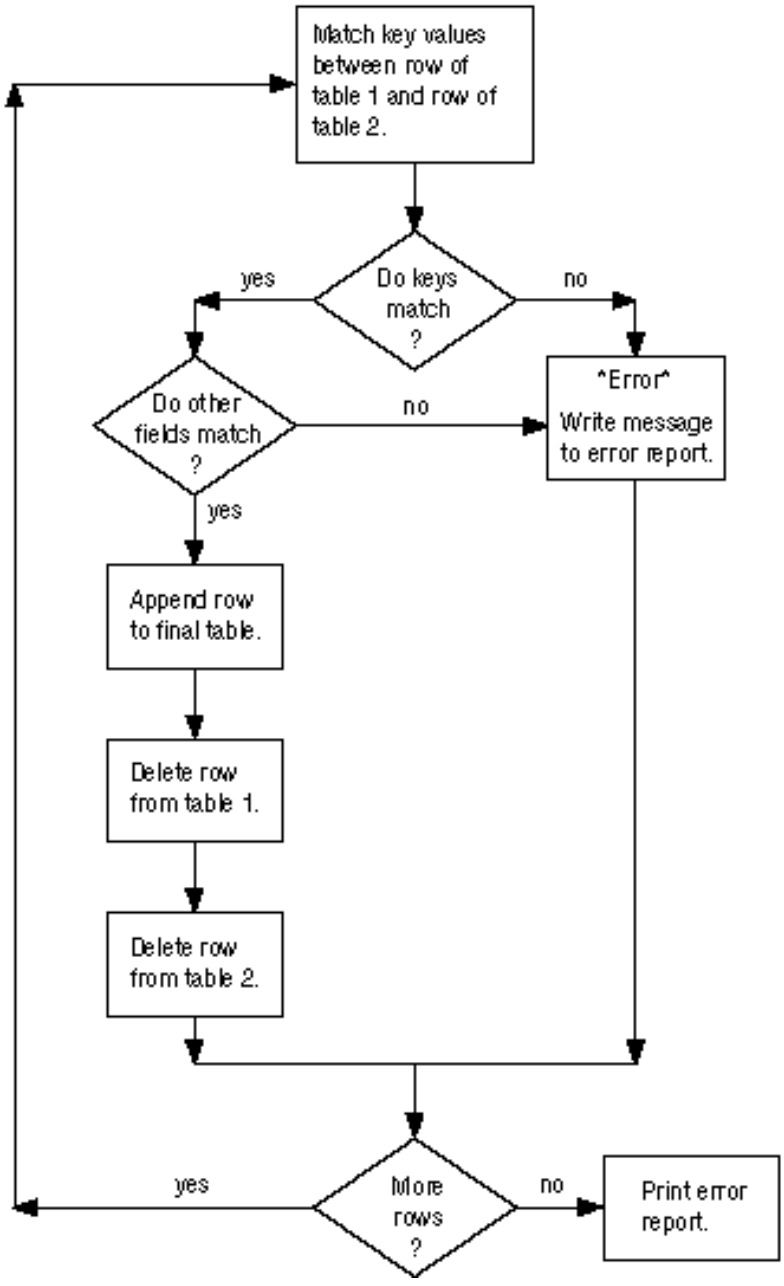


FIGURE 8. *Double-entry reconciliation algorithm*

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition.* Washington, DC: American Psychiatric Press, 1994. 886 pp.

Beck, A.T.; Steer, R.A.; and Gabin, M.G. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psych Rev* 8(1):77-100, 1988.

Date, C.J. *An Introduction to Database Systems.* 3d ed. Reading, MA: Addison-Wesley, 1982. 574 pp.

Date, C.J. *An Introduction to Database Systems, Vol. II*. Reading, MA: Addison-Wesley, 1983. 383 pp.

Dawson, K.S., and Parker, L.M.P. From entity-relationship diagrams to fourth normal form: A pictorial aid to analysis. *Computer J* 31:258-268, 1988.

Kent, W. A simple guide to five normal forms in relational database theory. *Commun ACM* 26:120-125, 1983.

McLellan, A.T.; Luborsky, L.K.; Cacciola, J.; Griffith, J.; McGahan, P.; and O'Brien, C.P. *Guide to the Addiction Severity Index: Background, Administration, and Field Testing Results*. DHHS Pub. No. (ADM)88-1419. Rockville, MD: National Institute on Drug Abuse, 1985.

World Health Organization. *International Statistical Classification of Diseases and Related Health Problems. 10th Revision.* Geneva: World Health Organization, 1992.

AUTHORS

Kathryn S. Dawson, Ph.D.
Senior Programmer Analyst
Department of Biostatistics
(804) 828-9824 (Tel)
(804) 828-8900 (Fax)
dawson@gems.vcu.edu (Internet)

Sidney H. Schnoll, M.D., Ph.D.
Professor
Department of Internal Medicine and Psychiatry
Chairman
Division of Substance Abuse Medicine
Medical College of Virginia
(804) 828-9914 (Tel)
(804) 828-9906 (Fax)
sschnoll@gems.vcu.edu (Internet)

Virginia Commonwealth University
1101 East Marshall Street, Box 980032
Richmond, VA 23298-0032

Compensating for Deficiencies in Perinatal Data Sets: Parametric Perspectives

A. Scott Tippetts and Paul R. Marques

INTRODUCTION

Considering the investments made in a research demonstration project—including the hours of work in research design, provision of services, data collection, data entry, money spent, and professional reputation—great care should be taken when testing hypotheses, but often it is not. A small amount of extra effort at the data analysis stage frequently can yield great benefits in terms of more accurate findings. The reason: Very few data sets are ideal, and among nonideal data sets, those of perinatal drug abuse treatment projects have few equals.

Data deficiencies are sometimes apparent without having to search for them, showing up as highly abnormal or extreme values or as missing data or responses. Other data deficiencies are not noticed until carefully pursued. When corners are cut (e.g., not checking distributions for normality, not performing basic diagnostics of residuals), the researcher risks missing the real relationships or reporting spurious relationships. Although some researchers excuse these shortcuts with the rationale that analysis of variance (ANOVA) and regression are relatively robust against violations of assumptions, it is easy for even a single overlooked data error to completely negate the outcome of a statistical test.

Even for researchers who regularly check for data deficiencies and irregularities, and therefore often discover them, the question arises, What should be done with the deficiencies? This chapter covers some basic "textbook" approaches for dealing with data deficiencies as well as some less used, more imaginative techniques.

NONNORMALITY

Before any hypothesis testing is undertaken, data *always* should be examined to assess the degree to which assumptions may be violated, especially assumptions that require a data set to be normally distributed. Even when a data set includes many variables, the time involved in quickly checking the distribution of each variable can pay big dividends in the long run. Histograms are the most basic visual check, and many statistical

software packages can superimpose the outline of the normal distribution over the histogram. Abnormal distributions, such as severely multimodal distributions, probably should be excluded from parametric analysis altogether, unless they can be modified in some way to become moderately normal (e.g., by collapsing them into fewer categories, such as by dichotomizing).

Conventional Transformations

Most distributions are fairly unimodal and taper into tails at each end, showing some semblance of normality. For such variables, ANOVA and regression are robust; nevertheless, the power of parametric procedures can be improved on—sometimes dramatically—by linearly transforming the variables into more normally distributed versions. One simple way to assess the need for a transformation is to check the degree of skewness, because skewed data represent perhaps the most common type of distribution abnormality. Any statistical package worth using should provide the skewness and standard error of skewness for each distribution. A good rule of thumb is to be skeptical of any distribution whose magnitude of skewness (in either direction) is greater than three standard errors of skewness, although in large samples (e.g., more than 1,000 cases), the standard error of skewness becomes so small that this rule of thumb becomes unreasonable and is likely to be exceeded even with only slight amounts of skewness. In cases where the distribution is positively skewed (i.e., having a long tail tapering to the high side), taking the square root, cube root, log, or reciprocal of the variable can compress the tail, "pushing" the distribution toward a more normal form. By contrast, when the distribution is negatively skewed (i.e., having a long tail tapering to the low end), using the square, cube, or exponential (*e* raised to the power of the variable) pushes the lower values of the distribution closer in. When more than one transformation seems to be successful in achieving normality, one can decide which transformation works best based on the new skewness statistics for the transformed variable because the standard error of skewness remains the same regardless of transformation.

A word of caution: Some transforms are sensitive to certain ranges of values, such as power transforms (e.g., squares, square roots) on data that straddle the value of 1. Root transforms converge toward 1 on *both* sides, and powers greater than 1 diverge away from 1 toward infinity or toward zero. For this reason, before using a power transform on a scale that includes values between 1 and zero *and* values greater than 1, it may be appropriate to first rescale via multiplication so that no values lie between zero and 1 (exclusive) or all values range from zero to 1. Log transforms also can be undesirable with values less than 1, and logarithms of values less than or equal to zero are undefined. Again, rescaling through

multiplication can make distributions more amenable to log transforms and may optimize the transform (see below).

Finding nonnormal variables that can be normalized through linear transformation does not imply that the original, untransformed variable should be discarded or ignored during hypothesis testing. The original variable may produce better relationships with other variables of interest than does the more normal transformed variable, even with sufficiently random residuals. In many such cases the use of the original variable is permissible because of the robustness of many parametric procedures. However, the linear transformation often not only helps to better satisfy the assumptions needed for the procedures but also may uncover a relationship not detectable until the variable is normalized. Indeed, one might wonder how many significant yet moderate relationships have been reported that were actually strong relationships, watered down by using untransformed variables that strayed only marginally from normality. This was the case with hair cocaine data from mother-infant pairs, where square root transformations normalized the distributions, strengthening the correlation from 0.33 to 0.41 (Marques et al. 1993).

Maximizing Normality

Some may question the practice of regularly checking and transforming variables that are in need of normalization, but such a view fails to recognize that many metrics are created in a somewhat arbitrary manner that may not necessarily be isomorphic with the phenomenon being measured. Some of the most obvious such measures are decibels (sound volume) and the Richter scale for measuring seismic activity, both of which have a logarithmic relationship between magnitude and the measured values. Standard radioimmunoassay values are usually based on a log-logit plot of drug concentration to the ratio of bound-to-free radioactivity. Even simple survey scales that sum positive (or negative) responses to a series of yes or no questions may linearly distort the metric of the underlying "ideal set" of response patterns. This is why some psychometric instruments are more sensitive than others; better instruments or scales produce values that correspond more closely to the natural progression of the underlying concept or latent variable. Such issues from measurement and data theory are abstract, but the process of mapping phenomena onto quantifiable scales often creates the frame of reference, or metric, leading often to the mistaken notion that the metric came first. The process of linearly transforming variables sometimes may be nothing more than restoring the isomorphism of the metric to the nature of the phenomenon.

