Family-Genetic Studies of Psychiatric Disorders

Developing Technologies

Myrna M. Weissman, PhD; Kathleen R. Merikangas, PhD; Karen John; Priya Wickramaratne, PhD; Brigitte A. Prusoff, PhD; Kenneth K. Kidd, PhD

- During the past decade new concepts and technologies have improved the conduct of family-genetic studies in psychiatry. We compiled and critically evaluated these advances, including study design, pedigree collection, diagnostic procedures in adults and children, and epidemiologic and genetic approaches to data analysis. These approaches have improved the collection of accurate information on the nature and patterns of psychiatric illness in families. The data generated from well-designed and well-conducted family studies are useful for the identification of homogeneous subgroups of psychiatric disorders, for understanding the spectrum of psychiatric disorders, for examining the associations between psychiatric disorders, and for studying the continuity between adult and childhood manifestations of psychiatric disorders. Findings from these studies also may enhance our capacity to identify the mode of transmission of the psychiatric disorders and to select potentially informative families for future genetic linkage studies using the new recombinant DNA techniques. The adaptation of these methods to routine clinical practice and new directions in the application of family-genetic studies employing more refined assessments and analytic methods are also discussed. (Arch Gen Psychiatry 1986;43:1104-1116)

The family has major importance in the dissemination of disease. It is the basic unit involved in the transmission of genes and of culture from one generation to the next. It also provides the immediate environment for the sharing of genes, diet, toxins, infections, stress, emotions, and social attachments that may interact to effect expression of underlying physical and behavioral traits.

The familial nature of human diseases has been recognized since biblical times. Diabetes has been noted to be hereditary since 1000 BC. Familial aggregation has also been observed for numerous other medical disorders such as hypertension, coronary heart disease, cancer, and neurologic disorders. Some now-classic studies (classic because they were the earliest well-designed studies) and many newer studies have also demonstrated that the major psychiatric disorders are familial, e.g., schizophrenia, depression, schizoaffective disorder, alcoholism, personality disorders, and anxiety disorders. Evidence of familial aggregation does not imply that the origin of a disorder is genetic. Aggregation could result either from shared genes or from common environmental factors such as infection, diet, stress, or social learning. It is both simplistic and erroneous to assume that only genetic or only environmental factors are involved in the etiology of the psychiatric disorders, for both undoubtedly contribute. Rather, understanding the degree to which each of the sources contributes and interacts to produce a given phenotype is the aim of genetic-epidemiologic studies. Study paradigms from which such information can be derived include adoption and twin studies, which have been extensively utilized to study the major psychiatric disorders.

Although family studies do not yield evidence of the amount of genetic variation, data from family studies in psychiatry can serve several purposes. One of the most fruitful has been to provide better understanding of diagnostic heterogeneity. It is quite likely that many of the psychiatric disorders are groups of conditions rather than single entities with different etiologic and modifying risk factors. The use of family data in the absence of specific neuropathological evidence is one approach to identifying homogeneous diagnostic subgroups and, thus, reducing diagnostic heterogeneity. If the diagnostic subgroup under study increases risk of the disorder and “breeds true” within families, potential evidence of the validity of the diagnostic group is suggested. One example is the work by Cloninger et al. in which family data were used to examine the relationship between alcoholism, antisocial personality, and hysteria. There are numerous other examples.
Because variation in expression of a particular trait within families is assumed to result from the same latent factor (ie, homotypy), family studies can yield information on variable expressivity of the gene or genes related to the phenotypic trait. This property of family studies can lead to the development of more precise clinical descriptions of the spectrum or the diathesis of disorders, as manifest in relatives' personality, symptoms, or social functioning. Studies of the young children of adult probands can yield information on the transmission of disorders and/or symptoms across generations, on the early signs and childhood forms of the disorder, and on the risk or protective factors that modify the development of the disorder.

The purpose of this article is to describe some of the recent advances in the methodology for the conduct of family studies of psychiatric disorders from a genetic-epidemiologic perspective. Approaches to identifying and assessing families, including design issues, systematic collection of pedigrees, standardized diagnostic procedures for adults and children, family history methods, and new statistical issues, will be discussed. General approaches to the design and concepts of epidemiologic or of genetic studies will not be described, nor will a general review of statistical methods be presented. There is excellent literature available for in-depth study of these topics. Kidd and Matthysee describe some of the general design strategies such as twin or cross-fostering studies to examine gene-environment interactions; others point out some of the obstacles to these studies.

**BASIC DESIGN**

Epidemiologic family studies can provide at least three types of evidence that suggest familial aggregation of a particular condition:

1. Relatives of cases have a higher prevalence of the condition than the population from which the case was selected.
2. The prevalence of the condition is higher among relatives of cases than among those of controls.
3. The distribution of a disorder within families is consistent with a known genetic pattern.

The first two types of evidence of familial aggregation of a disorder are derived from studies that employ a case-control, retrospective, or longitudinal cohort (high-risk study) design of a large number of families. The third type of evidence requires the study of extended pedigrees of a small number of selected families. These epidemiologic study designs and their variants are described in greater detail by Kleinbaum et al. In psychiatry, the following designs have most commonly been used.

**Case-Control Studies**

In a family case-control study, a proband with the illness under investigation is selected for study and is then matched with a control proband (ie, an individual without the illness under investigation but who is comparable on other characteristics). Usually, the prevalence of the condition among first-degree relatives is compared with prevalence among the relatives of controls. Earlier family studies often did not use control groups for the purposes of direct comparison. Rather, the rates of illness among relatives were compared with population rates. In either case, studies usually have a retrospective cohort design in that the lifetime rates of illness in relatives are obtained on the basis of the recall of lifetime course.

In family studies of psychiatric disorders, the probands, or index cases, with the disorder being investigated have nearly always been adults who were selected from treatment settings or psychiatric or case registries. Family studies that begin with the adult probands and study psychopathology among their offspring as well as other relatives have been termed **top down** by Puig-Antich.

