Clinical Psychology Review, Vol. 12, pp. 853-865, 1992
Printed in the USA. All rights reserved.

0272-7358/92 $5.00 + .00

EVALUATION PLANNING FOR AN INNOVATIVE CHILDREN'S MENTAL HEALTH SYSTEM

Leonard Bickman
Craig Anne Heflinger
Georgine Pion

Vanderbilt University

Lenore Behar

North Carolina Division of Mental Health, Developmental Disabilities and Substance Abuse Services

ABSTRACT. *We describe an evaluation of an innovative mental health system for children and adolescents. The Ft. Bragg Demonstration was developed to provide a complete continuum of care to children receiving health care benefits through CHAMPUS. The evaluation, jointly funded by the State of North Carolina and the NIMH, is designed to evaluate the implementation, service quality, clinical outcomes, and costs of this new system of care. Over a 5-year period, approximately 1,100 children and families will be studied in a four-wave longitudinal study at Ft. Bragg and two comparison sites. This project should provide important information about whether these new systems work better than the traditional array of services.*

Strong concensus exists concerning the problematic manner in which mental health services are provided to children. Many children do not receive any services, and others receive inappropriate services. In the past two decades, experts (Hobbs, 1982; Knitzer, 1982; Stroul & Friedman, 1986) have highlighted the vast discrepancy between the numbers of children and youth in need of mental health services and those who receive services. More than half of these children receive no treatment at all, and many who are treated are receiving inappropriate care (Saxe, Cross, & Silverman, 1988). Senator In-

---

Correspondence should be addressed to Leonard Bickman, Vanderbilt University, Vanderbilt Institute for Public Policy Studies, Center for Mental Health Policy, Box 7701 Station B, Nashville, TN 37235.

ouye (1988) claimed that 80% of the children who need services are receiving inappropriate care or none at all.

There is also agreement that unnecessarily restrictive treatment settings are overutilized (NMHA, 1989). Children with emotional problems are best treated in the least restrictive, most normative environment that is clinically appropriate. However, according to Congressional testimony, the number of children placed in private inpatient psychiatric settings increased from 10,764 such placements in 1980 to 48,375 in 1984—a 450% increase (Stroul & Friedman, 1986). Moreover, the number of private psychiatric hospitals continues to grow (Bickman & Dokecki, 1989).

Contributing to this problem is the fact that alternative treatment settings are generally unavailable. Knitzer (1982), Behar (1985), and Silver (1984) all reported that approximately 40% of inpatient placements were inappropriate because either the children could have been treated in less restrictive settings, or the placements that were initially appropriate were no longer appropriate, but less restrictive treatment settings were not available. This remains the situation in spite of evidence that even severely emotionally disturbed children can receive treatment while living in their own homes when a comprehensive system of care is present in the community (Behar, 1985).

Even where services are available, the lack of coordination between programs compromises the effectiveness of the interventions (Saxe, Cross, Silverman, Batchelor, & Dougherty, 1987; Stroul & Friedman, 1986). Given the developmental complexity and multiple needs of children and adolescents, services must be both available and coordinated (Behar, 1985).

CONTINUUM OF CARE AS AN ALTERNATIVE TO TRADITIONAL SYSTEMS

The continuum of care approach has emerged in response to the problems characterizing mental health service delivery systems for children and adolescents. The term *continuum of care* refers to the comprehensive range of services required to treat severely disturbed children and adolescents, which includes both nonresidential and residential services (Stroul & Friedman, 1986). This approach attempts to deliver needed services on an individualized basis and in a coordinated manner, relying on case management to integrate treatment programs and facilitate transitions between services. It also is designed to be community-based, involving various agencies pertinent to children's developmental, social, medical, and mental health needs. There is consensus among professionals that the most effective way to deliver mental health services to children and adolescents is in a continuum of care system (Stroul & Friedman, 1986). However, there has not been a definitive study that has demonstrated the superiority of the continuum of care model to the traditional method of service delivery. The Ft. Bragg Evaluation is the first comprehensive evaluation of this system of care.

THE FT. BRAGG DEMONSTRATION PROJECT

The high cost of providing mental health services to the children and adolescents of military personnel stimulated the Office of the Civilian Health and Medical Program of the Uniformed Services (OCHAMPUS) to consider alternatives to the existing delivery system. In 1983, CHAMPUS alone spent $74 million on inpatient mental health hospitalization for dependent children. In pursuit of better alternatives, the Department of the Army, in August of 1989, funded the Ft. Bragg Child and Adolescent Mental Health Demonstration Project (hereafter referred to as the Demonstration) through a contract with the North Carolina Department of Human Resources, Division of Mental Health, Developmental Disabilities, and Substance Abuse Services (MH/DD/SAS).