There is another rarely used (and to some controversial) variant of linearly transforming variables to satisfy assumptions of normally distributed data.

If one accepts the concept that many metrics or scales are somewhat arbitrarily mapped from the "ideal" metrics of the phenomena, then the process of transformation may uncover these "ideal," inherent metrics. If so, then rather than looking only at transformations involving a few limited exponents such as 0.5 (square root) or 2 (squared), one might solve for the power that maximizes the normality of the distribution along the entire scale of possible exponents. The procedure for finding the optimal linear transform involves creating an extra variable that consists of the *expected* normal scores from a normal distribution and corresponds to the ranks of the values of the variable to be transformed. If the Statistical Package for the Social Sciences (SPSS) (Norusis 1992) is used, these expected normal scores can be computed with the command RANK/INTO NORMAL. Then, the original variable (RAWVAR) is modeled in a basic nonlinear regression equation

$$NORMLVAR=((RAWVAR^{POWER})–MEAN)/STDDEV$$

that predicts the new normalized variable of expected normal scores (NORMLVAR). The parameters for MEAN and STDDEV are estimated for the new transformed variable. The estimated parameter POWER is the exponent to which the original variable RAWVAR must be raised to best produce the normal distribution NORMLVAR. The solution to the power transform question is sample dependent and may not differ much from one of the basic transformations (e.g., square, square root) previously mentioned. In addition, this process can add much time (and therefore cost) to the preanalysis stage, while gaining marginally little over the traditional transformation powers. If a variable requires normalization, this approach is as defensible as any other, because the original relationship among values is not altered.

The procedure described above is based on power transforms, which work well in many instances, but sometimes the distribution in question requires a log transform. This is fairly common when the scale is constrained at the lower end (often at zero), which produces something not unlike an F-distribution. As with the power transform, there is usually a particular base whose log produces the most normal distribution. Because most computers are limited to logarithms of the bases 10 and *e*, some scale adjustment is necessary before taking the transform. By multiplying/ dividing the distribution by a constant, then adding/subtracting a constant, the original distribution first is rescaled around the log's base so that the ensuing log transform will maximally approach normality. Although this three-step procedure may seem excessive, remember that multiplication and addition *do not change the shape of the distribution at all* but only shift the metric. Any analysis of these rescaled (but still untransformed) data will produce exactly the same results as the original distribution in

terms of test statistics and probability values. However, the log-transformed distribution of the rescaled data will be closer to normality than that resulting from the unrescaled original distribution, and the transformation is still completely linear.

BISERIAL CHANGE VARIABLES

Although there are various objections to consolidating repeated measures into a single "change" measure on an individual case basis, it is nevertheless common practice for researchers to do so for the sake of simplicity, lack of analytic sophistication, or the need to overcome the violation of the assumption of independent observations without having to resort to more complex repeated measures procedures like multivariate analysis of variance (MANOVA). Because the practice is certain to continue, it is helpful to be aware of some of the more basic alternatives that can be used when a single variable representing change is desired.

Simple Difference and Percent Difference

The first type of change is the simple difference, or posttreatment measure minus the pretreatment measure (or baseline). Much has been written about the mathematical issues relating to difference scores, and many researchers have proposed corrections to produce less biased estimates of a true difference (e.g., Chronbach and Furby 1970; Harris 1963). Many researchers continue to compute simple difference scores anyway, either because of unfamiliarity with the alternatives or to sidestep the complexity of the correction formulas. Yet numerous researchers have abandoned the use of difference scores (and their various corrections) because of a conceptual consideration, not a mathematical one: Nearly all fail to adequately address the idea of "relative change."

For example, regardless of how the difference score is computed or corrected, a change of –200 ng in measured hair cocaine levels may be an unimpressive reduction for a subject who measured 1,500 ng at basal, but such change arguably represents a dramatic improvement for a subject who began at a level of only 400 ng. This is particularly true in substance abuse research where, as a result of different degrees of drug tolerance, a low or moderate exposure level for one person may be an extremely high level for another. Data for such phenomena often manifest this relativeness through a high positive correlation between the baseline value and the absolute value of the simple difference score. For this reason, most researchers who compute single "change" variables often prefer the more subject-specific *percent-change relative to baseline*, in which the –200 ng from a baseline

of 400 ng would be represented as a 50-percent decrease but from a baseline of 1,500 ng would be represented as only a 13.3-percent decrease. The intuitive appeal of this computation is that change is relative to each case's baseline. If all cases have the same baseline (such as under the most ideal laboratory-controlled conditions), the same percent-change would also represent the same amount in terms of raw difference. In clinical studies a relative change variable is often used to estimate an effect size for treatment progress.

The problem with this common percent-change measure is that most data for which it is computed are constrained at the low end by zero or some other minimum constraint below which values are inherently not possible (e.g., less than zero treatment hours or a negative amount of drug found in urine samples). This would not be a problem if values rarely approached the minimum (i.e., the mean would be at least three or four standard deviations above this constraint), but in most types of clinical research, this minimum value constraint is often observed with some frequency so that the distribution of values is bunched up against the minimum, as shown in figure 1.

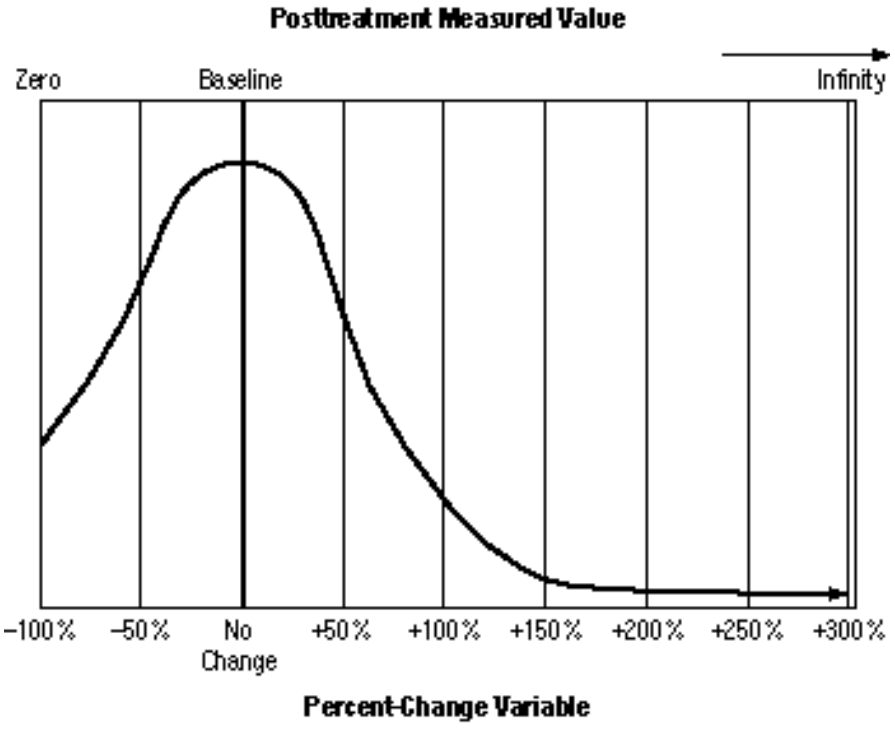


FIGURE 1. *Distribution of potential percent-change scores*

The result is that the *theoretical* universe of all *possible* values for the change variable, with a theoretical median of zero percent-change, suddenly disappears at –100 percent, chopping off the tail of what should be a normal distribution of values. (A 101-percent decrease for such scales is inherently impossible.) On the other end of the scale, the distribution has a normal tail approaching a theoretical limit of +infinity-percent increase. The observed values disappear well short of infinity for *empirical* reasons (i.e., sampling and/or measurement sensitivity), but this upper measurement constraint is usually well above the median, which allows the sampling distribution to have an upper tail that is virtually normal. Depending on what is being measured and the sensitivity of the instrument, the upper end of the percent-change distribution allows for increases as large as 500 or even 5 million percent! This asymmetry violates the assumption of normality in a serious manner; the irregular *potential* distribution is not the result of empirical problems (i.e., a sampling anomaly) but is inherent in the computation. In any theoretical sampling from the universe of all clinical trials of pre-post measures, researchers who expect to see a pre-post reduction have the deck stacked against them if they compute change as a percentage difference from baseline.

Log-Change Ratio

The solution to this situation is obtained through a slight modification in the calculation of the "change" variable. Similar in principle to log-likelihood ratio statistics, the log-change ratio (LCR) is obtained by computing the ratio of the postmeasure to the premeasure, then taking the log of this ratio. The theoretical distribution (with a null hypothesis of no change) of the resulting variable is perfectly normal and symmetrical about zero, which represents no change (i.e., when both premeasures and postmeasures are equal). Taking the log of the ratio also prevents creation of extreme values that result from a spuriously low value (baseline or postvalue) and allows a more complete spread when postvalues represent marginal change.

Although its qualities make the LCR more robust and conceptually meaningful than percent-change, its metric is not readily interpretable in a descriptive sense. For description, one can work backward from the LCR value to reconstruct an example of "representative change" using the mean basal value and the LCR-expected posttreatment value. The authors find the LCR to be best suited for reporting effect sizes (such as *d*, difference divided by standard deviation) because no change equates to an LCR of zero and proportional change is symmetrical about zero; simply dividing the LCR by its standard deviation produces the effect size.

However, the LCR is not without its bugs; values of zero for either measure produce undefined results, so zeros must be recoded to some nonzero value.

The most conservative method is to recode zeros to the lowest nonzero value observed in the sample, with more liberal procedures recoding zeros to some point, such as halfway between the empirical nonzero minimum and zero. Although choice of this recoding point for zeros is arbitrary, the effect on the *potential* distribution is symmetrical regardless of the value chosen because the zero value could occur in either the premeasure or postmeasure. On the other hand, the effect on the empirical distribution can be profoundly affected, depending on the recode value chosen, and the researcher should report explicitly the value to which zeros are recoded. The authors recommend that the lowest nonzero value observed or measured be used because it is (1) the most conservative option and (2) less arbitrary (empirically determined). The latter also implies that in a truly constrained distribution this "next-lowest" value should theoretically correspond to the smallest detectable increment above zero, assuming a fairly large sample. Other procedures for empirically deriving the appropriate minimum (based on various percentile ranges within the sample distribution) have been suggested but not yet fully evaluated.