With the increasing interest in childhood psychiatric disorders during the past decade, children have also begun to be defined as the probands in family studies, termed **bottom-up** studies by Puig-Antich and Strober and Carlson. Similar to the adult studies, children who serve as probands generally have been selected from inpatient or outpatient treatment settings.

Studies that begin with the child or adolescent as the proband, or index case, tend to find very high rates of illness in the adult relatives, possibly because of sampling bias. Although the proband is the treated child, it is the parent who brings the child for treatment and who grants permission for the child to be included in the study. Ill parents may be more likely than well parents to bring their children to treatment and to consent to the child's inclusion.

One method used to control for this ascertainment bias has been to select for comparison a control proband group of children with another treated psychiatric illness. The rates of all types of psychiatric illness will also tend to be high in the adult relatives of the child comparison group. The types of illnesses and the magnitudes of the differences in rates between the relatives of the cases and the relatives of the comparison control group can provide more important information than the absolute rates of illness in the relatives.

**High-Risk Studies**

The high-risk paradigm is a variant of the family study. The focus is usually limited to the young offspring of ill probands. The proband is a parent. Usually there is no assessment of the proband, and the outcome is based on the parent who is at high risk. High-risk studies are sometimes studied longitudinally to identify risk factors that are premorbid to, rather than concomitant with, the disorder or that are manifestations of it. Such factors may serve to identify vulnerable individuals and permit efforts toward prevention and intervention.

High-risk designs may be used to obtain information on the early signs and first episodes of the illness uncon founded by chronicity, the elucidation of factors that may protect vulnerable persons from the development of the disorder or minimize its impact, the minimization of sample size necessary to yield sufficient incidence rates for meaningful analysis, and the calculation of the absolute and attributable risk of factors that may be associated with the onset of the disease among offspring of cases compared with controls.

A major problem in the conduct of family studies of low-prevalence disorders, such as schizophrenia and bipolar illness, is that a very large sample of probands and relatives is required to achieve the necessary power to conduct meaningful data analyses, particularly with respect to the identification of subtypes of the disorder.

**Extended Pedigrees**

Studies of pedigrees beyond the first-degree relatives (ie, second- and third-degree relatives) of the probands have also been conducted in psychiatry. Multigenerational pedigree analyses can be used to determine if the distribution of the disorder within the family is consistent with any known genetic pattern.

**GENERAL PROCEDURES**

**Selection of Probands**

The aim of a family study, regardless of its design or the diagnostic group under study, is to collect detailed and unbiased information on the disorder of interest among
biological relatives, both living and deceased. The general procedure is to collect information on the psychiatric status of all biological relatives of a person with the disorder—the index case or proband—in a uniform manner.

With rare exceptions, probands in family studies of psychiatric disorders are patients receiving psychiatric treatment for the disorder being investigated. Since a majority of psychiatrically ill patients never receive treatment, a bias may be introduced into family studies in which probands are selected only from treated samples. For example, treated patients are generally among the more severely affected. If severity is related to familial prevalence, or pattern of illness, the selected families will not be representative of all those containing an affected individual.

Even more complex biases might occur. The proband's decision to seek treatment or to consent to participate in the study may be determined in part by how many and/or which relatives are affected. Such biases are very difficult to deal with statistically. As noted before, this source of bias may be especially troublesome in studies that employ young children as probands.

Control Group

There are several sources of control groups that may be utilized for comparison of the patterns of familial aggregation among the probands. Subjects who have the disorder, but who have never received psychiatric treatment, can be selected from primary care facilities or from community samples. Another method is to select a control group of patients who have a psychiatric illness other than that of interest and for which they have sought treatment. Control groups employed in previous family-genetic studies have included medically, but not psychiatrically, ill probands, normal, never psychiatrically ill controls, and controls with a psychiatric illness other than the one under investigation.

The purpose of the control group is to permit comparison of the rates of the disorder among relatives. Because many factors may contribute to the variance in the rates of psychiatric illness in families, data that are collected in an identical manner from the families of cases and controls can be used to estimate the relative risk for the disorder.

A matched control group design can be used to reduce further possible sources of bias. The guiding principle in this design is to match probands with the disorder to control probands without the disorder on confounding factors that are likely to affect the outcome, ie, the illness status in relatives. Because the age of the proband may affect the composition of the family structure, which in turn may affect illness rates in relatives, some consideration should be given to the comparability of probands with controls on age. Other factors that may have an impact on outcome are the proband's sex, social class, income, race, and ethnicity.

Relatives

As a first step, a systematic pedigree that identifies the spouse and all first-degree relatives (ie, parents, siblings, and offspring), their names, biological relationships, sex, date of birth, age, and year of death is obtained. Standardized methods of collecting and computer processing pedigree information for genetic analysis will be described.

Once a pedigree has been established and the family members have been identified, information on their clinical status is obtained by direct interview with the relative, supplemented by family-history data collected independently from multiple sources (eg, the other family members, health care providers, treatment records).

Ideally, the information from each relative should be obtained independently. Different interviewers should interview each relative and should be blind with regard to the clinical status of the proband. This procedure can be expensive, especially if relatives live at some distance. It may not always be feasible if the family members are reluctant to be interviewed or if the proband is reluctant to give consent to interview relatives. However, willingness to have family members interviewed should not be a criterion for including a proband in the study. The information obtained on each relative from multiple informants, clinical records, and direct interview is reviewed by an independent and experienced clinician who takes into account the sources and quality of the data, applies uniform criteria, and assigns a "best estimate" diagnosis to the relative.

The clinical status of the spouse (who is not a biological relative of the proband) must also be assessed, as the spouse's status can affect rates of illness in children. This information is necessary in the application of family data results for genetic modeling as well as for understanding the potential environmental impact of the family on the child.