The state has contracted through the Lee–Harnett MH/DD/SA Program with CARDINAL Mental Health Group, Inc. (CARDINAL), a private, not-for-profit corporation, to provide a continuum of care for the Ft. Bragg catchment area. For a period of 4 years, mental health and substance abuse services will be provided to those in need of the 41,500 children and adolescents of military personnel in the Ft. Bragg area. The range of services will include both nonresidential and residential components. CARDINAL has contracted with individuals and agencies in the community already providing traditional mental health services such as outpatient therapy and acute inpatient hospitalization. For the "middle" of the continuum, those services not currently available in Fayetteville nor typically available across the country, CARDINAL developed and operates services that include in-home counseling, after-school educational treatment services, day-treatment services, therapeutic foster homes, specialized group homes, and a 24-hr crisis management team. All children and adolescents requesting services participate in a comprehensive intake/assessment process to determine the appropriate level of service, and any client receiving more-than-outpatient care is assigned a clinical case manager. Using a closed system or exclusive provider organization (EPO) model, families seeking services for their children and adolescents are required to use the Demonstration's clinical services, which are free, or to pay for services on their own.

The clinical services of the Demonstration include a comprehensive diagnostic assessment, treatment planning process, and ongoing review of treatment progress, which involve the client, the family, and other professionals. Treatment services available within the continuum include individual, family, and group outpatient therapies; in-school support programs; after-school programs; day treatment; individual and group residential treatment; and inpatient services. Crisis intervention is available by telephone, by visit to the emergency room, and through in-home crisis stabilization services. For children using multiple community services, the clinical services are coordinated with the other child-serving agencies/practitioners in the community, especially pediatric, education, and protective services. Services within the continuum and across other agencies are linked together through a case management component. Related services to parents are also provided. In order to assess the effectiveness of the clinical component on multiple levels, an independent evaluation component also has been funded.

THE FT. BRAGG EVALUATION COMPONENT

The Center for Mental Health Policy of the Vanderbilt Institute for Public Policy Studies at Vanderbilt University has been awarded a subcontract by the North Carolina MH/DD/SAS to conduct an independent evaluation of the Demonstration. Four critical issues are addressed by the Evaluation Project:

1. *implementation* of the Demonstration and issues concerning its replication at other sites;
2. *quality* of services provided by the Demonstration;
3. *mental health outcomes* of the children and adolescents who receive services; and
4. *costs* of services delivered at the Demonstration.

These central issues are the focus of specific studies in the Evaluation Project.

Studying Implementation

Essential to the conduct of a high quality evaluation is the need to address questions of program conceptualization, design, and implementation (Hargreaves & Shumway, 1989; Rossi & Freeman, 1985). This aspect of the evaluation examines the theories and assumptions underlying the hypothesis that a specific intervention should be successful; work toward ensuring that the program's major goals, individual components, and specific activities do indeed "fit together"; follow a logical sequence; and appear likely to produce the desired outcomes. Moreover, as evaluators have been frequently reminded (e.g., Rezmovic, 1984; Scheirer, 1981), it is unwise to simply assume that the program will be delivered as planned to its intended recipients. Various problems can surface, despite the best efforts of program architects, including temporary or permanent obstacles to the program reaching all members of the target population, and inability to provide the required treatment "dosage" and consistent delivery of high quality treatment to all participants. Thus, it becomes important that structural, environmental, and/or political barriers responsible for diluting full-scale implementation of the program be documented.

Another reason for measuring program implementation is to gain better insight into the relationships between program inputs and outcomes. For example, determinations can be made as to which program elements or processes appear more effective than others, and which classes of program recipients benefit most from the intervention. In addition, the thorough description of services actually provided will advance the field in the effort to define various components of the continuum of care. Finally, implementation data gathered throughout the course of the program (i.e., from its initial "start up" phase through its "fully operational" stage) can be used as a guide to others who wish to replicate the program in different sites.

The overall strategy for examining program implementation is based on both Chen and Rossi's (1983) "theory-driven" approach to program evaluation and Bickman's program (1987, 1990) and component (1985) theories of evaluation. Whereas the "theory-driven" perspective essentially aims at developing models that identify the causal and operational linkages among program elements, the component approach proceeds one step further. Here the emphasis is placed on discerning distinct philosophies, "subtheories," and activities, along with the linkages among these, *within* the individual program elements/components (see Graham & Birchmore-Timney, 1989 for an example of this strategy). Thus, combining these approaches should result in a descriptive model of program structure, process, and outcomes for the Ft. Bragg service delivery system as a whole, and for each of the service components that are incorporated under its administrative umbrella.

