

Multi-site Evaluation and Research

edward mullen

This article examines multi-site evaluation and research. Differing from single-site research, multi-site research is appropriate for reasons that distinguish it from single-site research. This article examines forms and types of multi-site research to illustrate a variety of applications. The article presents examples of multi-site research conducted in the United States, at the national level with reference to mental health services and alcoholism treatment research applications. Additional mental health, child welfare, and social work practice examples are provided from research conducted at the Center for the Study of Social Work Practice in New York City. Advantages and disadvantages of multi-site research are described with suggestions for the conduct of multi-site research.

Edward J Mullen, B.A., M.S.W., D.S.W.; Willma & Albert Musher Chair Professor for Life Betterment through Science and Technology, Columbia University School of Social Work; Director, Center for the Study of Social Work Practice, a joint program of Columbia University School of Social Work & the Jewish Board of Family & Children's Services; Director, NIMH funded PhD Training Program in Mental Health Services Research.
The assistance of Haluk Soydan, Gretchen Borges, and Chito Trillana in the preparation of this paper is acknowledged.

Introduction

This article examines multi-site evaluation and research.¹ Differing from single-site research, multi-site research is appropriate for reasons that distinguish it from

¹ The term "research" is used and it is meant to include both "program evaluation" as well as other forms of intervention and survey research.

single-site research. In this article I examine forms and types of multi-site research to illustrate a variety of applications. I describe examples of multi-site research conducted in the United States, at the national level with reference to mental health services and alcoholism treatment research applications. Drawing from my experiences as director of research at a social work research center, the Center for the Study of Social Work Practice (CSSWP), I describe additional applications in mental health, child welfare, and social work practice.² I conclude with reflections on advantages and disadvantages of multi-site research drawing out

2 As Director of the Center for the Study of Social Work Practice I have had the opportunity to observe a wide range of multi-site research applications in social welfare. The Center has lent itself to multi-site research because of its organizational sponsorship as well as its mission. The Center is located at Columbia University in the City of New York. It is a joint program of the Columbia University School of Social Work (CUSSW) and the New York City based Jewish Board of Family and Children's Services (JBFCS). Research studies have been conducted by principal investigators who are faculty members at the university. Accordingly, their interests have been national and international in scope supporting studies located at sites in New York City as well as in other locations throughout the United States and to some extent in other countries. Because of the Center's affiliation with the JBFCS the Center has tended to carry out research at locations served by that agency's programs, oftentimes through multi-site research projects. Indeed, the population served by JBFCS and, therefore, the focus of the Center's work is comparable in size and diversity with that of many national service organizations in other

suggestions for the conduct of multi-site research.

Little has been written about the methodology of multi-site research. There are notable exceptions. An informative article published in French by Irene Elkin was written in 1992 based on her ten-year experience as Coordinator of the U.S. National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program (Elkin, I., 1992). Publications based on the experiences of Project MATCH, a multi-site study examining alternative alcoholism treatments, are another important contribution (Fuller RK, Mattson ME, Allen JP, Randall, CL, Anton, RF, Babor, TF, 1994;

countries. The agency has a target population of over eight million New York residents spread over five boroughs. Over the period of one-hundred years the JBFCS has grown into one of the United States' largest nonprofit mental health and social service agencies. Now, JBFCS is a comprehensive agency that serves over 54,000 New Yorkers annually from all religious, ethnic, and economic backgrounds through 140 community-based programs, clinics, residential facilities, and day-treatment centers. JBFCS employs 1,400 staff including professional social workers, licensed psychologists, and psychiatrists, as well as a cadre of clinical support personnel in continuing day treatment and residential treatment centers. In addition services are provided by approximately 1,700 volunteers. Many of the Center's studies have been carried out at one or more of the ten JBFCS' New York State licensed outpatient mental health clinics, which provide mental health services for a wide range of emotional and social problems. Services for adults and children include evaluation and assessment; crisis intervention; and time-limited, time-effective, and ongoing individual, couple, family, and group therapy.

Zweben A, Donovan DM, Randall CL, et al, 1994; Del Boca, FK, 1994). Another informative paper has been prepared by Robert Boruch and Ellen Foley, scheduled to appear in a book edited by Leonard Bickman (Boruch and Foley, in preparation). Also, Boruch and Lawrence Hedges have examined one type of multi-site evaluation research in their article which appears in this special issue of *Socialvetenskaplig Tidsskrift*. Indeed, there appears to be a scarcity of published analyses of multi-site research and evaluation methods. While multi-site studies have been common in recent years, it is as if there has been an assumption that it is sufficient to follow methods developed for use in single-site research. As a result, mistakes have been made and opportunities have been lost in many past multi-site studies.

Many questions need to be examined regarding multi-site research. What is multi-site research? Why conduct multi-site research? What is to be gained through multi-site studies? What is lost through multi-site investigations? What infrastructures are needed to successfully implement multi-site investigations? Elkin notes that when the NIMH Treatment of Depression Collaborative Research Program was considered a range of questions was posed (Elkin, I, 1992). These are important questions to consider more generally when undertaking multi-site research: Would it be possible to find researchers willing to undertake a lengthy collaborative effort? Would it be possible for individual researchers to place their research interests secondary to the general shared goals? Would it be possible to achieve uniformity

across sites so as to consider the study a replication across sites? Would it be possible to maintain the interests and commitment of the research groups through such a lengthy enterprise? Would it be possible to analyze, interpret, and write up findings in a collaborative and mutually satisfactory fashion? Would the collaborative multi-site model prove worthwhile and superior to independent, single-site studies regarding drawing inferences about the effectiveness of treatments?