In most family studies of psychiatric disorders, diagnostic information on offspring under 18 years of age has not been obtained. However, this practice is changing and procedures for studying younger children in family-genetic studies are now available.

The inclusion of second-degree relatives can become prohibitively expensive and probands may be more reluctant to involve distant relatives. However, certain types of family-genetic studies require information on the "extended pedigree," as noted above.

SYSTEMATIC COLLECTION OF PEDIGREES

The traditional procedure for collecting family data in clinical psychiatry has been to note the family history of psychiatric disorders during the intake examination. Such information concerning "positive family history" is inadequate for genetic-epidemiologic family studies because the population at risk within each family is not quantified in terms of person-years of risk, or even simply in terms of the number of relatives that is being considered. Without basic denominator information, assessment of risks among relatives is clearly problematic and subject to numerous biases resulting from differences in mortality, family size, and the age at which people have their children. Consequently, it is vital to identify all family members, living or deceased, by sex, current age (or age at death), and precise biological relationship.

Systematic methods for eliciting and recording pedigree information from probands have been developed to ensure that the data will be complete and unambiguous and in a form that can be transferred directly to a computer file. Moreover, various logical inconsistencies in pedigree data may be detected through computerized checking operations. An error-checking program is available, as well as a computerized method of reorienting the pedigree information that is obtained from interviews with the relatives of the proband.

The procedure developed by Thompson et al requires that a trained interviewer elicit and record the pedigree data, which takes about 30 to 50 minutes. Cole et al have also developed a questionnaire that subjects are required to complete on their own. However, the accuracy and completeness of the pedigree data obtained through a self-administered questionnaire have not been validated against such data obtained in a direct interview.

Recording data as a pedigree diagram while the information is being collected is a common procedure that has many shortcomings. When diagrams are used it is difficult to collect information systematically, and parts of a family may be omitted. Furthermore, a simple data-recording form is still needed for information, such as
The large family studies in the United States have used the SADS-R.

The DSM-III Criteria.—The DSM-III\textsuperscript{39} is a direct descendant of these earlier diagnostic approaches. The DSM-III employs a multiaxial diagnostic system and permits the concurrent diagnosis of multiple disorders on Axis I and II. While diagnostic hierarchies are reflected in the exclusion criteria for specific disorders (eg, a phobia is not diagnosed if it occurs in conjunction with a major depression), these hierarchical decisions have been suspended in some family studies to test the validity of the proposed hierarchies, and many will probably be eliminated in the forthcoming revision of the DSM-III (DSM-III-R). Most family studies have only assessed relatives on Axis I and not on a range of Axis II diagnoses.\textsuperscript{36}

The Diagnostic Interview Schedule (DIS).—A refinement of both the Renard Diagnostic Interview and the SADS, the DIS was developed by Robins et al\textsuperscript{108} for use by lay interviewers in epidemiological surveys. The DIS is highly structured to reduce the need for interviewer judgments and takes one to 1½ hours to complete. Diagnoses according to the Feighner criteria and the RDC as well as selected DSM-III Axis I diagnoses can be differentially generated by computer from information collected on the DIS. The DIS was used in the National Institute of Mental Health Epidemiologic Catchment Area Study.\textsuperscript{22,24} It can generate lifetime diagnoses and is applicable to family studies, but because of its recency it has not been widely used for that purpose.

The Structured Clinical Interview for DSM-III (SCID).—The newest of the structured instruments, the SCID\textsuperscript{22,24} is designed to enable clinically trained interviewers to make the major DSM-III, Axis I diagnoses. The SCID was modeled on an interview of the present illness (or personal problems) and past episodes of psychopathology precedes the systematic inquiry about specific symptoms. It contains some open-ended questions to encourage subjects to describe their symptoms. Furthermore, the sequence of questions approximates the differential diagnostic process according to the DSM-III decision trees. Its authors are extending the SCID to include all of theAxis I and II adult disorders in the DSM-III, as well as sections for completing Axis IV and V ratings. A modified version of the SCID is being developed for interviews with people who are not patients and, thus, will be suitable for family studies.\textsuperscript{106}

DSM-III Axis II Diagnosis.—Thus far, only a limited number of Axis II diagnoses have been included in structured interviews: Antisocial Personality in the SADS-E and the DIS, and Borderline Personality in some versions of the SADS-E. Because of the considerable research interest in these diagnoses, particularly regarding questions of the spectrum of disorders, several promising diagnostic instruments assessing Axis II diagnoses are currently being tested and will be useful for future family studies. (See Reich\textsuperscript{85} or Stangl et al\textsuperscript{109} for a review.) An instrument developed by Livesley et al\textsuperscript{109} for collecting information to make Axis II diagnoses is now being field tested in a study sponsored by the World Health Organization (WHO).

International Classification of Diseases, ed 9 (ICD-9)/Present State Examination (PSE).—Similar developments on standardized diagnostic interviews have taken place in the United Kingdom. These interviews are linked to the ICD-9. The most widely used is the PSE, developed by Wing et al.\textsuperscript{102,103} which incorporates standard methods of defining, eliciting, and recording information on 140 symptoms in an interview format. The PSE has good reliability, and modified versions have been widely used in large international studies. Only recently, however, have attempts been made to modify the PSE to include information on “lifetime” symptoms.\textsuperscript{104}

Composite International Diagnostic Interview (CIDI).—Currently being tested, the CIDI is a structured diagnostic interview that combines elements of the DIS and the PSE and is capable of generating DSM-III and ICD diagnoses.\textsuperscript{110} The CIDI will be compatible with the European, US, and other diagnostic schemes and will take into account international differences in psychiatric diagnosis, which will facilitate cross-national studies. The tested version is not yet available for general use.