For the purposes of understanding and evaluating the program at Ft. Bragg, this strategy seems particularly appropriate. The Demonstration is an attempt to develop and implement a model service delivery *system* for addressing the mental health needs of children and adolescents. At the same time, this system is composed of several different types of treatment settings and facilities. While all of these individual aspects of the program may subscribe to the overall philosophy and values held by Demonstration program administrators and staff, they also will have their own set of theories and values that define and guide the structure, recipients, process, and outcomes of their efforts; not all of these may overlap perfectly with those of the Demonstration program. For example, one key element of the philosophy encompassed by the Demonstration Project concerns the need to involve the family in treatment. While this thrust has certain overarching features, the way it is operationalized on a daily basis by different components (e.g., the types and amount of information on the child's progress reported to the family by staff in residential treatment centers vs. group homes), or even by different providers within these components, may vary. As such, understanding both the set of theories and values that underlie the overall Demonstration service delivery system and those influencing the operation of its individual "building blocks" is important to assessing program implementation.

Once accurate conceptual and operational models of the program have been developed, they will be translated into variables that can reflect the degree of program implementation. Here, the goal is to collect information that will assess coverage, bias, and outputs. It is anticipated that data for ascertaining the fidelity of the program to its intended conceptualization and design will be obtained from six basic sources: (a) program services records on client characteristics, diagnoses, client movement through the Demonstration, and services delivered; (b) client files (e.g., types of case management received); (c) reports by parents and significant others (e.g., perceptions as to the extent they were involved in treatment); (d) reports by service providers and others involved in the child's treatment regarding the characteristics of services delivered; (e) peer review of treatment received (e.g., the extent to which the child was treated in the least restrictive, appropriate setting); and (f) observational data.

Determining the Quality of Services

A significant issue concerning mental health systems in the coming decade is the need for research on assessment, monitoring, and improvement of the quality of mental health services (Bickman & Peterson, 1990; Peterson & Bickman, 1992; Wells, 1988). Providers currently must meet the typically minimal requirements of legislative mandates, hospital accreditation programs, and private insurance carriers for providing quality care. However, the changing nature of the mental health system has stimulated the need for systematic research on the nature of quality. The body of research concerned with defining, assessing, and assuring quality of mental health services is not well developed, and in fact lags far behind advances in the physical health area and, in general, other mental health research issues. While there exist considerable philosophical and methodological difficulties in defining and measuring the quality of mental health services, the importance of this topic warrants vigorous investigation (Palmer, Donabedian, & Povar, 1991).

Berwick (1989) contrasts what he calls the "Theory of Bad Apples" with the "Theory of Continuous Improvement." In the former, quality is best achieved through inspection followed by discovery and removal of the bad apples. It implies certain thresholds for acceptability and a search for outliers, and also emphasizes deficiencies in personnel who warrant close scrutiny. In health care, this is most evident in the Health Care Financing Administration's publication of mortality profiles of Medicare recipients in U.S. hospitals. Berwick believes that this use of quality results in gaming the system, blaming the victim, and no real improvement in quality. Moreover, one also must avoid the often minimalist standards found in quality assurance programs that follow the bad apple theory of quality. Such standards often become ceilings instead of floors, and prevent the achievement of excellence.

The quality assessment approach of the Evaluation will have two tracks. One track will review the quality assurance (QA) activities of CARDINAL at the Demonstration site, which, according to the stipulations of the Department of the Army contract, follow the requirements of the Joint Commission on Accreditation of Health Care Organizations (JCAHO). Consistent with the JCAHO model, QA is a complex management tool, including (a) credentialing and privileging of clinicians, (b) monitoring against indicators of quality programming, (c) clinical care studies, and (d) utilization review. Indicators are developed for each service component to reflect issues of quality and to identify areas needing further investigation through clinical care studies. Examples of such indicators are: (a) in emergency services, the number of clients moving from telephone interview to face-to-face interview to hospital admission per month, or (b) in diagnostic services, the number of days elapsing between the family's request for services and the scheduled intake assessment. Essentially, in areas where CARDINAL plans to implement QA activities, the Evaluation will assess the extent to which CARDINAL meets its own QA criteria and standards.