In some areas of research multi-site studies have become commonplace. However, because of their expense - in terms of time, money, and staff - as noted by Fuller, et al (1994) multi-site studies are generally undertaken only after preliminary data or policy considerations indicate the need for large, representative samples. As reported by Fuller, et al, (1994) funding has followed this trend. The United States National Cancer Institute and the National Eye Institute fund multi-site studies almost solely and the National Heart, Lung, and Blood Institute allocates about half of its clinical trial resources to multi-site studies. Other United States' National Institutes of Health units, such as the National Institute of Mental Health and the National Institute on Alcohol and Alcoholism, are increasingly encouraging cooperative or collaborative research which typically involves multiple sites. As noted by Elkin multi-site, large scale, collaborative research is not new to the field of mental health treatment research. This type of research has been called collaborative clinical trials. While collaborative clinical trials have been common in the field

of psychopharmacology and various areas of medicine for many years their use in psychotherapy and in social work research is relatively new (Elkin, I, 1992). Also, as noted by Fuller, et al (1994) in their discussion of Project MATCH, multi-site clinical trials have been used in medicine and psychiatry for some time, but only recently have they been used in alcoholism treatment research.

Differences between multi-site and single site research

What differentiates single-site and multi-site research? The most obvious distinction arises from geography. Single-site studies are carried out in one geographical location. Multi-site research is carried out in more than one site.

Single-site research

In single-site research units for study are drawn from that location. Generalization is achieved through replication at the same site using different samples or subpopulations. Also, generalization and external validity is achieved through subsequent replication at other geographical sites. Replication studies attempt to use identical or, at least, similar methods. Replications focus on the same questions and variables as those in the original study. However, subsequent single-site replications are often redesigned. As a result of the redesign original research questions as well as original study variables are most often modified based on what was learned in

prior single-site applications. Subsequent single-site studies may be better thought of as elaborations or specifications of prior studies in a research program, since the aim is most often not replication but application with different sub-samples, settings, organizational contexts, and so forth. This latter objective is pursued in an attempt to advance generalization. Indeed, such programs of progressive single-site studies may suffer from a frequent failure to replicate because of the continual modifications that occur along the way. In addition, while modern methods of meta-analysis have been developed which permit the aggregate analysis of data from multiple single-site studies, when such single-site studies are combined in meta-analytic reviews, less certainty is achieved than would be the case with true concurrent replications resulting from multi-site investigations.

Single site evaluation studies often suffer from small sample sizes due to the constraints of how many subjects are available at any given geographical location.³ For example it is my impression that this may be the case in Sweden wherein studies conducted in one municipality may have a built-in constraint on sample size. Unfortunately, replications of single-site studies at other locations often face the same constraint due to small samples used in subsequent replications. The repeated use of a series of small samples does not overcome problems inherent in small sample size research.

3 This is often the case for evaluation and intervention research studies. However, survey research studies often include large samples.

Multi-site research: A working classification of multi-site research

Multi-site studies use more than one geographical location within the context of a single study. Unlike single-site studies, multi-site studies can increase sample size by drawing from more than one location. The number of locations can be influenced by statistical power considerations.⁴ Accordingly, to achieve a desired level of statistical power, projections can be made regarding how large a sample is required, how many subjects can be made available in each site, and, accordingly, how many sites would be required. This is an important advantage over single-site studies, each one of which is limited by the constraints of the site.

Also, unlike single-site studies, multi-site studies can provide for replication, if that is the intent. Multi-site studies that include the replication aim can use identical or near identical procedures at various sites thus minimizing possible effects of procedure variation on outcomes. Furthermore, when designed so that the research is conducted concurrently at the various sites, confounding temporal differences can be minimized. Of course, if the purpose is replication at multiple sites, efforts can be made to assure that all other site related conditions are similar such as sample characteristics. In the case of intervention studies multiple site replications

⁴ Statistical power is necessary to detect hypothesized associations between variables when using inferential statistical methods.

would need to include special provisions to assure that the interventions administered in the sites are similar.

Sometimes in multi-site studies the intent is, not to replicate, but rather to examine variations across sample characteristics or procedures so as to facilitate elaboration and specification. In these cases, for example, sites with different populations could be sought so as to study an intervention's effects with populations of different ethnic composition.

Three forms of multi-site research

While there are no classifications of multi-site research that have become generally accepted, Elkin notes that at least three models of multi-site research exist (Elkin, I, 1992). These models reflect varying roles that individual researchers may take relative to central decision making. Elkin describes these models as: (1) centrally designed and coordinated models with sites competitively selected through peer review wherein collaborators are variously involved in decision making; (2) models wherein researchers decide on their own to collaborate, jointly designing a common procedure at the outset; (3) models wherein researchers use a common data base but do not collaborate in implementation of the study.

I find it useful to distinguish among three forms of multi-site research, namely, simple multi-site research; complex multi-site research; and, multi-site allocation research. Next I briefly describe each of

these types providing examples drawn from our experiences at the CSSWP as well as from other research conducted in the United States. Of necessity these descriptions are brief, highlighting multi-site dimensions only. Furthermore, no attempt is made to present study progress or findings. The interested reader is referred to the cited references for further information and findings pertaining to each study.

Simple multi-site research

Often the same research questions and procedures are used across sites. I will refer to these as simple multi-site studies. These differ from what I term complex multi-site studies which address different research questions and/or may use different procedures at the sites studied. I provide two examples, the first provides somewhat more detail since it is the first example of multi-site research and it is taken from our experiences at the CSSWP. The second is a well-known multi-site study funded by the U.S. National Institute of Mental Health (NIMH).