The LCR shares the same advantage of "relativeness" with the percent-change measure in that the magnitude of change is considered relative to the baseline or premeasure value. The symmetry about the expected mean of zero change (and ensuing normality in the potential distribution) of the LCR is a desirable feature and a definite advantage over percent-change. For example, calculated by the percent-change method, a change from 40 to 200 ng would produce a score of +400 percent, whereas the same change in the other direction (from 200 to 40 ng) produces a score of –80 percent. In this case the LCR would produce values of the same magnitude for both the increase and decrease (+1.609, –1.609), differing only in sign.

DATA ERROR

Up to this point the discussion of data deficiencies has centered on violations of the normality assumptions, which can and do occur in perfectly good and accurate data. Such violations also occur because of data error, which can be of many types and can creep into the data at many stages in the research process. Whether measurement, sampling, or data entry error, all error, regardless of source, should be minimized. In research where sample sizes are limited (e.g., perinatal research), a single substantial error or a handful of small errors may completely negate what otherwise might have been a strong finding. Cohen (1990) cites a study of 25 height-weight pairs in which reversal of data for a *single* subject would have changed the 0.83 correlation to a –0.26 correlation! The authors recommend always performing random spot-checks of data already entered and/or duplicating the data entry for small segments of

the data set to verify the integrity of the data entry. The rate of data entry error can be estimated by drawing periodic random samples for verification.

Outliers

Other sources of data error include measurement instrument ("sensitivity") error and sampling error. These are inherent in any data set and cannot be quantified or known in any practical sense. (If they could be identified and quantified, probabilistic methods would not be necessary.) Because they usually can be assumed to be normally distributed, they are "ignored" as the error term in the statistical model. However, remote "one-in-a-million" cases exist in the near-infinite populations that are sampled, and these "data spikes" inevitably pop up at times. The long-run expectation of finding such cases does not change the fact that they are significantly overrepresented in a sample of only 100. Therefore, when checking variables for normality in the preanalysis, it is important to check for extreme outliers as well.

The definition of the normal distribution stipulates the existence of tails, but given a sample size, the researcher can determine the probability of any z-score being a chance occurrence. For example, z-scores greater than 3 should occur only about three times per 1,000 cases, and z-scores greater than 4 should occur only about six to seven times in a sample of 100,000 cases. Unless working with large data sets (more than 10,000 cases), it is always worthwhile to investigate cases in which a value has a corresponding z-score greater than 3 and certainly always when the z-score exceeds 4. The value may turn out to be a data entry error just waiting to ruin a correlation, or it may belong to a case/subject that on further investigation is found to be so bizarre and unrepresentative of the population being studied as to warrant its exclusion as an anomalous case.

Simply excluding a case on the basis of one extreme value is a highly questionable (and arguably unethical) procedure; the decision to exclude a case from the data set should be based on some other *objective* basis, which may result in the exclusion of one or more "good" cases also. When any subject has a preponderance of extreme values on a number of measures (variables), a brief case study/descriptive explanation should be included in the research report to justify the basis for a "specific" (nonobjective) exclusion. This is not to say exclusions are not sometimes justified and proper. The authors recommend that, when the researcher locates a questionable case so extreme as to radically influence the outcome of an analysis, the report include the results of *both* analyses, including and excluding the questionable case, as well as the bases on which the case was deemed questionable. If the case is truly bad, producing a spurious or erroneous result, and the researcher fails to exclude it, the

scientific integrity of the research is jeopardized just as much as if the case is excluded simply to boost a correlation. If the results do not change much either way, presumption must rest with leaving the questionable data in.

The previous discussion of extreme cases refers only to univariate extremeness. It is possible that a case may be an extreme anomaly worthy of exclusion, but only in terms of its multivariate distribution. Such cases often cannot be identified as extreme on any single dimension. Perhaps the simplest way to locate the existence of such cases is to perform a cluster analysis specifying a small number of clusters (e.g., three to five). If any of these clusters comprises single cases (or relatively few cases compared with the other clusters), there is a possibility that these cases are multidimensional extremes. The researcher can follow up on such indications by computing for each case the mean pairwise distance score. Truly extreme cases should stand out. To avoid having the different metrics of variables influence the distance scores, it is usually preferable to first standardize the variables to be used in the clustering procedure. Also, highly correlated variables can be thought of as indicators of the same latent dimension, which then would contribute disproportionately more to the cluster solution if these somewhat redundant variables are all included in the analysis. The researcher can select the most appropriate variable from such a group or combine several through principal components analysis to keep them from exerting undue influence in defining the multidimensional space of the data set.

Other Errors

Data entry errors occurring within normal ranges of values may be undetectable when one looks at the univariate distributions yet still be different enough from their true values to greatly distort a relationship with another variable. These data errors are perhaps the most pernicious of all. For this reason, it is advisable to use a statistical package that can produce residual diagnostics, particularly lists of cases that have great influence (or leverage) on the outcome of the analysis. Some of the best measures to check are Cook's distance, Mahalanobis' distance, and deleted studentized residuals (Belsley et al. 1980; Cook 1977). Again, cases with unusual influence should be checked for data entry errors or on some other *general and objective* basis that would disqualify them.

One other source of error that can inflate the residual error is related to independence of sequential observations/measurements. It is useful to have variables that indicate sequence of measurement or collection in the data set so that sequence can be plotted against both the dependent variable and the residuals from any analysis. Sequential error can occur as the result of instrument calibration issues (such as warmup or drift),

measurer/collector change (e.g., interviewer fatigue/boredom, interviewer's becoming more proficient in measuring or using the instrument), or even from subject "learning" (such as becoming more proficient at tasks that are performed repeatedly). Such sources of sequential error can sometimes be explicitly modeled into an analysis (e.g., as a covariate in an ANOVA), thereby reducing the residual error and strengthening the estimate of the magnitude of the effect size being tested.

MISSING DATA

One type of instrument sensitivity error mentioned above is the failure to register a value at all, such as when a mechanical instrument fails to sense a stimulus or when a survey respondent fails to provide a response. In perinatal research, missing data can be common because of uncooperative subjects, scheduling difficulties, and access restrictions. When working with many variables in multiple-variable or multivariate analyses, the researcher may find that the usual default procedure of excluding cases with a missing value on any one of the many variables in the analysis (often called listwise deletion) will eliminate nearly all cases in the sample. Even if nearly every variable has valid data for 98 percent of the cases, if the 2 percent missing data are different for each of 50 variables, then the sample N can disappear completely. Whereas in most data sets missing data tend to congregate within the same variables (difficult questions or bad instruments) and/or within the same cases (e.g., problem respondents), a data set with many variables can have enough of a scattering of missing data to render an analysis fairly unreliable when excluding cases with missing data listwise.

There are at least three methods of dealing with missing data without having to exclude cases altogether. However, the researcher first must determine whether the missing data have a story to tell, because missing data can be evidence of a relationship. If missing data patterns are correlated with other variables, the researcher may have the serious problem of *nonresponse bias*, which can render any statistical inference invalid and unrepresentative of the entire sample. In cases where variables have missing values for more than a handful of cases, it is important for the researcher to test for nonresponse bias. The easiest way is to divide the sample into two subsamples: those with valid data for a particular variable and those without. Then the researcher can perform a simple difference of means test (such as the two-independent sample *t*-test) on other important variables for which both groups have most cases with valid data. Any significant difference implies that those cases with missing values may be different from the rest of the sample and that their actual values on the missing variable are possibly unrepresentative and cannot be "corrected"

by substitution or imputation methods with much reliability. Such patterns of nonresponse can be significant findings.

Mean Substitution

If the researcher decides to salvage cases with missing data without losing the good data, the most common method is to employ a mean substitution strategy, in which the missing data for a predictor variable receive the value of the mean for all those cases with a valid datum for that variable. This is a good method to use in situations where an analysis includes many "incomplete" variables, each of which may be missing values for only a very few cases, but the values are scattered among *different* cases for each variable. On the other hand, if the variables tend to have their missing values concentrated within the same few cases, then the researcher is probably better off excluding those cases. When the missing data are scattered among the cases, mean substitution can keep a fairly complete data set from losing a substantial portion of its cases in multivariate analyses, with relatively little risk of introducing a falsely positive bias.

The problem with the mean substitution method is that it usually flattens true correlations by reducing variance, erring on the side of being conservative (reducing magnitude of the effect). Because this is perhaps the easiest and least risky way to salvage a disappearing N, many software packages offer mean substitution as an option for multivariate procedures. As mentioned above, any indication that nonresponse patterns are significant makes mean substitution a highly questionable procedure. In any respect, mean substitution should be undertaken *only for independent (predictor) variables* in an analysis, and care should be taken to exclude cases that have missing values for the dependent variable, even when mean-substituting for independent variables.

Imputation

A second method for dealing with missing data is to impute, or predict, values for these cases based on other variables—preferably from variables *that will not be used* in the analysis (called external variables here). Much has been written on the topic of data imputation, and it is advisable to examine some of the literature if this path is chosen, as there are many complex issues (and procedures) dealing with imputation (e.g., Little and Rubin 1987). Imputation is perhaps most advantageous when a small number of important predictor variables have missing values and nonmissing values of these important variables can be reliably predicted from a combination of two or more external variables. If such other variables exist that are successful in predicting values of the "incomplete"

variable in question, the multiple regression or ANOVA model used for predicting the valid cases can be used to assign predicted values for those cases with missing values. The danger here is that the researcher may spuriously inflate the findings of the statistical tests by reinforcing the relationships used to impute the values—essentially perpetuating a result by increasing the N. For this reason, missing values for the independent variable should *always* be imputed from variables external to the analysis, although these external variables may be used in other unrelated analyses that do not involve the imputed variable.

As with mean substitution, it is arguably not appropriate ever to impute values for *dependent* variables. When imputing values for predictor variables, the dependent variable must *never* be used, and external variables also should be avoided if they have even moderate bivariate correlations with the dependent variable. One rule of thumb is that no external variable contributing to the imputation should have a bivariate correlation (Pearson's *r* [Snedecor and Cochran 1980]) with the dependent variable greater than half the multiple *r* of the imputation equation. This way the researcher is less likely to fall into the trap of circular logic. Although imputation can be a beneficial tactic for dealing with missing data, the dangers of using it improperly are great enough to warrant caution. When it is used, it is a good idea to weight cases with imputed data somewhat less than cases with all valid data. One good rule is to weight the case by the multiple *r* (or, more conservatively, the r^2) of the equation used to impute values for the missing variable, which means that the less successful the imputation formula, the less weight is given the cases relying on the imputed data. One implication of this policy is that cases utilizing imputed data carry case weights worth a fractional case, which often results in interesting sample Ns (e.g., N=121.37). The researcher should report the weighting scheme used to protect against suspicion of posthoc "data adjusting."