The second track will assess, at the program level, the quality of those operational service components that are unique and crucial to the continuum of care model. These are components that are neither direct treatment services (e.g., outpatient care, day treatment) nor indirect services (e.g., legal, educational services). Instead, the component level of evaluation will focus on two key aspects of the continuum of care—intake/assessment and case management. These two system components were chosen for assessment because they are especially vital to the effectiveness of the Demonstration model. Thus, they will likely be defined, developed, and implemented differently in the continuum of care than in typical treatment settings.

To measure quality of services at the component level, those services must first be described. The need for research on services description is perhaps best represented through an analogy. Many resources have been directed toward defining and measuring mental illness. This focus, however, has neglected the description of treatment that may induce change in mental health. Mental health research focuses almost exclusively on the dependent variable, and little information is provided concerning the independent variable—services. Often, the services the client receives are described as simply "day treatment," which is equivalent to describing an independent variable in a psychological experiment as occurring in the laboratory or in the field. This description of the environment of the independent variable provides little information.

Instead, it is argued here that one must begin to describe services at a more specific and informative level. In sum, previous research on services has taken a "black box" approach, where the inputs and outputs are studied, but not what goes on within the box (service). Our approach is an initial attempt to systematically represent some of the workings within the black box. In the present evaluation, the focus on intake/assessment and case management reduces the need to develop a complete taxonomy of services.

The approach proposed here is based on Bickman's (1985) component theory of evaluation. The component approach was developed as a comprehensive means of first describing and then assessing statewide services delivered to preschool children. The logic behind this approach is that evaluations can be designed to examine components of a program (or service) rather than the entire program/service. A component is viewed as the largest homogenous unit of a service. Each serves as a building block of services and, although not fully independent of one another, may be studied separately.

Once the program components and the activities that make up a component are explicated, the next step is to assign a weight or value to each activity so that evaluators can assess service quality by measuring the presence/absence, frequency, intensity, and appropriateness of the activities. Valuing the activities associated with high quality services is obviously a formidable task, particularly in relation to mental health services, where the diversity and sometimes incompatibility of theoretical orientations toward illness etiology and therapeutic approach makes consensus difficult. The technique to be used here in valuing service activities will be to convene an advisory panel of experts who specialize in the components of interest (e.g., assessment). The advisory panel will rate the appropriateness of each activity level and rate the importance of each activity to a high quality component. Once the expert panel has made its recommendations, a checklist will be developed that can be used by evaluators at Ft. Bragg. The results of the checklist observations will be analyzed by applying a weighting scheme and computing a summative quality rating regarding each component.

In addition to this approach, the Evaluation also plans to assess the proximal outcomes of these two components. For example, one outcome of a good intake/assessment process is the appropriateness of the initial diagnosis and placement of the child. Evaluation of actual records, combined with the data collected as part of the outcome study, will help us assess the appropriateness of the diagnosis. In a similar fashion, the case management function can be assessed by examining the pattern of service placements. The focus here is on whether the child is placed in the most appropriate level of service and whether the child is moved to other levels of service appropriately.

Measuring Mental Health Outcomes

Currently, little information is available on the effects of innovative models of mental health treatment on clinical outcomes. Several major efforts are underway to demonstrate and evaluate systems of care, including the Robert Wood Johnson's Mental Health Services Program for Youth (Beachler, 1990) and the Ventura Project (Jordan & Hernandez, 1990). Attempts to individualize services are also being reported, such as Kaleidoscope in Illinois, the Alaska Youth Initiative, and Project Wraparound in Vermont (Burchard & Clarke, 1990). These latter efforts, however, have been aimed at small populations of children and adolescents with severely maladjusted behavior who were receiving intensive and expensive services, often out of state. Results released to date have focused on costs and levels of service, with little information on mental health outcome for the clients in question. Furthermore, this work is difficult to generalize to a community-based effort involving children and adolescents with a wide range of types and severity of problems.

Given the importance of a client-level focus on mental health outcomes, the prior exclusion of such a focus raises the question of why client-level outcomes have not been more center-stage. Reasons for excluding the study of client-level mental health outcomes have included: (a) the expense of such a study, (b) the political pressures generally present in most evaluation efforts to focus on issues such as the equitable distribution of services rather than the impact of those services (Bickman & Rog, 1986; Newman, Hunter, & Irving, 1987; Schulberg, 1981), and (c) the conceptual and pragmatic difficulties in defining and measuring mental health outcomes.

Key questions, therefore, that address mental health outcomes in the Ft. Bragg Evaluation include:

1. Are there improvements in mental health outcomes of the children and adolescents served in the Demonstration?
2. Do the children and adolescents served in the Demonstration exhibit equal or greater improvement than those children and adolescents receiving typical mental health services?
3. What mediating factors and processes contribute to the outcomes?