From Research to Practice

From Research to Practice is an example of a simple multi-site study recently completed at the CSSWP in collaboration with investigators at the New York State Psychiatric Institute (NYS-PI). It is an example of a simple multi-site study because a single set of research aims, questions, hypotheses and a common procedure were applied at all research sites. A multi-site rather than

a single-site study was indicated so as to increase sample size beyond what could be achieved at any one site and so as to increase the geographic and demographic diversity of the patient population sampled.

The study was directed by a Principal Investigator together with two co-investigators at the coordinating unit, namely the NYS-PI. NIMH funding was awarded to the Principal Investigator at the NYS-PI through the New York State Research Foundation for Mental Hygiene.⁵ The study was designed by the NYS-PI research team. This team carried ongoing responsibility for project coordination, implementation, data analysis and reporting. The original plan called for subcontracts to a New Jersey site research team as well as to the New York City based CSSWP research team. The New Jersey site was composed of one outpatient mental health clinic. The CSSWP conducted the

⁵ This report is based on research conducted under subcontract to the Center for the Study of Social Work Practice - CSSWP (New York State Research Foundation for Mental Hygiene contract #SDMHCU00642601). The multi-site study was initially funded by the National Institute of Mental Health Grant #1R01MH052822-01 (02) (03). The Principal Investigator at the New York State Psychiatric Institute was David Shaffer, M.D. and the Co-Investigators were Prudence Fisher, Ph.D. and Christopher Lucas, M.D. The investigators for the Center for the Study of Social Work Practice subcontract were: Principal Investigator Edward J. Mullen, D.S.W.; Co-Investigators Robert Abramovitz, M.D., William Bacon, Ph.D. and Bruce Grellong, Ph.D. Prior investigators included Helene Jackson, Ph.D. and Jennifer Magnabosco, Ph.D.

research for the New York City sites. The New York City sites included eight community-based outpatient mental health clinics operated by the Jewish Board of Family and Children's Services (JBFCS). These clinics are located in four of the five boroughs of New York City.

Each subcontractor had a designated Principal Investigator and research staff. Each subcontractor carried responsibility for project implementation at their respective sites including human subjects review and approval, procedure administration, and data gathering. Publication of study findings pertaining to a site was the joint responsibility of the NYS-PI Principal Investigator and the individual subcontract site Principal Investigator.

From Research to Practice examined how a computerized, mental health, diagnostic assessment instrument, the Computerized Diagnostic Interview Schedule for Children (C-DISC), originally developed for use in epidemiological research affects clinical practice when used with child and adolescent outpatients. The C-DISC is a lay administered, computerized interview based on the American Psychiatric Association's Diagnostic and Statistical Manual for Mental Disorders, version IV (DSM-IV) and the tenth edition of the World Health Organization's (WHO) International Classification of Diseases (ICD-10). As originally planned, the evaluation was to use an experimental design. In each of the clinics data would be gathered during a prospective period detailing each clinician's normal assessment practice during routine practice prior to any experimentation. Following this prospective period, in each clinic,

half of the clinicians were to be randomly assigned to an experimental exposure to the C-DISC, and the other half would continue with routine practice completing a simple data gathering checklist. At the end of a year clinicians in each of the two conditions were to be assigned to the alternate condition in what is called a crossover design (Mullen, 1998; Mullen, et al., in preparation).

NIMH Treatment of Depression Collaborative Research Program (simple - intervention)

The NIMH Treatment of Depression Collaborative Research Program is another example of a simple multi-site study. The study was conducted simultaneously at several research sites (University of Pittsburgh, George Washington University in Washington, D.C. and the University of Oklahoma), addressing a common set of research questions and using a common research procedure. The study involved the collaborative efforts of these research sites, as well as additional training sites, NIMH and the Veterans Administration Data Analysis Facility (Veterans Administration Hospital, Perry Point, Maryland, USA).

The study used an experimental design in which treatments were randomly assigned to subjects. The basic purpose of this NIMH funded collaborative research was to assess the efficacy, efficiency, and safety of two well-defined, short-term psychological approaches, cognitive/behavior therapy and interpersonal psychotherapy, for outpatient treatment of non-bipolar,

non-psychotic depression. These two psychotherapies were compared to a medication treatment previously shown to be effective for this study population. In addition a pill-placebo condition was included. Outcomes measured were symptomatology, general clinical status, and social functioning. These measurements were taken at various points including during treatment, at termination and at several follow-up points (Elkin, 1994).

Complex multi-site research

A multi-site study can have a common set of aims, but within that common set of aims the study can allocate different research questions to different sites. Accordingly, different procedures could be used in different sites, specific to the questions addressed at each site. Such efforts are considered multi-site studies because they are a planned component in a larger research undertaking which is focused on a specific set of aims. I will call these complex multi-site studies, referring to the level of complexity of the research questions and procedures used. Two recent examples of complex multi-site research are Matching Patients to Alcoholism Treatments (Project MATCH) and Mental Health Service Use, Needs, Outcomes, and Costs in Child and Adolescent Populations (UNOCCAP).