In situations where more than one variable in the analysis carries imputed data and some cases have imputed data for different variables than do other cases, case weights depend on which variable(s) a case is missing. Cases that carry imputed data for more than one variable should be weighted by the product of all the multiple *r*s (of the imputation equations for these imputed variables) so that the more missing data elements a case has, the closer to zero that case's weight becomes.

Multiple-Donor Matching

A third method for dealing with missing data is a variation of "donor-recipient" matching. This technique usually produces results similar to that of the multiple-regression imputation. The idea behind

donor-recipient matching is to find the case or cases that are most similar to a case with missing data in terms of other important variables for which the "recipient" case has valid data. The matching "donor" case then contributes its score on the variable in question to the case that is missing a value for that variable. This can be done through a clustering procedure, which determines distance between cases in terms of multidimensional space or, inversely speaking, determines the most similar or closest "neighbor" in terms of these other nonmissing attributes. Because many such clustering procedures are affected by the particular metric scaling of each variable, the clustering should be based on either standardized scores (i.e., z-scores) or on ranks for the nonmissing variables. Also, having several variables that are similar and highly correlated among themselves can overemphasize the importance of these variables in determining the distance score. For this reason, some researchers prefer to extract principal components (factors) from among the set of variables and then to cluster the cases based on factor scores.

Some researchers propose a strict one-to-one donor-recipient match, but the authors recommend instead that the recipient case receive a mean value for the group of closest neighbors. Determining how many neighbors to use as donors can be based on a criterion that specifies the average size of clusters. For example, if one wishes the average cluster size to be approximately 5 percent of the total sample, a cluster solution that produces 20 clusters would be specified. Some clusters contain more than 5 percent of the cases and some less, depending on the distribution of multidimensional similarities. The recipient case then would have its missing value replaced by the mean of that variable for the other cases in the recipient case's cluster. The criterion to use for specifying cluster size should depend on sample size. The authors' somewhat arbitrary recommendation is that samples with fewer than 100 cases should average 5 cases per cluster; samples of 100 to 500 should average 5 percent of the sample per cluster; and samples with more than 500 cases should produce clusters averaging about 25 cases.

This type of donor-recipient matching can be rather time consuming. Although the end result will often be similar to the result of multiple-regression imputation described earlier, donor-recipient matching can provide more reliable estimates of the missing data when relationships among the various dimensions are nonlinear or nonadditive or have specific ranges in which residuals from multiple regression are not randomly distributed. Finally, as with the multiple-regression imputation technique, recipient cases should be weighted to some fraction less than a full, single case. One possibility for determining the weight for such cases involves the ratio of the cluster's standard deviation of the variable to be donated to the standard deviation of this variable for the total sample. For example,

if the cluster standard deviation is 3.57 and the total sample standard deviation is 16.22, then the weight for recipient cases in this cluster would be 1–(3.57/16.22), or 0.78. In this way, the more homogeneous the cluster, the closer the weight approaches that of a full case.

NONLINEAR RELATIONSHIPS

One fact of scientific life is that, for one reason or another, many true relationships are not strictly linear.[1] Although sophisticated procedures exist to test models that are more complex than a simple linear frame of reference, such as nonlinear regressions, dynamic/interactive systems modeling, and qualitatively interactive segmentation or "tree-splitting" procedures, these nonlinear/nonstatic/nonadditive models are beyond the scope of this chapter. Nevertheless, many times a relationship that is not strictly linear may be *intrinsically* linear through the use of a linear transformation, such as those discussed above in the section titled "Nonnormality." Many times a simple linear regression will produce a significant result, yet the true result may be a stronger relationship if the function is curvilinear, such as a log function relationship (as in figure 2). Failure to model the relationship as its actual log function or polynomial function (achieved by transforming some or all the variables in the analysis or adding square and cube components) results in underestimating the magnitude of the relationship—sometimes concluding incorrectly that there is no relationship, such as would result when the relationship resembles a basic parabolic function as shown in figure 3.

Thinking Nonlinearly

There are two ways to guard against mistakenly missing or underestimating a relationship. The first is to be more imaginative in conceptualizing the nature of the phenomenon and break away from the constraints of the strictly linear frame of reference. To do so means the researcher would have various transformations of variables available for evaluation by a regression procedure's stepwise selection method of including variables. If the researcher can conceptualize the nature of the phenomenon more accurately, he or she is more likely to provide the analytic procedure with

---

[1] The term "strictly linear" is used here to denote relationships that fit a straight line. Curvilinear functions, such as log, inverse, and polynomial, are also linear in the truest mathematical sense but are called "intrinsically linear" here to differentiate them from the functions that can be properly fit using general linear regression methods *without* transforming the data. For a more detailed explanation, see the introduction to the nonlinear regression procedure in *SPSS/PC+ Manual (Version 5.0), Advanced Statistics* (Norusis 1992, pp. 231-233).

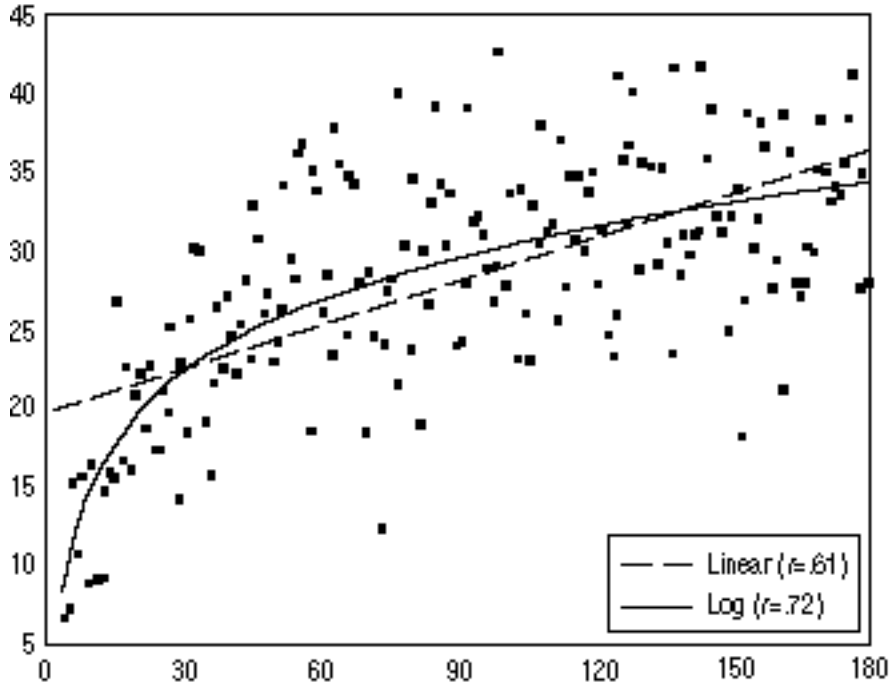


FIGURE 2. *Comparative fits of linear and log models to data generated by a quadratic function*

NOTE: Units are hypothetical; data were generated by simulations.

KEY: r=correlation coefficient

the appropriate transformed variables. In this way analyzing data becomes more than just feeding the computer a bunch of variables; the researcher must spend some time conceptualizing the model in abstract mathematical terms. Many phenomena exhibit relationships that, rather than simply being linear, are more properly conceptualized as "saturation," ceiling, or leveling-off models (log transforms); sudden "critical mass" functions that suddenly take off (exponential); transfer functions that combine both of these (sigmoidal or "S-curve"); and nonmonotonic curve functions that have reversals, such as the relationship between vitamin A dosage and healthy physiological functioning or between information input and performance/decisionmaking, where "too much of a good thing" becomes detrimental after a certain point (often modeled as simple polynomial functions, as shown in figure 3).

For example, in a recent study of cocaine use by mothers referred to treatment, the authors obtained hair measures of cocaine use every 4 months after the pretreatment baseline measure. Using a repeated

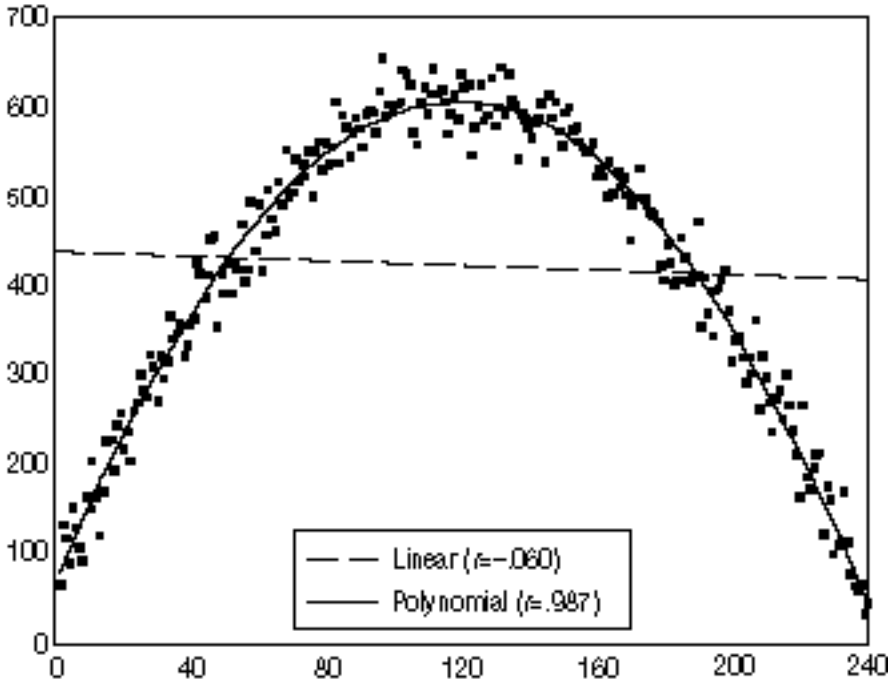


FIGURE 3. Comparative fits of linear and polynomial models to data generated by a quadratic function

NOTE: Units are hypothetical; data were generated by simulations.

KEY: r=correlation coefficient

measures MANOVA, the authors expected the contrast between consecutive measures to decrease over time for at least two reasons. First, prior research and common sense led to the belief that if a client were to improve at all over 2 years, much or even most of the improvement would be seen in the first 4 months of treatment. Second, there is a ceiling effect on improvement (or from another perspective, a floor effect on drug use because drug use cannot drop below zero). For these two reasons the authors expected plots of drug use over 2 years to drop suddenly at the beginning and then level off toward some asymptote, as in an inverted logarithmic function. By specifying a second-order polynomial contrast among the repeated measures rather than strictly linear or simple difference contrasts, the true relationship between treatment over time and cocaine use for the group became clear.