Mental health outcomes will be studied longitudinally to assess whether children's clinical conditions improve more and faster than children in typical treatment settings. Additionally, the Evaluation will try to determine whether the children and their families are more satisfied with continuum of care services than they are with typical services. After a child or adolescent is recommended for treatment by the intake/assessment team, the Evaluation staff will conduct the first of four comprehensive, in-person interviews with the child or adolescent and family. Additional interviews by Evaluation staff will take place 6, 12, and 18 months after intake. This last interview is supported by a National Institute of Mental Health (NIMH) grant to Vanderbilt to enhance the evaluation. In addition, two comparison sites have been designated. They are located at Ft. Campbell, Kentucky and Ft. Stewart, Georgia, where dependent children and adolescents are receiving care in typical mental health treatment settings. The CHAMPUS-insured services include only three types of services: hospitalization, care in a residential treatment facility, and outpatient. Moreover, there is no single point of entry and coordination of services through case management as in the Demonstration. The sites at Ft. Stewart and Ft. Campbell will serve as comparison sites to the Demonstration at Ft. Bragg. We plan to recruit 550 clients in the Demonstration site and 225 in each of the comparison sites. This will provide sufficient statistical power to detect effects as small as 0.2 of a standard deviation (Lipsey, 1990).

Data Collection Strategies. The primary source of mental health outcome data will be the research participants themselves—the children and adolescents and their families who are receiving mental health services at the Ft. Bragg Demonstration or one of the comparison sites. These interviews will use multiple measures and are designed to be (a) comprehensive, providing information on a multitude of child and family variables; (b) standardized, through the use of established instruments and trained interviewers; and (c) feasible, asking children and parents to provide adequate but not excessive amounts of information.

The primary data collection effort focuses on the children's clinical functioning and the functioning of their families. The data on children includes psychiatric status, behavior problems and social competence, level of functioning, self-esteem, and school adjustment and achievement. Family data includes family perceptions of stresses and hardships imposed by their child's mental health problems, family functioning, family life events, and family resources. Data are collected through similar interviews with the child and parent, as well as through self-report measures and questionnaires.

The instrument package developed for this study consists of a combination of structured and semistructured interviews, behavioral checklists, and self-report questionnaires. The domain of child psychopathology is measured by the Child Assessment Schedule (CAS) (Hodges, Kline, Fitch, McKnew, & Cytryn, 1981; Hodges, Kline, Stern, Cytryn, & McKnew, 1982), including the parallel form, the Parent-CAS (PCAS); selected modules from the revised Diagnostic Interview Schedule for Children (DISC-2.1) (Shaffer, Fischer, Piacentini, Schwab-Stone, & Wicks, 1989); the Child Behavioral Checklist (CBCL) (Achenbach & Edelbrock, 1983); and the Youth Self-Report (YSR) (Achenbach & Edelbrock, 1987) for teenagers. The Family Background Form (FBF) and the Treatment History Interview (THI), both developed at Vanderbilt, are used to collect background information, including the child's physical and mental health history, experiences with schools, and contacts with law enforcement and court systems.

To measure social functioning, a questionnaire, the Self-Perception Profile (SPP) (Harter, 1982), is used. The CBCL and the YSR also include items that measure social functioning. In addition, the interviewer will complete a modification of the Child Global Assessment Scale (CGAS) (Shaffer et al., 1983) and the Child and Adolescent Functional Assessment Scale (CAFAS) (Hodges, 1990), developed by Kay Hodges in conjunction with this project and modeled after the North Carolina Functional Assessment Scale (NCFAS), which was developed primarily for use with adults.

The domain of family functioning is assessed through several self-report instruments, including the Family Assessment Device (FAD) (Epstein, Baldwin, & Bishop, 1983; Miller, Epstein, Bishop, & Keitner, 1985), the Family Inventory of Life Events (FILE) (Olson et al., 1982), and the Family Resource Scale (FRS) (Dunst & Leet, 1987). Also, the Brief Symptom Inventory (BSI) (Derogatis & Spencer, 1982) is used to measure parental psychopathology. Using the above-mentioned FBF and THI, additional information will be collected on the family's physical and mental health history, mental health services used by immediate family members, and their contacts with law enforcement and court systems.