Matching Patients to Alcoholism Treatments - Project MATCH

Project MATCH is an example of complex

multi-site research. This was a multi-site client-treatment matching trial involving nine geographically diverse clinical research settings and one coordinating center. The study was funded by the United States National Institute on Alcohol Abuse and Alcoholism. The purpose of Project MATCH was to assess the benefits of matching alcohol dependent clients to three different treatments with reference to a variety of client attributes. This is an example of a complex multi-site study since two parallel but independent randomized clinical trials were conducted, one with alcohol dependent clients receiving outpatient therapy in five sites and one with clients receiving aftercare therapy following inpatient or day hospital treatment in four sites. Clients were randomly assigned to one of three twelve-week, manual-guided, individually delivered treatments: Cognitive Behavioral Coping Skills Therapy; Motivational Enhancement Therapy; or, Twelve-Step Facilitation Therapy. Clients were then monitored over a one-year post-treatment period. In addition a three-year follow-up study was conducted. Individual differences in response to treatment were evaluated for ten primary matching variables and sixteen contrasts specified a priori. The primary outcome measures were percent days abstinent and drinks per drinking day (Project MATCH Research Group, 1997).

Mental health service use, need, outcomes, and costs in child and adolescent populations (UNOCCAP)

Another research program was titled

Multi-site Study of Mental Health Service Use, Need, Outcomes, and Costs in Child and Adolescent Populations but referred to as the UNOCCAP study. Although the UNOCCAP research program was terminated following the developmental phase, it serves as a good example of large scale, complex multi-site research. In 1994 the NIMH invited cooperative agreement applications for a five-year study of child and adolescent mental health services. Applications were invited for two types of studies. One type involved multi-site, collaborative, longitudinal, community-level studies of the types and patterns of mental health service use by children and adolescents, the extent of unmet need for services, and the ways in which the organization and financing of services influence access to, use of, and outcomes of mental health services. The other type of study involved a national survey to address issues related to the prevalence and incidence of specific mental disorders among children and adolescents, rates of mental health service utilization across major service sectors, and costs and financing of care. The UNOCCAP initiative encouraged and required the collaboration of multidisciplinary research teams at both the community and national sites; and, it sought to enable independent teams of investigators to work together to develop common study procedures. The National Institutes of Health (NIH) cooperative agreement mechanism (U01) funded this research program. The National Institute of Health (NIMH) staff worked jointly with the awardees in a partnership role, to support, coordinate, and facilitate the awardees' activities, and to assist in moving the

study through its phases. Direction and principal responsibility for the conduct and implementation of the study remained with the awardees.

The UNOCCAP participating sites were the Johns Hopkins University, University of California at Los Angeles, University of Chicago, and Washington University, with additional collaboration from the Research Triangle Institute, Rand Corporation, Vanderbilt University, Yale University, and the NIMH staff. The NIMH budgeted \$45 million dollars over five years for the project, which also received support from the United States Administration for Children, Youth and Families; the United States National Institute of Child Health and Human Development; the United States Substance Abuse and Mental Health Services Administration; the United States Department of Education; and, the MacArthur Foundation.

UNOCCAP was originally conceptualized as providing both a national probability sample for determining the prevalence and incidence of specific disorders in children, and as a set of community level studies to address the types and patterns of mental health service use by children and adolescents, the extent of unmet need for services, and the financing of these services. During the two-year developmental phase of the UNOCCAP study, the participating researchers designed a nationwide household sample of approximately 10,000 children and adolescents, and made major efforts in instrumentation development. The collaborators also planned to assess additional samples of children in both outpatient, such as specialty mental health

and school-based services, and inpatient/residential services.

The study underwent scientific review in 1997 by an Oversight Board Appointed by the NIMH Director. Based on the Oversight Board's recommendations, the NIMH Director decided not to carry out the UNOCCAP study. I return to this study in my concluding discussion since it illustrates important issues pertaining to the conduct of multi-site research (National Institute of Mental Health, 1998).

Multi-site allocation research

Multi-site allocation research differs sharply from simple and complex multi-site research. In this third form of multi-site study the geographical unit is not used as the location-site for the study, but rather the geographical unit is the object of study. Accordingly, Boruch and Foley describe »sites and other entities, rather than individuals, as the units of allocation, treatment and analysis« in randomized trials (Boruch & Foley, in preparation, abstract). The sites they describe are geographical locations used as units of allocation. Other examples of allocation units are families, communities, and organizations. Their focus is on experimental research in which sites or other units are randomly allocated to differing interventions so as to study causal associations. However, sites could also be studied in non-randomized research as well as in randomized research. I will refer to this form of multi-site study, whether using randomized or non-randomized designs, as multi-site allocation research. Randomized

allocation studies are fully discussed by Boruch and Hedges in another article in this issue. Next I provide one example from our work at the CSSWP, namely the Sanctuary study.

Trauma Focused Intervention Targeting Risk for Violence (Sanctuary)

This multi-site intervention study is being conducted by the CSSWP at twelve JBFC residential treatment programs for children and youth operated in Westchester County, New York. The twelve units are being randomly assigned to either an innovative milieu treatment or to a standard residential treatment condition without the innovative enhancement. This multi-site research is examining the implementation and proximal effects of an intervention designed to reduce trauma-related symptoms of youths that place them at high risk for violent behavior, poor adjustment, and serious mental health difficulties.

The aims will be achieved by using a 2 x 5 design. Two service delivery conditions are provided, namely the experimental Sanctuary Model enhanced milieu treatment or the Standardized Residential Services. Measurement of outcome variables will be taken at five data collection points, namely at baseline, three months, six months, nine months, and 12 months. The twelve residential units have been randomly assigned to either the Sanctuary Model or Standard Residential Services. The twelve residential units serve one-hundred and fifty youths and have ninety-six staff. Implementation and effects of the model will be measured

at the provider level (i.e., perceptions of change in the therapeutic environment, changes in interaction patterns between staff and youths) and at the youth level (i.e., perceptions of change in the therapeutic environment, change in youths' behavior and skills).