Examining Residuals

The second way to guard against passing over relationships is to spend time looking at residual plots to check for violations of the assumption of randomly distributed and uncorrelated error. Such violations do not necessarily mean that the data are not amenable to parametric analysis but often indicate that the relationship can be better discerned through a different conceptual model (such as those previously described). Regularly performing residual diagnostics is something everyone is (or certainly should be) taught in the first course on regression, but few busy researchers feel they have time for such tedium after the results are in. One of the easiest checks is to look at normal probability plots of the residual for deviations from a straight diagonal line. Because most software packages can generate these plots automatically on request, there is no reason why the researcher should not *always* check them. Second, one should examine plots of the raw and studentized residuals against predicted values of the dependent variable and against each of the independent variables. A glance at each is all it takes to check that there are no patterns (i.e., that these scatterplots look like random shotgun blasts) and that the spread of residuals does not increase as the *predicted* values for the dependent variable increase (which indicates that the homogeneity of variance assumption is violated). Most of the better statistical packages allow the researcher to request that the software automatically generate these plots as part of the regression or ANOVA procedure.

When multiple regression is used, it can be difficult to see from those residual plots where transformation of a single variable would improve the model. Although somewhat more time consuming (because many software packages do not have the capability to automatically generate the output), one of the best regression diagnostics is the *partial regression plot*, which removes the variance of the dependent variable that is explained by the other predictor variables, thereby making nonlinear relationships easier to see. The partial regression plot consists of plotting for each independent variable *j*, the residuals when the dependent variable is predicted from all other independent variables *except j*, against the residuals produced when independent variable *j* is predicted from the other independent variables. Such plots should produce a straight line pattern of data points; nonlinear relationships show up as curvilinear patterns. Such patterns can help the researcher reassess the proper conceptualization of the model and locate variables that need linear transformation. Just as with the preanalysis diagnostics of univariate distributions discussed at the beginning of this chapter, checking residuals for the assumption of independent and random error to ensure that the procedure is valid may not be necessary given the robustness of these parametric procedures; often the greatest value

in doing so is to optimize the results of analyses that might otherwise be biased in the direction of a Type II error, that is, underestimating or finding no relationship.

CONCLUSION

Although some topics mentioned here merit greater detail (especially missing data imputation and nonresponse bias), the authors have presented some new techniques that others may find useful in dealing with difficult or problematic data sets and have emphasized issues that may motivate reassessment of data sets and analyses that were assumed to be valid. Finally, in fairness to the data, perhaps the perspective of this chapter's title should be rephrased: The most common data deficiencies result from the researcher's imperfect methods of scaling, measuring, and collecting data as well as from the use of often overly simplistic perspectives in modeling and diagnosing relationships among them. Thus, researchers need to "minimize their deficiencies in dealing with data."

REFERENCES

Belsley, D.A.; Kuh, E.; and Welsch, R.E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley, 1980.

Chronbach, L.J., and Furby, L. How should we measure "change"—or should we? *Psychol Bull* 74:168-180, 1970.

Cohen, J. Things I have learned (so far). *Am Psychol* 45(12):1304-1312, 1990.

Cook, R.D. Detection of influential observations in linear regression. *Technometrics* 19:15-18, 1977.

Harris, C.W., ed. *Problems in Measuring Change.* Madison, WI: University of Wisconsin, 1963.

Little, R.J.A., and Rubin, D.B. *Statistical Analysis With Missing Data.* New York: Wiley, 1987.

Marques, P.M.; Tippetts, A.S.; and Branch, D.G. Cocaine in the hair of mother-infant pairs: Quantitative analysis and correlations with urine measures and self-report. *Am J Drug Alcohol Abuse* 19(2):159-175, 1993.

Norusis, M. *SPSS/PC+ Manual (Version 5.0), Advanced Statistics*. Chicago: SPSS Inc., 1992.

Snedecor, G.W., and Cochran, W.G. Correlation. In: Snedecor, G.W., and Cochran, W.G., eds. *Statistical Methods.* Ames, IA: Iowa State University Press, 1980.

AUTHORS

A. Scott Tippetts
Chief Statistician
(301) 731-9891, ext. 113 (Tel)
(301) 731-6649 (Fax)
tippetts@pire.org (Internet)

Paul R. Marques, Ph.D.
Senior Research Scientist/Director of Drug Research
(301) 731-9891, ext. 102 (Tel)
(301) 731-6649 (Fax)
marques@pire.org (Internet)

National Public Services Research Institute
Suite 220
8201 Corporate Drive
Landover, MD 20785

Analysis and Use of Qualitative Data

Souraya Sidani and Lee Sechrest

INTRODUCTION

Perinatal drug use is a topic of obvious importance to society in general and to clinicians specifically. Despite its importance, myriad questions exist about it, in part because of a relative dearth of research and in part because of the uncertain nature of the information that has been obtained to date.

Questions indicate gaps in knowledge about a phenomenon; "they point either to problematic phenomena, observed events that are puzzling in terms of currently accepted ideas, or problematic theories, current ideas that are challenged by new hypotheses" (Brewer and Hunter 1989, p. 55). Lack of adequate knowledge creates ambiguity or uncertainty that limits understanding of what is happening and consequently restricts the ability to intervene or help those in need. Inquiry is conducted to answer questions and reduce uncertainty. Thus, the intent of inquiry is to produce information relevant to a particular question or uncertainty. Information is contained in any message that reduces uncertainty about a phenomenon of interest (Shannon and Weaver 1949). That is, information is not merely a fact such as an observation or a significant finding; it subsumes the ability to give structure and meaning to the fact, in relation to the phenomenon under examination. For example, treatment of substance abuse is a formidable task, and little if any information is provided by the failure of yet another program unless reasonably precise reasons for the failure can be delineated.

NEED FOR METHODOLOGICAL DIVERSITY

When questions arise, different approaches to resolving them are available. These approaches can be generally classified as empirical and nonempirical. Empirical approaches involve a systematic process of data collection, analysis, and interpretation (Sechrest and Sidani, in press). Data refer to information gathered during an inquiry. These data are then summarized, weighted, or analyzed in a way that allows valid conclusions to be drawn that provide logical explanations of the identified problem, answer questions, and reduce uncertainty. Nonempirical approaches include reversion to authority, revelation, intuition, and logical reasoning.

The extent to which uncertainty is *legitimately* reduced depends on the validity of conclusions. In turn, ascertaining the validity of conclusions

is a difficult, even "uncertain," task (Brewer and Hunter 1989). The uncertainty regarding interpretation of findings stems from the inability to rule out alternative interpretations. Critical rival hypotheses usually cannot be tested when data are collected from only one source, using one method of data collection and reflecting a single perspective. In contrast, the likely validity of conclusions is enhanced when multiple methods of data collection and analysis are used and when alternative theoretical perspectives are represented in an empirical inquiry (Cook and Campbell 1979; Shadish 1993, pp. 13-57; Webb et al. 1981).

Empirical approaches to inquiry include a variety of research designs that often are divided into two broad categories: quantitative and qualitative. Quantitative and qualitative methodologies are likely to differ on several grounds: theoretical perspective, research purpose and design, and methods for data collection and analysis. Some authors claim that the two methodologies cannot be mixed because they are founded on incommensurable paradigms (Denzin and Lincoln 1994, pp. 1-17; Guba and Lincoln 1988, pp. 89-115), but triangulation of research strategies should be sought when an inquiry is conducted; the differences in perspective, data collection, and analysis require that multiple research designs be used when a phenomenon of interest is examined. Information obtained from different sources provides a complementary and comprehensive explanation of what is occurring, thus increasing confidence in the validity of conclusions and reducing uncertainty. However, little explicit attention is paid to the exact methods by which information obtained from diverse sources may be integrated (Dennis et al. 1995).

ADVANTAGES OF TRIANGULATION

Triangulation refers to the use of a combination of methodologies in the study of the same phenomenon (Jick 1979). Triangulation may involve multiple investigators, multiple theories, multiple data collections, multiple measures, multiple analysis methods, or any combination thereof. The assumption underlying this pluralism is that each type of research methodology, whether quantitative or qualitative, has an inherent bias or weakness; however, the biases or weaknesses of research methods are not identical. Consequently, when several methods are used conjointly, the bias inherent in one method may be canceled out by the bias introduced by another. That is, the weaknesses or limitations of one method may be compensated for by the strengths of the other. This balancing at least will be likely if methods or approaches are chosen because they do not share the same biases or weaknesses. For example, when a community-based drug rehabilitation program is being evaluated, teenagers may be asked

to self-report weekly on their drug use, and their friends also may be asked to report on their friends' drug use. The self-reports may be biased toward underreporting; friends' reports should not be biased to the same degree. Still, both groups' reports may be biased toward underreporting compared with results of a urine test. Thus, one advantage of triangulation is reduction of bias. However, elimination of bias requires careful analysis of the sources of error in different methods.

Furthermore, with triangulation, the strengths of one method add to or even enhance the strengths of the other, thus providing complementary information that gives a more complete and comprehensive explanation or picture of the phenomenon under study. For instance, a urine test can determine only whether drugs have been used within a certain timeframe; it reveals nothing about patterns or circumstances of drug use. Other methods of inquiry, including self-report and reports by friends, may be of value in the latter respect.

Thus, a second advantage of triangulation is improvement of the validity of findings. When findings of multiple methods converge, confidence in the validity of conclusions is enhanced; results in agreement indicate that the methods are assessing the same phenomenon and that the unique bias of each method is reduced, if not eliminated. The plausibility of rival explanations of the phenomenon also is diminished because various perspectives are represented in all stages of the research process.

Findings from multiple methods that sharply diverge are as important to the understanding of a phenomenon as convergent ones. Contradictory results provide the researcher with an opportunity to uncover deviant or unexplored dimensions of a phenomenon; identify substantive or methodological sources of divergence, thus enriching the understanding of the phenomenon and potentially leading to synthesis or integration of theories explaining it; and rethink and reanalyze the problem from a new perspective (Brewer and Hunter 1989; Jick 1979; Mathison 1988). Discrepant findings about the effects of an intervention program, such as those that may arise from comparing objective measures with information obtained from clinical interviews, should not be a source of despair but should be regarded as both challenging and useful in defining issues more precisely.