Finally, two additional questionnaires were developed by Vanderbilt. The Client Satisfaction Survey (CSS) is designed to measure how satisfied clients and their families are with the services they receive at the Demonstration and comparison sites. Issues addressed at both the individual service component and global levels include (a) access and convenience, (b) involvement in treatment decision-making, (c) relationships with therapists and other staff members, and (d) perceived effectiveness of services. The Burden of Care Questionnaire (BCQ) measures the extent to which families are burdened by having an emotionally disturbed child in the family, and how that burden changes as the child enters and receives treatment. Also, there is an attempt to discern the level of burden associated with having a child in residential as opposed to nonresidential treatment settings.

The CSS and BCQ were developed to address the issues faced by children, adolescents, and their families in the mental health system that may involve several components of care, including the innovative programs of the Demonstration. The psychometric properties of the CSS and BCQ are under examination. Additional collateral data will be collected from the child's teacher using the Teacher Report Form (Edelbrock & Achenbach, 1984), and from the child's therapist using a survey specially developed for this project.

In the development of this package, each instrument has undergone a series of pilot tests and refinements based on feedback received. Several of the instruments have been adapted for use in this package and altered to eliminate duplication of items among instruments and to enhance readability. The instrumentation package has undergone review by members of a family advocacy organization as well as black and Hispanic mental health experts for possible cultural biases.

Data Management/Quality Assurance. To assure high quality interview data, all interviewers participate in an intensive 5-day training program and subsequent 10 days of independent work. To qualify to collect data, each interviewer must reach criteria (kappa = .75) on four out of five Child Assessment Schedule (CAS) training tapes after completing five practice tapes. In order to maintain quality, every interview (with the subject's permission) is recorded on audiotape. A 10% sample of each interviewer's tape is reviewed by a trained instructor.

Cost Analysis

The primary objective of the cost study of the Evaluation Project is to determine whether the cost of delivering continuum of care services is comparable or lower than the cost of delivering care in typical treatment settings (i.e., at the comparison sites).

Society as a whole is anxious for information on mental health care delivery systems that may promise reduced use of expensive and restrictive inpatient care, and smoother transitions from critical episodes back to fully normal lifestyles. To serve this broader interest, a different definition of relevant costs is necessary. All the costs borne by any segment of society are potentially relevant, though some of these costs may prove on close inspection not to involve the use of scarce resources, but merely transfers of titles to resources among individuals, firms, and government units.

One of the most influential benefit–cost analyses in mental health was conducted by Weisbrod (1983). Although Weisbrod did not develop a new method of analysis, his work represents the first application of an experimental benefit–cost analysis to a mental health program. In addition to examining the direct costs at primary treatment centers, Weisbrod considered other important social costs (e.g., law enforcement and legal activities and burdens placed upon the client's family and neighbors) and social benefits (e.g., improved physical health, labor productivity, and consumer decision-making efficiency) in his analysis. Weisbrod's topology of costs is especially useful in identifying the costs to consider for the Ft. Bragg Project.

Cost data will be assembled from both the Demonstration site and the two comparison sites. Efforts will be made to express all costs in dollar terms, either through measurement or estimation. However, as noted by Weisbrod (1981), there will likely be some costs that are very difficult to express in dollar terms (e.g., psychic losses). The magnitude of these will be estimated and compared across study sites without conversion to dollar units.

At the Demonstration site and the comparison sites, the research team will collect cost data at the system level as well as for individuals participating in the study. The system-level data will allow estimation of total and average resource consumption for client subpopulations, while the individual-level data will permit estimation of costs associated with different treatment regimes. In addition, the individual cost data will serve as a check on system data. For example, if the system-level data suggest a reduction in costs for children with behavioral disorders, the research team will look to the sample of such children (using the diagnoses determined by the research team) to determine whether the apparent reduction in costs is actual or is due to changes in diagnostic procedures by mental health providers.

The general strategy for assessing the cost of each service will include three steps: (a) development of a list of resources consumed, including units of each resource; (b) estimation of a unit dollar value for each resource; and (c) estimation of total dollar costs by multiplying resources consumed by appropriate unit dollar values and summing these products. Developing the list of resources and estimating unit costs will often rely on the same data source. For example, billing records may provide lists of resources consumed as well as initial estimates of the dollar value of those resources. However, in many cases alternative data sources (i.e., sources other than those used to estimate resource consumption) may be used as estimates of unit costs. For example, with some resources, national estimates of unit cost may be used in place of local figures.

CONCLUSIONS

The preceding sections have summarized the major questions of interest to the Evaluation and the key elements of the research plan designed to answer these questions. Since the award of the evaluation subcontract in November 1989, most of the effort has been devoted to developing the evaluation plan and pilot-testing data collection instruments and strategies. Full-fledged data collection began in November 1990.