This is a collaborative, multi-site study involving the CSSWP, the JBFC Westchester facilities, the JBFC Center for Trauma Program Innovation, the Columbia University School of Social Work and the Columbia University New York State Psychiatric Institute. The study Principal Investigator is an affiliate of the CSSWP and a faculty member at the CUSSW. Co-Investigators include members from each of the collaborating organizations. The research is funded by the NIMH (Rivard, 2000).

Two types of multi-site research

In addition to these three forms of multi-site research, two broad types of multi-site research can be specified. The two types of multi-site research are: (1) multi-site intervention research; and, (2) multi-site survey research. These are described next with examples of each provided.

Multi-site intervention research

In multi-site intervention research the aim is to study and draw conclusions about intervention programs. Examples of multi-site intervention research are the previ-

ously described From Research to Practice, the NIMH Treatment of Depression Collaborative Study, Project MATCH, and the Sanctuary study. Examples of multi-site survey research are provided next.

Multi-site survey research

A second type of multi-site research has as its aim description of populations, using survey methods. Both probability sampling and non-probability sampling methods can be used. I will call this second type of multi-site research multi-site survey research. In survey research multiple sites are used for sampling to increase representativeness and to increase sample size. This is a well-known application of multi-site research with well-developed methodologies.

Probability sampling in multi-site research

Four examples of multi-site survey research are provided next including three that use probability sampling methods and one that uses non-probability methods. Both simple and complex multi-site research examples are provided. Examples include studies conducted by the CSSWP and as well as through NIMH funded research.

The Patient Profile Study

Multi-site survey research can use probability sampling methods to address research questions. Cluster sampling methods have long been used in survey research. These clusters can be geographi-

cal sites. An example of such a multi-site study is the Patient Profile Study. This is a simple multi-site study using probability sampling survey methods. This survey was conducted by the CSSWP. The geographical sites sampled were a number of JBFCs operated community mental health clinics located in four New York City boroughs. The research aim was to describe the child and adult population of patients who came to these clinics for mental health services. A sample of approximately 20% of adult and child clients were randomly selected from these clinics, stratified on the basis of clinic, ethnicity, and age. Accordingly, findings were specified by site as well as ethnicity and age. Because multiple sites were used conclusions were drawn pertaining to each clinic as well as to the total JBFCs clinic population (Mattaini, M. A., Grelong, B. A., & Abramovitz, R. (1992).

Epidemiologic Catchment Area Study (ECA)

The Epidemiologic Catchment Area Study (ECA) is a well known example of multi-site survey research. It was a simple multi-site study in the sense that common research questions and procedures were used across sites. The study used multistage probability sampling drawing from five geographical locations. The purpose of the ECA research was to collect data on the prevalence and incidence of mental disorders and on the use of and need for services by the mentally ill. Independent research teams at five universities (Yale University, Johns Hopkins University, Washington University, Duke University, and University of California at Los Angeles), in collaboration with the

National Institute for Mental Health, conducted the studies with a core of common questions and sample characteristics. The sites were areas that had previously been designated as Community Mental Health Center catchment areas. Each site sampled over 3,000 community residents and 500 residents of institutions, yielding 20,861 respondents overall. The longitudinal ECA design incorporated two waves of personal interviews administered one year apart and a brief telephone interview in-between (for the household sample).

While the ECA used probability methods to sample within each of five sites, these sites can not be assumed to be representative of the United States population (Robins & Regier, 1991). To address this limitation another NIMH funded study, the National Co-morbidity Survey (NCS), was subsequently conducted drawing a probability sample from the total United States population (Kessler, et al., 1994). Currently, the NCS is being replicated.

Methods for the Epidemiology of Child and Adolescent Mental Disorders Study (MECA)

This NIMH funded study is an example of a complex multi-site survey using probability sampling methods. The MECA collaborative study was conducted to develop methods for surveys of mental disorder and service utilization in unscreened population-based samples of children and adolescents. Probability household samples of youths were selected at four sites and interviews were conducted with a total of 1,285 pairs of youths and their adult caretakers in their homes. Lay interviewers administered

a computer-assisted version of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC) and structured interviews to assess demographic variables, functional impairment, risk factors, service utilization, and barriers to service utilization. More than 7,500 households were enumerated at four sites. Since sites varied in terms of procedures used and research questions examined this can be considered a complex multi-site survey.

Non-Probability sampling in multi-site research

Non-probability survey research methods can also be used in an attempt to describe a population. Multiple sites frequently serve as the source for such research. An example from our work at the CSSWP is provided next.

Odyssey Project

The Odyssey Project is an example of a simple multi-site survey using non-probability sampling methods. The Odyssey Project is a collaborative multi-site study in which the CSSWP is a participant. The Odyssey Project is a descriptive and prospective study of children in residential group care, group homes, and therapeutic foster care in the United States. The Child Welfare League of America, in cooperation with and support of its members, is conducting this national, multi-site, descriptive and prospective study of children and youths in residential care. The purpose of the descriptive phase is to determine what types of settings and services are serving what kinds of children and youths.