Although advantageous, triangulation should not be applied mindlessly (Patton 1988, pp. 116-137). That is, the decision to use multiple methods should be based on the nature of the problem under study and the availability of measures sensitive to the problem. Planning for collecting data from various sources using different methods is one aspect of triangulation that can be done by researchers to draw accurate

and valid conclusions. Examples of studies in which data and method triangulation were implemented are increasing in number (Breitmayer et al. 1993; Connides 1983; Dennis et al. 1995; Helitzer-Allen and Kendall 1992; Roter and Frankel 1992; Steckler 1989). Both quantitative and qualitative methods of data collection and analysis frequently have been used, thus providing additional empirical support for the complementary nature of the two methodologies and the benefits of data and method triangulation. Using alternative methods for measuring or capturing a phenomenon is useful in identifying the direction of bias, estimating its extent, validating or verifying results, and providing complementary and complete information about the phenomenon (Fielding and Fielding 1986).

VALUE OF QUALITATIVE METHODOLOGY

As stated above, overreliance on only one type of research method may lead to systematic bias in results and therefore threaten the validity of conclusions. Evaluation of drug use prevention, treatment, and rehabilitation programs has relied primarily on quantitative research methods. Although some questions about the effectiveness of such programs have been answered, the uncertainty regarding the validity of conclusions has not been completely resolved because of the potential bias inherent in quantitative research. For example, results of a structured questionnaire administered to subjects who participated in a drug rehabilitation program may indicate that the program was effective in reducing drug use, improving well-being, and helping participants adjust to the demands of the role of teenage mother. Nevertheless, social scientists, skeptical by nature of their training, question the validity of the success of the program. At least two factors might have contributed to observed findings. One factor relates to the data collection method used: self-report. The validity of self-reports is not always maintained: Honesty of participants, social desirability, acquiescence, and clarity of directions, questions, and response options are issues that must be taken into consideration when interpreting results obtained from a structured questionnaire. A second factor relates to the participants' knowledge that they are being studied and their willingness or desire to please the investigator or clinician. In addition, the mechanisms underlying the observed changes in the participants' behaviors might not have been precisely determined; factors other than those theoretically expected might have been responsible and might have led to a "turning point" in participants' lives.

Qualitative methodology offers an alternative to and a complementary perspective of problems addressed in perinatal drug research. Qualitative research is described as "multimethod in focus, involving an integrative,

naturalistic approach to its subject matter. This means that qualitative researchers study things in their natural settings, attempting to make sense of or interpret phenomena in terms of the meanings people bring to them" (Denzin and Lincoln 1994, p. 2). Qualitative methodology differs from quantitative methodology on several grounds, including ontological and epistemological assumptions and subsequent methodological strategies (Bednarz 1985; Bryman 1984; Denzin and Lincoln 1994, pp. 1-17). Although there is not adequate support for some of these distinctions (Sechrest and Sidani, in press), the authors provide a summary of the essential features that characterize qualitative methods because, as discussed in the literature, qualitative data collection and analysis are ingrained and connected to the philosophical assumptions held by researchers.

In general, qualitative research is based on a relativistic, constructivist ontology that posits that there is no objective reality. Rather, there are multiple realities constructed by human beings who experience a phenomenon of interest:

> People impose order on the world perceived in an effort to construct meaning; meaning lies in cognition not in elements external to us; information impinging on our cognitive systems is screened, translated, altered, perhaps rejected by the knowledge that already exists in that system; the resulting knowledge is idiosyncratic and is purposefully and effortfully constructed (Lythcott and Duschl 1990, p. 458).

These constructed realities exist in the minds of the individuals and cannot be broken into parts but must be examined as a whole. Individuals are viewed as open systems, engaged in continual dynamic interactions with their environments, which results in an ongoing evolvement and refinement of their perceptions of reality. To capture these subjective realities that are continually constructed, the "knower" or researcher interacts with the "known" or participants for the purpose of gaining an understanding of the participants' points of view, also called the insider's perspective. Because the insider's perspective is assumed to be evolving, the inquirer aims at capturing this dynamic process that occurs at a specific time and in a specific context. Therefore, many qualitative methodologists assert that investigations should be conducted under natural conditions (Guba and Lincoln 1988, pp. 89-115).

These epistemological and ontological assumptions are translated into distinct methodological strategies. The goal of a qualitative investigation is to understand the complex world of human experience and behavior from the point of view of those involved in the situation of interest. Therefore,

the investigator is expected not to have an a priori, well-delineated conceptualization of the phenomenon; rather, this conceptualization is to emerge from the interaction between participants and investigator. Flexibility in design, data collection, and analysis of research is strongly recommended to gain "deep" understanding and valid representation of the participants' viewpoints. Data analysis is interpretive, guided by the inquirer's insights or intuition in identifying intersubjective, common meanings or regularities—the patterns of observed events. The outcome of this analysis is a rich, often called thick, description and formulation of "working" hypotheses regarding the examined phenomenon (Guba and Lincoln 1988, pp. 89-115; Hughes 1992).

In summary, qualitative methodologies are founded on a relativistic ontology, an interpretive or constructivist epistemology, naturalistic or interpretive methods, and inductive reasoning. Their methodological flexibility of qualitative approaches is valuable when exploring new problems, when observing unanticipated variations in a plan and results, and when evaluating developing programs, that is, at the stage of formative evaluation (Fetterman 1988; Woodhouse and Livingood 1991).

QUALITATIVE DATA ANALYSIS

Despite the commonly held perspective of multiple realities and interpretive, naturalistic methods, qualitative researchers differ in their data analysis techniques. Qualitative research methods include ethnography, grounded theory, phenomenology, feminist research, and critical social theory research. However, feminist and critical social theory research are ideologies more than methodologies in that they have a preconceived sociopolitical agenda, that of raising the "consciousness" of participants toward their current status for the purpose of improving it. As such, these ideologies defy the goal of qualitative methodologies—that of understanding human phenomena from the insider's viewpoint. Consequently, this chapter discusses methods of data analysis that are pertinent to ethnography, grounded theory, and phenomenology only.

Ethnography

Ethnography is defined as "an analytic description of the behaviors that characterize and distinguish cultures or sociocultural groups" (Walters 1980, p. 16). The emphasis is on learning about aspects of life, such as norms, values, and beliefs held by a group of people, and about the nature of behavior and social phenomena observed among members of this group. To achieve a comprehensive, holistic understanding of the sociocultural

characteristics of the study group, the ethnographer "steps in and out of society" (Walters 1980, p. 16): The ethnographer becomes immersed as fully as possible in the participants' world, observing the participants in their natural environment; actively participating, in an acceptable role, in their activities; and collecting verbal or written reports of their perceptions of events. At the same time, the ethnographer, usually an outsider, provides a personal account or interpretation of the participants' world. The ethnographer deals with data from two sources: the participants' accounts of their world and the ethnographer's interpretation of his or her observations.

Transcriptions of participants' interviews are analyzed, using an "emic" frame of analysis. That is, these data are first coded using words given by participants, thus representing their own perspectives. The coded materials are examined for similarities and differences and then sorted by topic. Codes that relate to each other are grouped into categories that illustrate the various domains or topics. Relationships in the data are sought, and a taxonomy is developed to describe the knowledge, beliefs, values, and behaviors of the cultural group under study. Regularities or patterns of observed behaviors and events that emerge from this analysis are used to clarify, extend, and interpret the field notes taken by the ethnographer from observations regarding the meanings and functions of human actions (Atkinson and Hammersley 1994, pp. 248-259).

During the process of data analysis, the ethnographer uses intuition, introspection, and reasoning in finding regularities or patterns of behavior (Burns and Grove 1987). The process of data analysis is facilitated by the ethnographer's (1) familiarity with the participants' vernacular, (2) length of interactions with participants, and (3) degree of participation in the participants' world (Walters 1980, pp. 15-20). These characteristics of an ethnographer are essential for providing a valid and reliable description of the patterns of behavior within the natural setting and from the point of view of the cultural group.

An ethnographic approach is particularly well suited to exploration of treatment milieus that, for example, might exist in a community-based program or therapeutic community. An ethnographer provides a useful account of the customs of treatment groups; the informal, unwritten rules by which they operate; the networks of social relationships and influence; and so on. Such a perspective and description are of great value in understanding how treatment gets done, why it might be more effective for some persons than for others, and how it might be improved—among other issues.

Grounded Theory

Grounded theory is defined as a "general methodology for developing theory that is grounded in data systematically gathered and analyzed. Theory evolves during actual research, and it does this through continuous interplay between analysis and data collection" (Strauss and Corbin 1994, p. 273). Grounded theory is based on symbolic interactionism (Wilson and Hutchinson 1991), an approach to the study of human conduct and a human group's life; it focuses on the meaning of events to people in natural, everyday settings. Symbolic interactionism is concerned with how people define events or reality and how they act in relation to their beliefs. In other words, individuals act toward things—objects, other human beings, institutions, activities, and situations—on the basis of the meanings the things have for them (Chenitz and Swanson 1986). These meanings are created by people from their interactions with their social world. Therefore, the main concern of grounded theory researchers is the discovery of basic social processes, patterns of action and interaction between and among individuals and groups, and the meaning of personal experiences as they are constructed within a specific social and interpersonal context. The emphasis is on understanding the processes, that is, the reciprocal changes in patterns of actions and interactions leading to the individuals' perceptions of events or a social situation. Thus, the purpose of grounded theory is the development of a theory. Strauss and Corbin (1994, pp. 273-284) define theories as interpretations of phenomena made from given perspectives: those of the participants and those of the researchers. Multiple perspectives are sought, analyzed, and incorporated into theoretical conceptualizations of phenomena, bound by time and context. Because the aim of grounded theory is to understand and explain the meaning of experience and behavior as presented by participants, the investigation should be conducted in the "natural setting" where the inquirer can observe behaviors and interactions as they occur. Glaser and Strauss (1966) explained that any observation made is quickly accompanied by "hypothesizing"; that is, the inquirer frames his or her personal account of the observation in the form of an alternative hypothesis. Such hypotheses are generated from an interaction among what is observed, an existing theory or body of knowledge, and the investigator's personal beliefs, meanings, and perceptions. Furthermore, the hypotheses guide the investigator in collecting data, such as by refocusing the topic of the interview; in analyzing data, such as the way data are coded and categorized; and in interpreting findings and integrating them into a meaningful theory.

Data are collected through interviews with participants, field observations, review of documents, or a combination of these. The procedure for sampling participants is purposive or theoretical: Participants are selected who can provide rich descriptions of the experiences under study; they

should be able to articulate their experiences and be willing to give complete and sensitive accounts (Wilson and Hutchinson 1991). In addition, when the inquirer has formulated hypotheses and is seeking their verification, he or she systematically chooses diverse cases and events that represent a wide continuum for the purpose of maximizing similarities and differences of information gained, a step necessary for verifying emerging theories. The diversity and number of participants or events to be selected are determined by the level of saturation of categories. Saturation of categories means that the investigator is not learning anything new from the available data that will substantially modify the concepts and hypotheses reached (Glaser and Strauss 1966). If saturation is high, the investigator will make a decision, based on personal judgment, either to terminate a line of inquiry or to seek a more diverse sample of participants. This characteristic of grounded theory implies that the processes of data collection and data analysis are conducted simultaneously, with one feeding back into the other.