It is our belief that the Ft. Bragg Demonstration Project and the accompanying Evaluation will be significant for several reasons. First, an innovative model for providing appropriate and high quality services to children and adolescents with mental health problems, one that is consistent with and capitalizes on our current knowledge base about this population, will be rigorously tested. In addition, a multitude of questions concerning such aspects as type, amount, quality, cost, and effectiveness of Demonstration services will be addressed. These questions have been designed to provide information that will be relevant to all major constituencies—that is, the families of children and adolescents with mental health problems, individuals who provide and administer services to these clients, the military, and those responsible for funding and setting future policy for children's mental health services. However, answering these questions is both expensive and difficult. The Demonstration project itself is estimated to cost over $50 million over the 4 years, while the Evaluation will cost over $7 million, including funds from both the NIMH and the State of North Carolina. There are many difficulties to overcome in fielding an evaluation of this magnitude in a new and previously unexplored area.

Finally, we would like to note that the next decade can prove extremely propitious in terms of increasing our knowledge about how better to address children's mental health problems. This project, along with others asking similar evaluation questions of different models of services for children, should result in a more comprehensive view as to what works best, with whom, and in what way for children and adolescents with mental health problems.

Acknowledgements—This research is supported by the U.S. Army Health Services Command (DA-DA10-89-C-0013) as a subcontract from the North Carolina Department of Human Resources/Division of Mental Health, Developmental Disabilities and Substance Abuse Services, and a grant to Dr. Leonard Bickman (RO1MH-46136-01) from the National Institute of Mental Health. An earlier version of this paper was presented at the 3rd annual conference "A System of Care for Children's Mental Health: Building a Research Base," Tampa, FL, February 1990.

The U.S. Army Health Services Command has reviewed this publication. This review does not necessarily imply that the Army or the U.S. Army Health Services Command approved the content of the publication or the conclusions reached therein.

REFERENCES

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and revised Child Behavior Profile.* Burlington, VT: Queen City Printers.

Achenbach, T. M., & Edelbrock, C. (1987). *Manual for the Youth Self-Report and Profile.* Burlington, VT: University of Vermont, Department of Psychology.

Beachler, M. (1990). The mental health services program for youth. *Journal of Mental Health Administration,* **17**, 115–121.

Behar, L. (1985). Changing patterns of state responsibility: A case study of North Carolina. *Journal of Clinical Child Psychology,* **14**, 188–199.

Berwick, D. M. (1989). Continuous improvement as an ideal in health care. *The New England Journal of Medicine,* **320**, 53–56.

Bickman, L. (Ed.). (1987). *Using program theory in evaluation. New directions for program evaluation, No. 33.* San Francisco: Jossey-Bass.

Bickman, L. (Ed.). (1990). *Advances in program theory. New directions for program evaluation, No. 47.* San Francisco: Jossey-Bass.

Bickman, L., & Dokecki, P. (1989). The for-profit delivery of mental health services. *American Psychologist,* **44**, 1133–1137.

Bickman, L., & Peterson, K. A. (1990). Using program theory to describe and measure program quality. In L. Bickman (Ed.), *Advances in program theory. New directions for program evaluation, No. 47.* San Francisco: Jossey-Bass.

Bickman, L., & Rog, D. (1986). The use of stakeholders in planning evaluations of early intervention programs. In L. Bickman & D. Weatherford (Eds.), *Evaluating early intervention programs for severely handicapped children and their families* (pp. 335–355). Austin, TX: Pro-Ed.

Bickman, L. B. (1985). Improving established statewide programs: A component theory of evaluation. *Evaluation Review,* **9**, 189–208.

Burchard, J. D., & Clarke, R. T. (1990). The role of individualized care in a service delivery system for children and adolescents with severely maladjusted behavior. *Journal of Mental Health Administration,* **17**, 48–60.

Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review,* **7**, 283–302.

Derogatis, L. R., & Spencer, P. M. (1982). *Administration and procedures: BSI, Manual—I.* Baltimore, MD: Johns Hopkins University School of Medicine.

Dunst, C. J., & Leet, H. E. (1987). Measuring the adequacy of resources in households with young children. *Child: Care, Health, and Development,* **13**, 111–125.

Edelbrock, C. S., & Achenbach, T. A. (1984). The teacher version of the Child Behavior Profile: I. Boys aged 6–11. *Journal of Consulting and Clinical Psychology,* **52**, 207–217.

Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy,* **9**, 171–180.

Graham, K., & Birchmore-Timney, C. (1989). The problem of replicability in program evaluation: The component solution using the example of case management. *Evaluation and Program Planning,* **12**, 179–187.