Assessment of Children and Adolescents

Although there has been a debate about using unmodified adult criteria to diagnose affective and other disorders in children,
particularly those who are prepubertal, a number of studies have established the feasibility and clinical usefulness of doing so. Most investigators agree that children over 8 years of age can report accurately about their own symptoms. There are still questions regarding when and about what parents should be queried. As a result, the new structured diagnostic interviews that have been developed for use in conjunction with the DSM-III, are intended to be designed for administration by parents about their children, as well as from parents about their children. These child interviews are very similar in structure to the adult interviews that preceded them. (For an excellent review of psychiatric assessment instruments for children, see Psychopharmacology Bulletin or Orvaschel et al.)

The Diagnostic Interview for Children and Adolescents (DISCA). One of the first diagnostic interviews for young people, the DICA was devised and tested by Herjanic and Campbell. Considerable disagreement was found between parent and child reports of symptoms that were not "concrete, observable, severe and unambiguous." The authors concluded from their findings that parents may be the best informants for behavior problems, while children may be the best informants for subjective symptoms that reflect inner feelings, and that both the parent and child should be interviewed. Similar conclusions were reached by Rutter in his classic epidemiologic studies of children.

The DICA has been revised to generate criteria for most of the DSM-III diagnoses of disorders applicable to children and adolescents. The highly structured DICA can be administered by minimally trained interviewers and usually takes no more than one hour to complete. There are parent and child forms of the interview, which cover the child's lifetime. The information collected on the forms can be summarized individually or in combination on a DSM-III diagnostic form. The DICA has been used in high-risk studies.

The Schedule for Affective Disorders and Schizophrenia for School-Age Children, Epidemiologic Version (K-SADS-E).—This K-SADS was designed as a semistructured interview that was developed by Puig-Antich and associates for administration to 6- to 18-year-old children. Structured similarly to the SADS-L, from which it was adapted, the K-SADS-E is intended for use by highly proficient clinical raters, who have received ten to 15 hours of training in its administration. The interview requires about 30 minutes with subjects who have no clinical disorders, but can require as long as 1½ hours with older subjects who have multiple disorders. It is designed to generate both current and lifetime RDC and DSM-III diagnoses. The authors recommend that the K-SADS-E be administered first to a parent about the child and then to the child about himself or herself, although the same interview format is used for both. The K-SADS-E has been used with prepubertal children and with inpatient adolescents. At the same time, the K-SADS-E has been demonstrated to produce diagnostic information similar with more than 200 children of depressed and normal probands in a high-risk study.

The Diagnostic Interview Schedule for Children (DISC).—The Center for Epidemiologic Studies, National Institute of Mental Health, Rockville, Md, is developing a national program of epidemiologic research on mental health and behavior problems in children and youth. The first stage in the program involved the construction of a new instrument, the DISC, by a panel of child psychopathology researchers with experience in the design and testing of childhood and adolescent diagnostic assessments. Like the DIS for adults, the DISC was designed to be administered by lay interviewers and, like the DICA and K-SADS-E, to be administered to parents about their children and to children about themselves. The interview, which focuses on problems in "the past year," requires less than one hour with a child and from one to 1½ hours with a parent. If used in a family study, it would need to be modified to include lifetime diagnoses. The DISC has been tested for feasibility and reliability with a psychiatric clinic sample and a primary-care clinic sample by Costello and Edelbrock in Pittsbrugh. The results of that study indicated that DISC is acceptable to children, parents, and interviewers, that lay interviewers administer it as reliably as clinically trained interviewers, that interrater reliability of ratings from videotaped interviews was very high, that parents' reports were reliable over a few days to a few weeks, but that children reported many fewer symptoms on a second interview than they did on a first. Those involved in the development and testing of the DISC caution researchers about using the interview before their methodological studies are completed on nonpatient populations.

All of the diagnostic interviews that are available for use with children and adolescents have undergone only limited testing. Results of the assessment of children are rapidly accumulating. The evidence thus far suggests that there is considerable discrepancy between parent and child reports of the child's psychopathology if both parents and children are interviewed independently and that children usually report more disorders than do their parents about them. These findings are similar with the new diagnostic assessments, the DICA, the DISC, the K-SADS-E, and symptom scales. Until the discrepancies can be resolved, diagnostic information on the child should be obtained independently from the child and the parents. Reliance should not be placed exclusively on parents' reports.

Which Diagnostic Assessment Should Be Used? There is often a question about which diagnostic assessment procedure to use. For family studies, a structured or semistructured interview that assesses the lifetime history of diagnosis is necessary. Those reviewed above are all useful and share a common heritage. To permit comparability with other studies, widely used methods should be chosen. The instruments that are selected can be adapted for a study, but marked changes in an instrument decrease the potential for comparing and replicating results across studies. Regardless of which instrument is selected, the interviewers should undergo a period of training, should be tested for interrater reliability, and should be periodically monitored for reliability to avoid rater drift. The selection of the interview should be guided by the level of clinical experience of the potential interviewers. The general rule is that the less clinical experience, the more structured the interview. A rule for gauging training time is that the more structured the interview, the less time required for training.

FAMILY HISTORY METHOD

Despite heroic efforts to obtain direct interviews with all family members, in a family study this is rarely possible due to proband or relative refusal, unavailability, geographic distance, or death. Furthermore, although direct interviews have been found to be the best source of diagnostic information on individuals within a family, family members who are in close contact with their relatives are often able to supply additional information about their relatives' symptoms, or about problems that were denied or minimized (e.g., alcoholism) in the direct interview with the relative.