When data are analyzed in grounded theory, the emphasis is on conceptualization rather than mere description. Conceptualization involves some level of abstraction from the observed, the concrete, for defining concepts and relating them in a meaningful way to explain the phenomena under study. Consequently, data analysis proceeds by a continual, reciprocal interplay between concepts and theories held by the researcher and the data provided by participants. This interplay between concepts and data is reflected in a constant making of comparisons (Strauss and Corbin 1994, pp. 273-284). More specifically, the process of data analysis in grounded theory includes:

- Open coding, meaning that labels are selected to represent significant data bits; participants' own words can be used as labels. An alternative method for coding the data is to categorize the data through "themes" that illustrate various aspects of the phenomena described in the data.

- Making notes while collecting data and reviewing interview transcriptions. These notes may be (1) observational—related to events experienced during field work, such as "who said or did what, under which circumstances" (Schatzman and Strauss 1973, pp. 100-101); (2) theoretical—"represent self-conscious, controlled attempts to derive meaning from any one or several observation notes" (Schatzman and Strauss 1973, pp. 100-101; or (3) methodological—"reflect an operational act competed or planned: an instruction to oneself, a reminder, a critique of one's own tactics" (Schatzman and Strauss 1973, pp. 100-101). These notes provide further support for the researcher's conceptualizations of the phenomenon.

- Comparisons of codes and associated data bits to identify similarities and differences, followed by clustering into categories with common ideas.

- Saturation of categories; that is, the investigator accumulates additional examples or data bits related to the category until confident of the meaning represented in specific categories and of the ability to categorize new data bits in respective categories with no hesitation. An abstract definition of the categories, in terms of the qualities or properties characteristic of the category, is helpful for later classification of new data bits.

- Developing linkages or relationships among the categories in the form of hypotheses.

- Identifying conditions under which the relationships hold; this step is facilitated through use of conditional matrices to lay out the findings.

Throughout this process, comparisons are continually made of the codes and categories emerging from the data, the notes made during data collection or analysis, and the sample characteristics (Chenitz and Swanson 1986). The result of this analysis is the development of a theory regarding the phenomena studied. The theory explains a behavior, action, or interaction among individuals, as viewed by participants. Thus, the theory is an "abstraction grounded directly or indirectly on perspectives of the diverse actors [participants and researcher] toward the phenomenon studied . . . . [It] connects this multiplicity of perspective with patterns and processes of action/interaction that in turn are linked with carefully specified conditions and consequences" (Strauss and Corbin 1994, pp. 280-281).

The authors are interested in the phenomenon referred to as the "turnaround" outcome; that is, for a wide variety of troublesome human behaviors, including chronic drug abuse, some individuals seem to be able to turn their lives around—or to have them turned around. These people have not just "improved" to some extent; they have been dramatically changed. How these turnarounds happen is unknown, as is how to make them happen. A grounded theory approach might be exceptionally useful in developing an understanding of this phenomenon.

Phenomenology

Phenomenology is a distinctive philosophy, theory, and method for studying human phenomena, with a focus on the lived experience of everyday life; the unique, personal interpretation of the experiential world; and commonsense knowledge and sense of reality (Oiler 1982). As a

philosophy and theory, phenomenology is based on a well-defined set of assumptions:

- Human beings live in the world, in a specific context that exists as an outside object. They perceive the outside world and interact in it. However, their perception of objects, events, and actions or behaviors is not passive; experience is consciously and actively constituted by individuals (Holstein and Gubrium 1994, pp. 262-271).

- Reality is, therefore, subjective, perceived and interpreted by an individual. The world becomes real through contact with it. In other words, reality is lived, subjective experience. Human beings make sense of the world they see and experience; facts are not taken for granted but are perceived, interpreted, and ascribed meanings. Human actions and interpretations of their experience are guided by a "stock of knowledge" handed down to them through language and cultural and social practices. Stock of knowledge refers to images, ideas, theories, rules or principles, and values, feelings, and attitudes that provide "resources with which a person interprets experiences, grasps the intentions and motivations of others, achieves intersubjective understandings, and coordinates actions . . . . Stocks of knowledge produce a familiar world, one with which [individuals] already seem to be acquainted" (Holstein and Gubrium 1994, p. 264), which is reflected in shared constructs and categories. Nonetheless, people's perceived realities are more important than any objective reality because people act on what they believe (Fetterman 1988).

- Truth is viewed as a composite of the individual's perceived realities. Because these realities derive from the individual's experience, they are based on the person's perceptual skills and preconceived stocks of knowledge. Perceptual skills and knowledge evolve with experience; that is, with acquired experience, perceptual skills and knowledge are refined. Consequently, truth is viewed as constantly changing. In addition, truth is context dependent; it is specific, not universal: Truth stems from interpretation of experience, which is influenced by the individual's background (Dzurec and Abraham 1993).

As a method, phenomenology embraces a holistic approach, meaning that a person is looked at as a whole that is different from the sum of his or her parts. The person's physical, psychological, social, and spiritual aspects are examined simultaneously, in addition to the spatial, temporal, and cultural contexts in which the phenomenon is or was experienced (Oiler 1982). The investigator is viewed as another person who possesses stocks of knowledge and is actively involved in the participants' uncovering of their experiences, interpretations, and meanings. The investigator becomes

part of the social context that influences the construction of participants' realities. Therefore, the uncovered meanings, truths, or realities are the product of the interaction and agreement between the participant and the researcher. Consequently, the task of the researcher is to bring to an explicit level, to describe and understand, the meaning of life or the human experience in everyday activities. However, steps in phenomenological inquiry are not clearly defined. Data collection and data analysis take place simultaneously (Burns and Grove 1987), guided by the following operations:

- *Reflecting* is a method used to attain a rich and comprehensive description of the lived experience. It is achieved through a dialectical process between the researcher and the participant; that is, data collection takes the form of a "conversation," which is an interview with a focus, but not one-sided.

- *Bracketing* refers to the act of suspending or laying aside the investigator's stocks of knowledge and bias regarding the experience being examined (Holstein and Gubrium 1994, pp. 262-271). Bracketing is used during both data collection and analysis.

- *Intuiting* requires looking at the experience from an unrestricted perspective, with an open mind, and setting aside previous knowledge, facts, and theories. It involves examining the experience with a "fresh" look, concentrating on the experience, and "becoming absorbed in the phenomenon without being possessed by it" (Oiler 1982, p. 180).

- *Analyzing* refers to (1) identifying recurrent elements in the experience by comparing and contrasting the descriptions obtained from participants and (2) using the familiar analytic scheme that asks for who, when, how, and why (Oiler 1982).

- *Describing* entails a detailed, comprehensive description of the experience or phenomenon that conveys what has been found or seen and the central characteristics of the phenomenon, also called recurrent themes.

Omery (1983) summarized steps of three phenomenological methodologies originally presented by Giorgi (1970), Van Kaam (1966), and Spiegelberg (1960). Although these methodologies differ in the various steps of the research process, the procedures advocated for data analysis are similar. They include (1) carefully reading transcriptions of interviews, (2) identifying units or categories that reflect the essential constituents of the phenomenon or experience under study, (3) eliminating redundancies in the elements, (4) clarifying the meanings of the elements, (5) relating

the elements to one another and to the experience as a whole, and (6) describing the experience. The essential elements of a phenomenon are identified by intuition, a frequency count, or judges who are expert in the field of phenomenology. The final product of a phenomenological inquiry is a narrative description of the phases of the experience or elements of the phenomenon, within its context.

As an example of an application of phenomenology that might be productive and helpful, the authors suggest the experience of drug craving and attempts at self-control. It is remarkable how little systematic work has been done on drug craving—how it is experienced, what efforts people make to control their urges, what they tell themselves about giving in to those urges, how they deal with the sequelae of relapse, and so on. Phenomenological study is well suited to the sort of "getting into minds" required to achieve depth of understanding of the craving experience and responses to it.

Issues in Qualitative Data Analysis

Qualitative research is, in general, interpretive; it is concerned with the researcher's subjective understanding of a phenomenon, experience, or behavior from the participant's point of view, within a specific natural context. Subjective understanding requires personal contact or interaction with participants, that is, dialectic process, openness, need for setting aside previous knowledge of the phenomenon examined, intuition, creativity, attention to various aspects of the phenomenon (holistic approach), and absorption or submersion in the phenomenon. These characteristics of qualitative data analysis raise two critical issues.

Self-as-Instrument.

To gain subjective understanding, the investigator attempts to obtain the trust of and rapport with participants, interviews or "converses" with them to collect data, makes notes of observations and interpretations or hypotheses, manages and analyzes data, and provides a description and interpretation of the phenomenon or experience. Thus, the investigator is the "instrument through which data is collected" (Rew et al. 1993, p. 300), processed, and analyzed. Consequently, data collection and analysis are inescapably influenced by the researcher's own beliefs, perspectives, biases, underattention or overattention to various aspects of the studied setting, selective memory (Firestone and Dawson 1988, pp. 209-221), communication and interview skills, and intuitive and analytic processes. Although the subjective understanding is expected to be reached through the exchange of ideas, interaction, and agreement between the researcher and participant, the researcher should avoid imposing his or her views, should set aside any preconceived knowledge, and should be open, sensitive, and empathetic to the participants'

responses, a task known to be difficult to accomplish. Therefore, qualitative investigators are encouraged to record their own biases, feelings, and thoughts and to state them explicitly in the research report (Creswell 1994). Nonetheless, the extent to which characteristics of the investigator will have played a role in or even interfered in data analysis cannot be known.

Lack of Guidelines for Data Analysis. To understand a phenomenon, experience, or behavior from the participants' point of view, data are collected in an unstructured format that allows participants to freely express their conceptualizations. In addition, data analysis, consisting of data reduction, display, and conclusion drawing, is an interpretive process guided by the available data as well as by the openness, intuition, and creativity of the researcher. Intuition and creativity are "private" processes that are difficult to describe and understand (Firestone and Dawson 1988, pp. 209-221). The process of qualitative data analysis is described as "eclectic," and there is no "right way" (Creswell 1994). Therefore, the conclusions drawn from such an interpretive, intuitive analysis are uncertain, may lack credibility (Miles and Huberman 1988, pp. 222-244), and may misrepresent participants' responses unless researchers describe the method of analysis used and show how the conclusions were drawn from the data "via warrants that are supported by backings" (Lythcott and Duschl 1990, p. 446). The lack of guidelines for data analysis is an issue that needs to be addressed by qualitative researchers because it is a threat to the credibility or reliability (i.e., internal validity) of their investigation and their conclusions.