Hargreaves, W. A., & Shumway, M. (1989). Effectiveness of services for the severely mentally ill. In C. A. Taube & D. Mechanic (Eds.), *The future of mental health services research* (ADM 89-1600, pp. 233–283). Washington, DC: U.S. Government Printing Office.

Harter, S. (1982). The perceived competence scale for children. *Child Development,* **53**, 87–97.

Hobbs, N. (1982). *The troubled and troubling child.* San Francisco: Jossey-Bass Publishers.

Hodges, K. (1990). *Child and Adolescent Functional Assessment Scale.* Unpublished manuscript.

Hodges, K., Kline, J., Fitch, P., McKnew, D., & Cytryn, L. (1981). The Child Assessment Schedule: A diagnostic interview for research and clinical use. *Catalog of Selected Documents in Psychology,* **11**, 56.

Hodges, K., Kline, J., Stern, L., Cytryn, L., & McKnew, D. (1982). The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology,* **10**, 173–189.

Inouye, D. K. (1988). Children's mental health issues. *American Psychologist,* **43**, 813–816.

Jordan, D. D., & Hernandez, M. (1990). The Ventura planning model: A proposal for mental health reform. *Journal of Mental Health Administration,* **17**, 26–47.

Knitzer, J. (1982). *Unclaimed children.* Washington, DC: Children's Defense Fund.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research.* Newbury Park, CA: Sage.

Miller, I. W., Epstein, N. B., Bishop, D. S., & Keitner, G. I. (1985). The McMaster Family Assessment Device: Reliability and validity. *Journal of Marital and Family Therapy,* **11**, 345–356.

National Mental Health Association. (1980). *Invisible children report.* Alexandria, VA: National Mental Health Association Press.

Newman, F. L., Hunter, R. H., & Irving, D. (1987). Simple measures of progress and outcome in the evaluation of mental health services. *Evaluation and Program Planning,* **10**, 209–218.

Olson, D. H., McCubbin, H. I., Barnes, H., Larsen, A., Muxen, M., & Wilson, M. (1982). *Family inventories: Inventories used in a national survey of families across the family life cycle.* (Available from Family Social Science, 290 McNeal Hall, University of Minnesota, St. Paul, MN 55108)

Palmer, R. H., Donabedian, A., & Povar, G. J. (1991). *Striving for quality in health care: An inquiry into policy and practice.* Ann Arbor, MI: Health Administration Press.

Peterson, K. A., & Bickman, L. (1992). Normative theory and program quality in mental health services. In H. Chen & P. Rossi (Eds.), *Theory-driven evaluations in analyzing policies and programs.* Westport, CT: Greenwood.

Rezmovic, E. L. (1984). Assessing treatment implementation amid the slings and arrows of reality. *Evaluation Review,* **8**, 187–204.

Rossi, P., & Freeman, H. (1985). *Evaluation: A systematic approach.* Beverly Hills, CA: Sage.

Saxe, L., Cross, T., & Silverman, N. (1988). Children's mental health: The gap between what we know and what we do. *American Psychologist,* **43**, 800–807.

Saxe, L., Cross, T., Silverman, N., Batchelor, W. F., & Dougherty, D. (1987). *Children's mental health: Problems and treatment.* Durham, NC: Duke University Press.

Scheirer, M. A. (1981). *Program implementation: The organizational context.* Beverly Hills, CA: Sage.

Schulberg, H. C. (1981). Outcome evaluations in the mental health field. *Community Mental Health Journal,* **17**, 132–142.

Shaffer, D., Fischer, P., Piacentini, J., Schwab-Stone, M., & Wicks, B. (1989). *Diagnostic Interview Schedule for Children (Disc-2.1C): Child Version.* New York: New York Psychiatric Institute.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry,* **40**, 1228–1231.

Silver, A. A. (1984). Children in classes for the severely emotionally handicapped. *Journal of Developmental and Behavioral Pediatrics,* **5**, 49–54.

Stroul, B. A., & Friedman, R. M. (1986). *A system of care for severely emotionally disturbed youth.* Washington, DC: CASSP Technical Assistance Center.

Weisbrod, B. A. (1981). Benefit-cost analysis of a controlled experiment: Treating the mentally ill. *The Journal of Human Resources,* **16**, 423–548.

Weisbrod, B. A. (1983). A guide to benefit-cost analysis, as seen through a controlled experiment in treating the mentally-ill. *Journal of Health Policy, Politics, and Law*, *7*, 808–845.

Wells, K. B. (1988). Quality-of-care in mental health: Policy and personal perspectives. *Focus on Mental Health Services Research, 3*.