General Procedure for Collecting Family History Data

Structured family history interview methods for systematically obtaining the data necessary to make RDC diagnoses in family members were developed by Andreasen et al and these methods have been extended and further developed to include the DSM-III diagnosis in a new format—the Family Informant Schedule and Criteria (FISC). After basic information on first- and second-degree relatives has been listed on a pedigree collection form, the proband and each of his first-degree relatives can be asked about the psychiatric status of family members systematically. The usual procedure is to first complete a direct diagnostic interview with the proband or relative about himself, and then to ask the informant whether his first-degree relatives (one by one and by name) have had any of the problems that were just mentioned in the interview about himself. This procedure ensures that the informant is better acquainted with the moods, behaviors, and symptoms that he is asked to identify in his family members. When the informant indicates that a relative has had a problem, the interviewer is instructed to probe for the information necessary to generate diagnoses by using the diagnostic interview as a guide and by completing symptom checklists for relevant disorders. Other questions about
each relative’s treatment history, marriage, social, and job problems may be asked and a narrative summary is completed for each relative in which age of onset and duration of symptoms, diagnosis, treatment, and problems are detailed. There are now companion diagnostic symptom checklists available for schizophrenia, depression, anxiety disorders, personality disorders, alcoholism, and so on, that can increase the precision of family-history data.

**Reliability of the Family History Method in Adults**

Comparison of diagnoses based on family history with diagnoses based on direct interview indicated that the specificity for the family-history method is high, but that the sensitivity is generally low. Accuracy is better for affective disorders and alcoholism than for schizophrenia. Spouses are usually the most knowledgeable sources. 

**Diagnostic information from relatives is scanty and the treated family member cannot be interviewed directly.** Chart-review coding forms that consist of checklists of diagnostic criteria have been used to obtain symptom information systematically from treatment records. It is best to code the symptoms noted in the chart and to derive diagnoses based on symptoms rather than to focus exclusively on the diagnosis given in the chart. In general, charts rarely contain sufficient information necessary to arrive at a diagnosis in the absence of either direct or family history interview data.

**ADDITIONAL ASSESSMENTS OF ADULTS AND CHILDREN**

Basic sociodemographic data (age, sex, marital status, occupation, ethnicity, etc.) and medical and treatment history data (including medications taken for both medical and emotional problems currently and in the past) are generally collected before beginning the direct diagnostic interview. In addition, a number of other psychosocial assessments may be obtained in a self-administered format. For example, measurement of cognitive functioning, recent and past stressful life events, social adjustment, marital reports, or family members may be useful in identifying environmental factors that may be implicated in the development of the disorder. There are a number of psychosocial measures that may be used to describe the family environment or to determine risk or modifying factors or consequences of diseases in families. A comprehensive interview that includes assessment of the family environment as well as lifetime psychiatric status, family and helplessness has been developed by the Washington University (St Louis) group. Since psychosocial measures rate current status rather than trait, they may not be useful as correlates of lifetime psychiatric diagnoses, and attention should be given to the time frame assessed.

For family members under 18 years of age, questions on sociodemographic data and on medical and treatment history should be included. In addition, parental attitudes and behavior toward the child and family interactional patterns can be assessed through self-administered reports completed by parents and children or through brief or elaborate family-interactional procedures as described earlier. The assessment of psychological history, early neurotic signs, and medical history of children is considered to be particularly important in studies concerned with childhood psychopathology.

**MULTIPLE SOURCES OF INFORMATION ON DIAGNOSES: BEST-ESTIMATE PROCEDURES**

Since diagnostic data on each family member often may come from multiple sources (different relatives, records, direct interview), resolution of the discrepancies that inevitably will occur is critical. The best-estimate procedure has been developed for this purpose. This method is based on all available information and is made by at least one clinician, usually a psychiatrist who is blind to the diagnostic status of the proband and who is not involved in direct interviews of any of the probands or relatives. An independent assessment is obtained by a second diagnostician. Any substantial disagreements between the two diagnosticians may be reviewed on a case-by-case basis until agreement is reached, but blindness with regard to the clinical status of the proband should always be maintained. If the best-estimate diagnoses show that the probands had other diagnoses that could have excluded them from their original proband group, those probands and their families
should be excluded from the study. The best-estimate procedures, which the Yale group used as part of a family case-control study of major depression among more than 1500 relatives, are described elsewhere.  

**ISSUES IN THE EPIDEMIOLOGIC ANALYSIS OF DATA**

The focus of the epidemiologic analysis of data from family studies is the investigation of the effect of various proband factors on the risk of the disorder developing in relatives. There are several important statistical issues that affect these analyses: (1) the concept of lifetime risk of disorder in relatives and the associated problems of variable follow-up or observational times; (2) controlling for confounding effects of extraneous factors while simultaneously investigating risk factors of interest; and (3) adjusting for lack of independence of observations among members of each family. Recent advances in methodology for dealing with these issues will be discussed.

**Methods of Estimating Morbid Risk**

The major outcome in a family study is rate of illness in relatives over their lifetime, termed morbid risk or lifetime prevalence. The term lifetime risk as used here refers to the risk of onset of a particular disorder between birth and some particular age $t$. Since it is impractical to follow-up individuals over a lifetime until their death and determine their risk of onset of illness, lifetime risk in a family study is obtained retrospectively. For purposes of epidemiologic analysis, the design of most family studies can be viewed as a retrospective cohort study. A cohort study involves a design in which information about the study factor is known at the beginning of the follow-up period. The population at risk of developing the disease is followed up for a given period, after which new cases are identified. In nearly all family studies this information is obtained retrospectively from the relatives. The limited exception would be longitudinal high-risk studies in which children are followed up prospectively for some limited period of time. In a family study, the study factor of interest is the ill proband, and the population at risk consists of the relatives of the proband. The follow-up time is equal to the age at onset of disease for those individuals who developed the disease, age at death for those who died before developing the disease, and current age for those who were alive and did not develop the disease.