The purpose of the prospective phase is to determine what types of settings and services are related to what outcomes with what kinds of children and youths. The research questions were designed to explore the impact of residential group care, group homes, and therapeutic foster care on children and youths in care. Children who enter care between April, 1995 and July, 1999 were eligible to participate in the project. Twenty-eight agencies from 15 states participated providing approximately 3,100 children and youths to the study sample. Many of these participating agencies included multiple facilities so that the actual number of distinct sites was larger than 28. Accordingly, in this example multiple facilities (sites) are nested within agencies which in turn are nested within the larger CWLA study.

Oftentimes multi-site studies are referred to as »collaborative« in nature. This is the case with the Odyssey Project. As collaborating partners in the research, member agencies had clearly defined rights and responsibilities.⁶

Issues in multi-site studies

In identification of issues in multi-site studies I draw from several sources. At the outset I wish to give full attribution to these sources for the insights provided. I have been struck by the common observations about issues in multi-site research found in these sources. Elkin's paper presents her personal reflections on experiences with the previously described NIMH Treatment of Depression Collaborative

Research Program (Elkin, 1992). The papers authored by members of the previously described Project MATCH research team identify issues in the context of that multi-site study (Carroll, et al., 1998; Del Boca, et al., 1994; Fuller, et al., 1994; Zweben, et al., 1994). The observations of the previously described UNOCCAP Oversight Board are based on a careful review of early experiences with that research program (National Institute of Mental Health, 1998). Finally, as Director of the CSSWP I have drawn upon my own experiences with the many multi-site studies conducted by CSSWP (Practice & Research, Spring 2000). Given the limits of space available in this article I will only outline a number of issues.

Advantages of Multi-site Research

The advantages of multi-site research have already been mentioned and they are clear. The ability to draw subjects from multiple sites can increase sample size. Accordingly, a study's statistical power is increased. Also, multiple sites can increase the sample's diversity on variables of interest. An additional advantage of multi-site research pertains to replication. Properly designed multi-site research can provide for simultaneous replication limited only by the number of comparable sites included in the study. In addition to these advantages Elkin notes two additional benefits. One pertains to the potential for increasing resources, both financial and in terms of expertise. The potential availability of multi-site team members with a range of methodological, statistical and clinical

expertise can be invaluable. Associated with this is the benefit of collaborative decision making, resulting in enriched discussion and improved outcomes. As noted by Fuller, et al., in multi-site research there is also the potential to create a centralized unit with resources for core functions such as data analysis.

Disadvantages of Multi-site research

If not properly managed there are a number of potential disadvantages associated with multi-site research. Elkin, Fuller, et al, Del Boca, and Zweben describe potential disadvantages:

- Multi-site studies can be costly and methodologically complex.
- The process of joint decision making, consideration of procedural and methodological problems and joint resolution of these problems requires considerable time and effort.
- Idiosyncratic site effects can create problems in management as well as in data analysis and interpretation.
- Logistic difficulties can be problematic when implementation spans across multiple sites. Examples include allocation of staff across sites, communication among geographically dispersed team members, and so forth.
- Statistical issues can be problematic such as how best to combine data collected from multiple sites. Site effects need to be addressed in the statistical analysis.
- Addressing new requirements after

- the study has begun can be troublesome. For instance sites can experience policy and fiscal changes that may impact on the study requiring procedural adjustments. When there are many sites these changes can be difficult to track in a timely way.
- Issues of reliability of data collection across sites are often problematic. Sites may implement data collection differently creating difficulties in maintaining uniform procedures.
 - Issues of subject recruitment and eligibility criteria in multi-site studies can be complex. Again, these issues can arise from a lack of uniformity among the sites as well as from the logistical difficulties of monitoring recruitment practices among many sites.
 - Redundancy of staffing across sites can increase cost unnecessarily. Careful planning is required to assure that staff is used efficiently across sites.
 - There can be problems establishing and maintaining cooperation and commitment across sites since no one site »owns« or is totally responsible for the research. Accordingly, in many multi-site studies individual sites have limited ownership and get limited credit.
 - Multi-site studies oftentimes experience difficulty maintaining cross-site uniformity of procedures. Maintaining consistency across sites can be problematic (e.g., similarity of treatment implementation across sites, of data collection, etc). There is greater difficulty in maintaining integrity of the treatment and research procedures across sites in multiple locales.
 - Mechanisms need to be established for how the data will be analyzed and the findings published that are protective of individual site interests as well as of the multi-site collaborative effort as a whole.

References

- Allen, J., Anton, R.F., Babor, T.F., Carbonari, J., Carroll, K.M., Connors, G.J., Cooney, N.L., Del Boca, F.K., DiClemente, C.C., Donovan, D., Kadden, R.M., Litt, M., Longabaugh, R., Mattson, M., Miller, W.R., Randall, C.L., Rounsaville, B.J., Rychtarik, R.G., Stout, R.L., Tonigan, J.S., Wirtz, P.W., & Zweben, A. (1998). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcohol Clin Exp Res*, 22 (6): 1300-1311 Sep 1998.
- Allen, J.P., Mattson, M.E., Miller, W.R., Tonigan, J.S., Connors, G.J., Rychtarik, R.G., Randall, C.L., Anton, R.F., Kadden, R.M., Litt, M., Cooney, N.L., DiClemente, C.C., Carbonari, J., Zweben, A., Longabaugh, R.H., Stout, R.L., Donovan, D., Babor, T.F., Del Boca, F.K., Rounsaville, B.J., Carroll, K.M., Wirtz, P.W., Bailey, S., Brady, K., Cisler, R., Hester, R.K., Kivlahan, D.R., Nirenberg, T.D., Pate, L.A., Sturgis, E., Muenz, L., Cushman, P., Finney, J., Hingson, R., Klett, J., & Townsend, M. (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, 58 (1): 7-29 Jan 1997.

- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders. Washington, DC: Author.
- Anon (1998). Matching alcoholism treatments to client heterogeneity: Treatment main effects and matching effects on drinking during treatment. *J Stud Alcohol*, 59 (6): 631-639 Nov 1998.
- Bacon, W.F. (2000). From research to practice: The C-DISC in clinical services. *Practice & Research*, Spring 2000, 26-28.
- Blazer, D., D. Huges, & George, L.K. (1987). The epidemiology of depression in an elderly community. *The Gerontologist*: 27 (1987), 281-287.
- Boruch, R. & Foley, E. (in preparation). The honestly experimental society: Sites and other entities as the units of allocation and analysis in randomized trials. In Bickman, L. (ed.). *Validity and social experimentation: Donald T. Campbell's legacy*. Thousand Oaks, Ca: Sage.
- Carroll, K.M., Connors, G.J., Cooney, N.L., DiClemente, C.C., Donovan, D.M., Kadden, R.R., Longabaugh, R.L., Rounsaville, B.J., Wirtz, P.W., & Zweben, A. (1998). Internal validity of project MATCH treatments: Discriminability and integrity. *Journal of Consulting and Clinical Psychology*: 66 (2): 290-303 Apr 1998.
- Del Boca, F.K., Babor, T.F., McRee, B., (1994). Reliability enhancement and estimation in multi-site clinical-trials. *J Stud Alcohol*: 130-136 Suppl. 12 Dec 1994.
- Eaton, W.W., et al. (1985). The epidemiologic catchment area program of the National Institute of Mental Health. *Public Health Reports*: 96 (1981), 319-325.
- Eaton, W.W., & Kessler, L. (eds) (1985). *Epidemiology field methods in psychiatry: The NIMH epidemiologic catchment area program*. Orlando, FL: Academic Press, Inc., 1985.
- Elkin, I. (1992). Multi-site studies: Advantages and disadvantages. In Gerin, P & Dazord, A., *Recherches Cliniques "Planifees" sur Les Psychotherapies: Methodologie*. Paris: INSERM, 1992.
- Elkin, I. (1994). The NIMH treatment of depression collaborative research program: Where we began and where we are. In Bergin, A.E. & Garfield, S.L., *Handbook of psychotherapy and behavior change*. New York: John Wiley and Sons, Inc.
- Fuller, R.K., Mattson, M.E., Allen, J.P., et al. (1994). Multi-site clinical-trials in alcoholism-treatment research – organizational, methodological and management issues. *J Stud Alcohol*: 30-37 Suppl. 12 Dec 1994 .
- Guterman, N.B. (2000). The Odyssey project and community violence exposure among children & youth living in residential treatment settings. *Practice & Research*, Spring 2000, 23-25.
- Journal of the American Academy of Child and Adolescent Psychiatry* (1996). Special Section on Epidemiology of Child and Adolescent Mental Disorders. 35:7.
- Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H. & Kendler, K.S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Archives of General Psychiatry*, 51, 8-19.
- Lahey, B.B., Flagg, E. W., Bird, H.R., Schwab-Stone, M.E., Canino, G., Dulcan, M.K., Leaf, P.J., Davies, M., Brogan, D., Bourdon, K., Horwitz, S.M., Rubio Stipeck, M., Freeman, D.H., Lichtman, J.H., Shaffer, D., Goodman, S.H., Narrow, W.E., Weissman, M.M., Kandel, D.B., Jensen, P.S., Richters, J.E., & Regier, D.A. (1996). The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) study: Background and methodology. *Journal of the American Academy of Child and Adolescent Psychiatry*: 35 (7): 855-864 Jul 1996.
- Mattaini, M. A., Grellong, B. A., & Abramovitz, R. (1992). The Clientele of a Child and Family Mental-Health Agency - Empirically Derived Household Clusters and Practice Implications. *Research on Social Work Practice*, 2(3), 380-404.
- Mullen, E.J. (in preparation). From research to practice: A preliminary report of the C-DISC on dimensions of clinician, patient and caretaker satisfaction (presented as an unpublished paper under title of: "Using assessment instruments in social work practice«. October

- 7-8, 1999, Inter-Centre Consortium of Social Work Research Centers, Stockholm, Sweden). In preparation as an on-line electronic pre-print to be available: London, UK: National Institute for Social Work, <http://www.intsoceval.net/Stockholm/stockholmpapers.htm>.
- Mullen, E.J. (2001). Outcome measurement in social work: health and mental health. (plenary session, 3rd International Conference on Social Work in Health and Mental Health, July 1-5, 2001, Tampere, Finland; [On-line electronic pre-print]. Available: London, UK: National Institute for Social Work, <http://www.intsoceval.net/Utrecht/utrechtpapers.htm>.
- Mullen, E.J. (1998). Linking the university and the social agency in collaborative evaluation research: Principles and examples. *Scandinavian Journal of Social Welfare*. nr2.
- Mullen, E.J. and Magnabosco, J. (eds) (1997). *Outcomes measurement in the human services*. Washington:D.C.: National Association of Social Workers.
- National Institute of Mental Health (1994). Multi-site study of mental health service use, need, outcomes, and costs in child and adolescent populations. NIH Guide: vol 23, number 15, April 15, 1994 RFA.
- National Institute of Mental Health (1998): Report from the UNOCCAP oversight board to the National Advisory Mental Health Council: Charting the mental health status and service needs of children: Recommendations from the UNOCCAP oversight board.
- Practice & Research: Journal of the Center for the Study of Social Work Practice (Spring 2000).
- Project MATCH Research Group (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol* 58: 1, 7-29. (also see: *Journal of Studies on Alcohol*, Supplement No. 12, December 1994.)
- Regier, D. A. (1994). ECA contributions to national policy and further research. *International Journal of Methods in Psychiatric Research*. 4, 73-80.
- Rivard, J.C. (2000). Evaluating the implementation and impact of an intervention designed to enhance residential treatment for adolescents with histories of trauma. *Practice & Research*, Spring 2000, 33-36.
- Robins, L. N., & Regier, D. A. (Eds.) (1991). *Psychiatric disorders in America: The Epidemiologic Catchment Area Study*. New York: The Free Press.
- Zweben, A., Donovan, D.M., Randall, C.L., et al. (1994). Issues in the development of subject recruitment strategies and eligibility criteria in multi-site trials of matching. *J Stud Alcohol*: 62-69 Suppl. 12 DEC 1994.