In addition, explicit methodology for qualitative research is necessary and should be followed with at least as much fidelity as for quantitative research. Claiming to be engaged in qualitative research is no license for doing whatever seems a good idea at the time or whatever an investigator feels like doing. If results of qualitative inquiry are to be credible, they must be buttressed by referencing the best authoritative sources available and by explicit descriptions attesting to the fidelity of implementation of recommended procedures.

Finally, the case for qualitative research often is set forth in the context of world views, ideologies, and even epistemologies that are not necessary foundations for the work. The position can be taken that the meaning an individual attaches to some phenomenon is important in its own right without extending that position to incorporate any further propositions about whether there is any independent reality. Qualitative research may be hospitable to political liberalism, but it does not require it. Qualitative methods are additional tools to the more established quantitative approaches in research and are potentially available to all researchers who wish or need to use them.

CONCLUSIONS

Miles and Huberman (1988, pp. 222-244) summarized qualitative data analysis, using ethnography, grounded theory, or phenomenology, into three activities:

- Data reduction: the process of selecting, focusing, simplifying, abstracting, and transforming raw data

- Data display: assembling information in an organized manner to help understand what is occurring

- Conclusion-drawing and verification: noting regularities, patterns, explanations, possible configurations, causal flaws, and propositions that are verified and tested for their plausibility, robustness, sturdiness, and validity (Miles and Huberman 1988, p. 230)

As described in the literature, these activities rely on the inquirer's subjective understanding, intuition, familiarity with the phenomenon studied or participants' viewpoint, and analytic skills. Moreover, no structured guidelines for describing the process of qualitative data analysis are provided so that researchers are allowed enough flexibility in understanding a phenomenon from the participants' perspective. However, reliance on intuition and lack of analysis guidelines threaten the credibility of conclusions and therefore add to, rather than reduce, the uncertainty regarding the phenomenon studied. Consequently, encouraging qualitative researchers to explicitly describe their analytic processes and to supplement their qualitative analysis with quantitative exploratory analytic procedures, such as cluster analysis, multidimensional scaling, and exploratory factor analysis, will enhance the credibility and validity of their conclusions.

Triangulation is a promising strategy for a cumulative science: using multiple methods for examining its extent, validating results (i.e., convergence), and providing complementary information about a phenomenon (i.e., completeness) (Breitmayer et al. 1993; Fielding and Fielding 1986).

REFERENCES

Atkinson, P., and Hammersley, M. Ethnography and participant observation. In: Denzin, N.K., and Lincoln, Y.S., eds. *Handbook of Qualitative Research.* Thousands Oaks, CA: Sage, 1994. 386 pp.

Bednarz, D. Quantity and quality in evaluation research: A divergent view. *Eval Program Plann* 8:289-306, 1985.
Breitmayer, B.J.; Ayres, L.; and Knafl, K.A. Triangulation in qualitative research: Evaluation of completeness and confirmation purposes. *Image J Nurs Sch* 25(3):237-243, 1993.
Brewer, J., and Hunter, A. *Multimethod Research. A Synthesis of Styles.* Newbury Park, CA: Sage, 1989. 224 pp.
Bryman, A. The debate about quantitative and qualitative research: A question of method or epistemology. *Br J Sociol* 35:78-92, 1984.
Burns, N., and Grove, S.K. *The Practice of Nursing Research: Conduct, Critique and Utilization*. Philadelphia: W.B. Saunders, 1987. 790 pp.
Chenitz, W.C., and Swanson, J.M. *From Practice to Grounded Theory: Qualitative Research in Nursing.* Menlo Park, CA: Addison-Wesley, 1986. 259 pp.
Connides, I. Integrating qualitative and quantitative methods in survey research on aging: An assessment. *Qualitative Sociol* 6:334-352, 1983.
Cook, T.D., and Campbell, T.D. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton-Mifflin, 1979. 349 pp.
Creswell, J.W. *Research Design. Qualitative and Quantitative Approaches.* Thousands Oaks, CA: Sage, 1994. 200 pp.
Dennis, M.L.; Fetterman, D.M.; and Sechrest, L. Integrating qualitative and quantitative evaluation methods in substance abuse research. *Eval Program Plann* 18(1):77-87, 1995.
Denzin, N.K., and Lincoln, Y.S. Entering the field of qualitative research. In: Denzin, N.K., and Lincoln, Y.S., eds. *Handbook of Qualitative Research.* Thousands Oaks, CA: Sage, 1994. 643 pp.
Dzurec, L.C., and Abraham, J.L. The nature of inquiry: Linking quantitative and qualitative research. *ANS Adv Nurs Sci* 16(1):73-79, 1993.
Fetterman, D.M., ed. *Qualitative Approaches to Evaluation in Education. The Silent Scientific Revolution*. New York: Praeger, 1988. 312 pp.
Fielding, N.G., and Fielding, J.L. *Linking Data*. Newbury Park, CA: Sage, 1986. 96 pp.
Firestone, W.A., and Dawson, J.A. Approaches to qualitative data analyses: Intuitive, procedural, and intersubjective. In: Fetterman, D.M., ed. *Qualitative Approaches to Evaluation in Education. The Silent Scientific Revolution*. New York: Praeger, 1988. 312 pp.
Giorgi, A. *Psychology as a Human Science, A Phenomenologically Based Approach.* New York: Harper & Row, 1970. 240 pp.
Glaser, B.G., and Strauss, A.L. The purpose and credibility of qualitative research. *Nurs Res* 15(1):56-61, 1966.

Guba, E.G., and Lincoln, Y.S. Do inquiry paradigms imply inquiry methodologies? In: Fetterman, D.M., ed. *Qualitative Approaches to Evaluation in Education. The Silent Scientific Revolution*. New York: Praeger, 1988. 312 pp.

Helitzer-Allen, D.L., and Kendall, C. Exploring differences between qualitative and quantitative data: A study of chemoprophylaxis during pregnancy. *Health Educ Q* 19:41-54, 1992.

Holstein, J.A., and Gubrium, J.F. Phenomenology, ethnomethodology, and interpretive practice. In: Denzin, N.K., and Lincoln, Y.S., eds. *Handbook of Qualitative Research.* Thousands Oaks, CA: Sage, 1994.

Hughes, C.C. "Ethnography": What's in a word—Process? Product? Promise? *Qualitative Health Res* 2(4):439-450, 1992.

Jick, T.D. Mixing qualitative and quantitative methods: Triangulation in action. *Adm Sci Q* 24:602-611, 1979.

Lythcott, J., and Duschl, R. Qualitative research: From methods to conclusions. *Sci Educ* 74:449-460, 1990.

Mathison, S. Why triangulate? *Educ Res* 17(2):13-17, 1988.

Miles, M.B., and Huberman, A.M. Drawing valid meaning from qualitative data: Toward a shared craft. In: Fetterman, D.M., ed. *Qualitative Approaches to Evaluation in Education. The Silent Scientific Revolution*. New York: Praeger, 1988. 312 pp.

Oiler, C. The phenomenological approach in nursing research. *ANS Nurs Res* 31(3):178-182, 1982.

Omery, A. Phenomenology: A method for nursing research. *ANS Adv Nurs Sci* 5(2):49-63, 1983.

Patton, M.Q. Paradigms and pragmatism. In: Fetterman, D.M., ed. *Qualitative Approaches to Evaluation in Education. The Silent Scientific Revolution*. New York: Praeger, 1988. 312 pp.

Rew, L.; Bechtel, D.; and Sapp, A. Self-as-instrument in qualitative research. *Nurs Res* 42(5):300-301, 1993.

Roter, D.L., and Frankel, R. Quantitative and qualitative approaches to the evaluation of medical dialogue. *Soc Sci Med* 34:1097-1103, 1992.

Schatzman, L., and Strauss, A.L. *Field Research.* Englewood Cliffs, NJ: Prentice-Hall, 1973. 176 pp.

Sechrest, L., and Sidani, S. Quantitative and qualitative methods. *Eval Pract*, in press.

Shadish, W.R. Critical multiplism: A research strategy and its attendant tactics. In: Sechrest, L., ed. *Program Evaluation: A Pluralistic Enterprise. New Directions in Program Evaluation.* No. 60. San Francisco: Jossey-Bass, 1993. 101 pp.

Shannon, C.E., and Weaver, W. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press, 1949. 125 pp.

Spiegelberg, H. *The Phenomenological Movement, Vol. 1*. The Hague, Netherlands: Martinus Nijhoff, 1960. 315 pp.

Steckler, A. The use of qualitative evaluation methods to test internal validity. *Eval Health Prof* 12:115-133, 1989.
Strauss, A., and Corbin, J. Grounded theory methodology. An overview. In: Denzin, N.K., and Lincoln, Y.S., eds. *Handbook of Qualitative Research.* Thousands Oaks, CA: Sage, 1994.
Van Kaam, A. *Existential Foundations of Psychology, Vol. 3*. Pittsburgh, PA: Duquesne University Press, 1966. 386 pp.
Walters, J.M. What is ethnography? In: Akins, C., and Beschner, G., eds. *Ethnography: A Research Tool for Policymakers in the Drug and Alcohol Fields*. DHHS Pub. No. (ADM)80-946. Rockville, MD: National Institute on Drug Abuse, 1980. 128 pp.
Webb, E.; Campbell, D.T.; Schwartz, R.D.; Sechrest, L.; and Grove, J.B. *Nonreactive Measures in the Social Sciences*. Boston: Houghton-Mifflin, 1981. 394 pp.
Wilson, H.S., and Hutchinson, S.A. Triangulation of qualitative methods: Heideggerian hermeneutics and grounded theory. *Qual Health Res* 1(2):263-276, 1991.
Woodhouse, L.D., and Livingood, W.C. Exploring the versatility of qualitative design for evaluating community substance abuse prevention projects. *Qual Health Res* 1(4):434-445, 1991.

AUTHORS

Souraya Sidani, Ph.D.
Assistant Professor
Faculty of Nursing
University of Toronto
50 St. George Street
Toronto, Ontario M5S 1A1
CANADA
(416) 978-2856 (Tel)
(416) 978-8222 (Fax)
s.sidani@utoronto.ca (Internet)

Lee Sechrest, Ph.D.
Professor
Department of Psychology
University of Arizona
Psychology Building
Tucson, AZ 85721
(520) 621-9182 (Tel)
(520) 621-6320 (Fax)
sechrest@u.arizona.edu (Internet)