There are various methods of estimating morbid risk to adjust for variable age at onset in relatives, such as the methods using an a priori age-at-onset distribution developed by Weinberg and by Stromgren, which are well known and therefore will not be discussed. More recently, Thompson and Weissman have shown that the product-limit method may be used to quantify lifetime risk, while adjusting for a variable age of onset. This method falls into the general category of survival-time analysis. Survival-time analysis can be used to estimate the morbid risk in relatives by recognizing that age at onset of disease can be treated as a survival time with censoring occurring when no disease appears by the end of the follow-up period. The survival function at time $t$, also known as the lifetime risk at age $t$, may be estimated by actuarial life-table methods or by the product-limit method. If $t$ was desired to complete the morbid lifetime risk in relatives between different levels of a proband factor, one would complete the survivor function for each group of relatives with a specific level of the proband factor by one of the above methods. These survivor functions can then be compared by using the log-rank test. A parametric approach to estimating the age-at-onset distribution in the "susceptible" population has been developed for genetic studies.

**Simultaneous Estimation of the Effect of Multiple Risk Factors**

When it is known or suspected that factors other than proband illness, such as sex and age of relative, may account for the differing rates of disorder in relatives, one must control for these factors while simultaneously estimating the effect of the proband factors. This may be accomplished by using multivariate regression models for survival data, such as Cox's proportional-hazards model or Holm's linear model for survival analysis.

**Investigating Effects of Proband Factors on Age at Onset of Disease in Relatives**

The regression models described above assume that the age-specific risk for any two levels of a proband factor remains constant. These models are not appropriate for investigating the hypothesis that relatives of probands with a given level of a proband factor are more likely to develop the disorder at an earlier age, when compared with relatives of probands with a different level of this factor. However, suitable extensions of these models have been developed to investigate this hypothesis. Cox's proportional-hazards model has been extended by Kalbfleisch and Prentice to investigate, through graphic methods, relative risks that vary with age. Cox's model, with time-dependent covariates, and Holford's model, which includes interaction terms between time and the covariate of interest, may also be used to study this hypothesis. These latter models are extensions of the regression models described in the preceding sections.

**Lack of Independence of Observations in Family Data**

The methods of survival analysis described in the preceding sections assume that the observations are independent of one another. This assumption may not be justified in family data. The problem of nonindependence of the observations in family data, when the outcome variable is continuous, has recently received attention in the epidemiologic literature. Methods for dealing with this issue, when the outcome is categorical and a variable number of subunits are available for a primary unit of analysis, as in familial data, have been developed. However, these methods assume that the risk remains constant over the study period and do not adjust for a varying age at onset. Methods for handling this problem, and the follow-up time variability, currently do not exist, and their development would be a significant contribution to the epidemiologic analysis of data from family studies.

**Adjusting for Cohort Effects**

There is increasing evidence that the rates of some psychiatric disorders, such as major depression, mania, and alcoholism, vary with birth cohort, ie, lifetime rates of disorder in younger cohorts are higher than those in older cohorts. Methods of age correction that use an a priori age-at-onset distribution, such as those developed by Weinberg and Stromgren, may not be appropriate in those disorders in which such a cohort effect exists. Traditionally, these methods have been used to adjust for the fact that individuals belonging to younger cohorts have not passed through the entire period of risk. Adjustment is made by using suitable weighing factors based on an a priori age-at-onset distribution to increase the rates in these younger cohorts. In disorders in which the unadjusted rates in younger cohorts are already higher than those in older cohorts, adjusting to increase these already high rates is inappropriate.

When using survival-analysis techniques to estimate morbid risk, combining information from multiple cohorts...
requires the assumption that there is no substantial change over time in the age-specific risks for developing the disorder.\textsuperscript{189} Otherwise, presenting morbidity risks averaged over multiple birth cohorts would obscure important features of the data. Furthermore, estimates of lifetime risk would vary according to the relative distribution of various birth cohorts in the sample. When it is known that the rates of the disorder under investigation vary with the birth cohort, it is important to control for this variable by including the year of birth as an independent variable in the multiple-regression model.

**GENETIC ANALYSIS OF FAMILY STUDY DATA**

If a trait is partly determined by genetic factors, it is expected to show familial aggregation. Therefore, one of the initial steps in any genetic analysis of family study data is an attempt to determine whether or not the phenotype (the observed trait or disorder) shows familial aggregation. Familial aggregation can be defined in many ways statistically and the method chosen will depend in part on the nature of the phenotype and on the type of family data available. For a quantitative trait, familial aggregation can be defined as a greater correlation for the trait between pairs of relatives than between pairs of unrelated individuals. Alternatively, one can measure the variability of the trait within and between families; the greater the between-family variance relative to the within-family variance, the higher the familial aggregation of the trait.\textsuperscript{64} For a qualitative trait, such as the simple presence or absence of a disorder, familial aggregation is shown as a greater probability of being affected for a relative of an affected proband than for a random individual in the population. In a case-control family study, the relatives of unaffected probands would be examined instead of random unrelated individuals in the population; this gives greater statistical power for common disorders.

For complex disorders, such as most of the major psychiatric disorders, the choice of phenotype to examine is not necessarily obvious. As more is learned about these disorders, it is possible to define relevant phenotypes at several different levels: the clinical, neurophysiological, biochemical, and, eventually, the molecular.\textsuperscript{186} Cloninger et al\textsuperscript{186} have described a multistage approach to analysis of data when it is possible to measure more than one of these multiple dimensions. Their approach involves (1) quantifying the amount of observed variation and clarifying the relationships among the variables to derive the inferred or latent factors underlying the variables; (2) evaluating these latent factors in a general population sample; (3) evaluating the inheritance of the inferred components of the “ultimate phenotype”; (4) evaluating the familial relationships among the individual components; and (5) predicting the risk of developing the ultimate phenotype by examining information on all of the component phenotypes in each individual and his relatives.