Summary and Conclusions

This article has examined multi-site research. Differing from single-site research, multi-site research is appropriate for reasons that distinguish it from single site research. In this article I have examined forms and types of multi-site research to illustrate a variety of applications. I have described some examples of multi-site research at the national level with reference to mental health services and alcoholism

treatment research applications. Drawing from my experiences as director of research at a social work research center I have described additional applications in mental health, child welfare, and social work practice. I have outlined some advantages and disadvantages of multi-site research. In concluding this analysis I provide suggestions for those considering engaging in multi-site research. While

there are many matters to consider I limit my comments in this concluding section to three areas that strike me as most often neglected and of particular importance to successful multi-site research. Since multi-site research can be expensive, involving relatively more resources and time than single-site research, it should not be undertaken unless preconditions are met. Several of these preconditions are described next.

Concepts and methods previously developed and tested

An important precondition is that preliminary studies should have shown that key research questions are conceptually sound and that the research methods needed to examine these questions have been adequately developed. For example, the ambitious UNOCCAP research program was undertaken after extensive methodological work had been done in the MECA study (described above). Nevertheless, the UNOCCAP research program was halted by NIMH because it was determined that the original UNOCCAP aims and research questions were too broad, unevenly developed, and too extensive for any one study to address. The NIMH Review Board concluded:

»--- no one study can address each of these questions well simultaneously. There is no way to combine these questions into one method or design without sacrificing the quality of data. --- Some areas require additional conceptualization, better instrumentation, or more empirical work to generate or

test hypotheses. This critical basic research must be conducted to lay the groundwork for answering all of these questions. --- More important, science is not ready for such an approach. Quite simply, the necessary conceptualization, tools, and designs are not uniformly available in all substantive areas. If resources were diverted into prematurely conducting such an immense and elaborate effort, the Board is concerned that it would generate data that policymakers should not use and that a substantial portion of the researchers would not accept as credible.« (National Institute of Mental Health, 1998, 2-4)

In addition to the lessons learned from the UNOCCAP experience additional lessons can be learned from the above described From Research to Practice study. In that study unusual problems were encountered securing the cooperation of sites as well as in engaging adequate numbers of clinicians and patients. Although during the planning stage the site central administration was supportive of the study, problems pertaining to motivation and commitment were encountered during study implementation at the local level. Also, while preliminary analysis indicated that sites had sufficient numbers of clinicians and clients, during implementation it was difficult to engage adequate numbers of both. In this example it could be argued that a multi-site study was premature. Rather, pilot testing the intervention and the research methodology at one site may have been prudent before launching an ambitious multi-site investigation.

Resources in place

A second precondition for multi-site research is that necessary resources should be secured before undertaking multi-site studies. Multi-site studies require adequate resources to support a central core as well as to support site-specific activities. The central core can carry out common functions such as data processing which can relieve sites of this responsibility. Basic funding and adequate personnel are required to make sure that all program and research functions can be properly implemented throughout the life of the study. However, often overlooked in planning of multi-site studies is the necessity of providing financial resources and qualified staff for the »relationship« side of multi-site research. Ongoing, frequent communication among all parties is essential so that commitment, motivation, and day-to-day problem solving can occur smoothly. While necessary in single-site research this aspect of multi-site research is critically important (Mullen, 1998).

Governance clarified and agreed to by collaborators

A third precondition is that a clearly defined and agreed-to governance and management structure and process be developed before implementation of multi-site research. Because of the complexity of such studies and because of the unusual nature of study »ownership« the rules and procedures should be clear from the outset. Differential responsibilities of the central coordinating team, the site teams, practi-

tioners and researchers, program administrators, the funders, and other partners need to be specified and agreed-to. Since there are different models for multi-site governance consideration should be given to which model best fits the circumstances of a given multi-site study.

If these (and others not addressed) conditions are met then multi-site research should be considered if it is determined that single-site research cannot provide an adequate sample size, the desired sample heterogeneity, or an adequate basis for drawing conclusions regarding a population of interest. Multi-site research should also be considered when concurrent replication is desired. As discussed elsewhere in this issue by Boruch, et al. when geographical sites or other complex units are the units of study, allocation multi-site research should be considered.

Finally, it is important for the success and development of future generations of multi-site research that a solid and detailed literature be generated regarding the planning, implementation, analysis and reporting of multi-site research. Researchers engaged in multi-site studies need to take responsibility for contributing to this literature by reporting not only multi-site findings but also what they have learned about the conduct of multi-site research. Currently, with the exception of survey research, little has been written about the methodologies of multi-site research as applied to social interventions.