There are several genetic models for complex disorders that serve as hypotheses in such analyses. These models relate the allelic forms of genes to genotypes, the population prevalences of the genotypes and other relevant variables to the phenotypes, and the distributions of phenotypes within families and within the general population. These models and their assumptions and analytic methodologies have been reviewed by Kidd.\textsuperscript{177,178} Segregation analysis and its extension beyond the nuclear family, usually referred to as pedigree analysis, are general approaches to analyzing genetic models. The relative strengths and weaknesses of the two approaches have been discussed by Kidd and Matthysse\textsuperscript{61} and by Pauls and Kidd.\textsuperscript{179} The programs that are available encompass a wide variety of genetic models including (1) one locus with two alleles, (2) one locus with three alleles, (3) two loci with two alleles each (with linkage analysis as a specific case), (4) polygenic, and (5) a mixed model with a major locus plus a polygenic background.\textsuperscript{180-182}

Several computer programs are available for testing sophisticated genetic models; among the more commonly used and more readily available are PA\textsuperscript{183} and POINTER.\textsuperscript{182} Though the two programs represent complex pedigrees in different ways and consequently use different algorithms that have different capabilities, both can handle a large class of genetic models.

Tests of genetic hypotheses are possible through likelihood-ratio tests comparing an unrestricted mixed model of major locus and polygenic inheritance to nested models. Tests of mendelian transmission at the major locus are also possible.\textsuperscript{184}

There are limitations to the use of genetic models in studies of familial transmission of complex human disorders.\textsuperscript{185} On the one hand, the methods for determining the mode of inheritance from family data using statistical approaches have become more powerful over the years. Better statistical procedures and more elaborate computer programs allow for more sophisticated segregation analysis with better power for determining the mode of inheritance. The power of these methods has been shown in several simulation studies.\textsuperscript{185,186} However, the sensitivity of these statistical approaches to violation of the underlying assumption and the potential for spurious results are considerable.\textsuperscript{187-189} While statistical analyses seem more likely to yield results that are apparently definitive, such approaches have rarely been able to demonstrate conclusively a specific mode of inheritance for any complex disorder.

**Genetic Linkage Studies**

The strategy for determining modes of inheritance has been to examine the linkage of illness to a classic chromosomal marker using extended pedigrees. Genetic linkage studies are useful for delimiting genetic heterogeneity; identifying the mode(s) of inheritance; determining the specific biochemical abnormality and the pathogenesis of the illness; identifying regulatory variants that cannot be studied directly; determining which nongenetic, e.g., environmental, factors are relevant to onset and/or course of illness; and determining gene-environment interactions.\textsuperscript{190,191} All of those are predicated on having first found a linked marker. Past linkage studies produced limited results since only about 22% of the genome could be reasonably scanned.\textsuperscript{192,193} However, the search for genetic markers linked to psychiatric illness now has more promise because of recent advances in genetics.

Recombinant DNA methodologies have revolutionized the field of human genetics.\textsuperscript{190,194,195} Research questions that were heretofore unanswerable can now be simply addressed with these new methods. Several new research approaches are relevant to the study of psychiatric disorders. Normal genetic variation identified directly in human DNA is allowing construction of a human genetic map.\textsuperscript{196} As of mid-1985 there were over 800 of these DNA markers, termed restriction-fragment length polymorphisms. These allow construction of a detailed genetic map of considerably more than 50% of the human genome.\textsuperscript{196,197} and the total map of the human genome should be saturated with such markers by the end of the current decade.

The availability of hundreds of genetic markers mapped throughout the human genome makes genetic linkage a powerful approach for studying major psychiatric disorders. Genetic linkage studies use the organization of the genetic information to follow the inheritance through a
single family of whole blocks of genes that occur near each other. One or two genes in a block can serve as genetic markers if there are variants to follow; the restriction-fragment length polymorphisms will serve this function because they have common alleles that are distinguishable. A linkage study compares the pattern of inheritance of blocks of genes among relatives with the inheritance pattern of the trait of interest. A significant intrafamilial correlation between an individual and a DNA marker locus indicates that some genetic information in the region around that marker is involved in determining who is and who is not affected. This approach has recently been successful in identifying the location in the human genome of the gene for Huntington's disease, as well as for other nonpsychiatric disorders.

The elements of a strategy for finding a linked marker include (1) starting with a (diagnostic) subtype that seems to be highly familial as a specific subtype; (2) finding at least one very large multigenerational family with multiple individuals affected with this subtype; and (3) conducting the linkage study using multilocus mapping methods that incorporate the known relationships among the markers, thereby allowing recombination to be distinguished from incomplete penetrance.

Family characteristics that are desirable for the genetic linkage analyses that examine patterns of transmission of complex psychiatric disorders are (a) a large number of siblings; (2) multigenerational extended pedigrees; (3) a high density of illness in the families; and (4) an absence of matings involving two affected individuals.

Attempting to perform linkage analysis with a trait that is not known to be a single-gene disorder is not only difficult, it is controversial. The argument is sometimes made that one cannot perform a linkage analysis until there is evidence of a major locus. One of us (K.K.K.) has argued that possibly the only way to obtain convincing evidence of a major locus of a complex psychiatric disorder is by linkage.

Other obstacles to linkage studies of psychiatric disorders include a lack of standardization for systematically collecting pedigrees, a lack of reliability and specified diagnostic criteria, and likely genetic heterogeneity of psychiatric disorders.

The methodology developed or refined over the past decade for family studies is complementary to these new linkage techniques and the substantive results of the family studies suggest which diagnostic subgroups may be most promising to pursue. Several DNA laboratories with investigators interested in psychiatric disorders are working on putting together the human genetic map and developing multipoint mapping methods.

A training program with guidelines to identify families that are potentially informative for linkage studies and to standardize the assessment procedures of the families selected has been developed by some of us (M.M.W., K.K.K., and K.J.) at Yale University in 1985.

NEW METHODS AND FUTURE STUDIES

There are new methods under development that will improve the precision of family studies. For example, availability of diagnostic instruments for Axis II diagnoses will make possible more comprehensive study of the familial aggregation of Axis II diagnoses and their relationship to Axis I disorders.

Problems remain in reconciling the results between studies conducted in the United States and those conducted in the United Kingdom and continental Europe in view of the different diagnostic criteria used (DSM-III in the United States and ICD abroad). The development by Robins et al. of a diagnostic instrument, the CIDI, which shares the heritage of the DIS and PSE and will generate both DSM-III and ICD-9 diagnoses, may provide some resolution.

As data continue to accumulate on the reliability of diagnostic information on children under 18 years of age, future family studies may be able to include assessment of a greater number of these younger offspring. A major unresolved problem is that of the considerable discrepancy between parents' and children's reports of the children's clinical status. There are no guidelines as to who is the better informant. Thus, it is unclear which children should be included as affected in epidemiologic or genetic analysis. Furthermore, the issue of the continuity of childhood and adult disorders needs to be resolved. In a longitudinal study of children of psychiatrically ill parents, Rutter et al. have shown that the majority of neurotic children do not develop major psychiatric disorders in adulthood. Therefore, at this time, one cannot predict which of these children will continue to manifest psychiatric disturbance in adulthood.

No family studies, to our knowledge, have included assessment of offspring under 6 years of age. Here, there is an opportunity for fruitful collaboration with developmental psychologists who use observational techniques that can be modified for field studies. As an example, the approaches used by Kagan et al. in their elegant studies of very young shy children may be important in assessing children of adult probands with agoraphobia and panic disorder. Is shyness an early form of agoraphobia?

The development of training material or videotapes for the procedures of pedigree collection and family history data and training programs for diagnostic interviews by L.N. Robins, PhD, on the DIS, and the training program by J. Endicott, PhD, on the SADS, will help to improve reliability and standardization of procedures across studies.

Comparison of findings between studies will also be improved by the explicit descriptions in published reports of the diagnostic hierarchies used in statistical analysis of family data. The lack of replication of findings between studies may be merely a function of different inclusion criteria for diagnosis in the analytic procedures.

The emerging capacity to perform chromosomal linkage studies using the new advances in molecular genetics provides the promise that the etiology of some of the psychiatric disorders may be found. It may be possible to isolate a specific genetic factor that may predispose an individual to the expression of a particular psychiatric disorder. Less important will be the search for those environmental factors that are risk or protection mediators for the expression of the disorders in high-risk families.

The identification of informative families for genetic linkage analysis requires that diagnostic criteria be applied reliably and systematically. The application of highly refined molecular techniques to haphazardly obtained and poorly assessed families will obscure findings in these studies. Moreover, in the selection of families, attention needs to be paid to the emerging data on the heterogeneity of psychiatric disorders.

The familial nature of the major psychiatric disorders—affective and anxiety disorders, schizophrenia, and alcohol abuse—are reasonably well established. However, the precise risk, diagnostic specificity, and overlap are less clear, in part because of the imprecision in the diagnostic criteria used, the lack of control groups, and the use of diagnostic hierarchies. There are numerous studies that could yield useful information on familial risk and risk factors, clarify the relationship between diagnostic categories, and have implications for treatment and detection.
Potentially fruitful areas that could be addressed by properly designed family studies include, but are by no means limited to, these questions. Is there an overlap between major depression and panic disorder? Is familial aggregation observed in schizophrenias with paranoid symptoms? Is schizophrenia in early childhood related to adult phobia? Is suicide familial? Is delusional depression on the spectrum of bipolar disorder? Is there a relationship between drug abuse and other addictive behaviors and psychiatric disorder? Are there familial vulnerabilities in persons who develop psychiatric disorders following drug ingestion, eg, mania in association with tricyclic antidepressants, major depression in association with β-blockers for hypertension, or psychosis following hallucinogens, or following life stress such as depression that first develops in the context of a divorce? The new nuclear magnetic techniques may be fruitfully applied to ill patients with high familial loading for the disorder, as well as to some of their more mildly ill or well relatives.

Adaptation for Clinical Practice

The methods described herein for collecting information on biological relatives of psychiatric patients can be readily adapted for routine clinical practice. In principle, the methodology for a family-genetic and family-dynamic approach to psychiatric disorders should not differ. In fact, family therapists have for years used genograms to track psychiatrically significant events in families. Family-genetic studies are primarily concerned with family resemblances in a disorder among biologic relatives (living and deceased), regardless of whether there is contact with the proband under study. Family-dynamic or systems studies are concerned with biologic as well as nonbiologic relatives living in the immediate family unit and do not focus on specific diagnoses. From the perspective of an epidemiologist, the family-dynamic or systems view may be seen to be more consistent with an infectious disease model in which contact among family members, regardless of their biologic relationships and exposure to noxious factors (however defined), is the relevant area of study. With some training and ideologic flexibility, the same personnel engaged in routine clinical practice can collect data compatible with the family genetic and family dynamic perspectives. The systematic collection of pedigrees followed by the collection of the family history of psychiatric disorders and, as indicated, direct interview of the patients' first-degree relatives, can be incorporated into clinical evaluations. This information on familial resemblances of psychiatric disorders can be quite useful not only in planning treatment of the individual patient but also in determining which family members may be of assistance and which members may themselves require assistance. The biological family members of psychiatric patients, both adults and children, are at increased risk for psychiatric disorders. The use of family study information to gain an understanding of their clinical state may be a potential area for early case finding and preventive intervention.

This study was supported in part by Alcohol, Drug Abuse, and Mental Health Administration grants MH 82744 and 30197 from the Affective and Anxiety Disorders Research Branch, National Institute of Mental Health, Rockville, Md; Yale-MacArthur Foundation Research Center, National Institute of Mental Health grant MH 30929; the John D. and Catherine T. MacArthur Foundation Mental Health Research Network on Risk and Protective Factors in the Major Mental Disorders (Dr Weissman); and Research Scientist Development Award MH 00490 (Dr Merikangas) from the National Institute of Mental Health.

References

null


