Assessing Children in Psychiatric Epidemiologic Studies

HELEN ORVASCHEL, PH.D.
DIANE SHOLOMSKAS, M.A.
MYRNA WEISSMAN, PH.D.

Introduction

Reliable epidemiologic data on rates of psychiatric disorders in children are nearly nonexistent. In particular, data are lacking on the incidence and prevalence of specific psychiatric disorders, rather than global impairment or symptoms independent of diagnosis that were characteristic of past research. Information on the rates of specific disorders would aid investigators in the identification of risk factors, advance our knowledge of etiology, contribute to the determination of treatment needs, and provide the groundwork for preventive strategies.

Rates of incidence and prevalence of psychiatric disorders cannot be based on treated populations alone. This source of data is subject to error, since many children with psychiatric disorders are likely to be untreated and therefore remain uncounted. In addition, considerable selective bias operates in determining which children are seen in treatment. For example: (1) Boys are referred for treatment more often than girls; (2) Acting-out behavior is referred more often than less overtly troublesome or withdrawn behavior; (3) Milder disorders are less likely to be treated than more severe disorders; (4) Socio-demographic variables and the availability of facilities are likely to influence who seeks and/or receives treatment. Generalizations based on clinical data are therefore, subject to distortions as to the sex and social class of the patient and the severity and type of disorder treated.
Methodological Issues

A major obstacle to epidemiologic studies of psychiatric disorders in children in the community has been the controversy regarding the reliable definition of a case. This problem is similar to previous difficulties encountered in epidemiologic studies of adult psychiatric disorders. However, there have been major achievements in the field of adult psychopathology which recently have been applied to community epidemiologic studies (Weisman and Kleinman, 1978). These achievements include improved concepts, definitions, and criteria for adult psychiatric disorders such as the Research Diagnostic Criteria (RDC), American Psychiatric Association Diagnostic and Statistical Manual-Third Edition (DSM-III), and the Feighner Criteria (Feighner et al., 1972). Such advances in definition have accelerated research in adult psychiatric epidemiology by reducing the problem of criteria variance in diagnosis (Spitzer et al., 1978). This work in the classification of disorders has now been expanded to include psychiatric disorders of children (i.e., C.A.P. DSM-III), and will eventually have a similar impact on research in child psychiatric epidemiology.

Advances in adult psychiatry have also been made in techniques for symptomatic, behavioral, and diagnostic assessments. These structured assessment tools include the Schedule for Affective Disorders and Schizophrenia (SADS), the Present State Exam (PSE), and the Diagnostic Interview Schedule (DIS) (Spitzer and Endicott, 1978; Wing et al., 1974; Robins, 1979). Such techniques have reduced information variance in assessment and have resulted in considerable improvements in the reliability of psychiatric diagnosis (Endicott and Spitzer, 1978). In an effort to extend these procedures to psychiatric disorders of children, several useful scales have been developed to assess psychopathology in children.

Assessment Techniques

This chapter will review five instruments which use direct interview techniques and seem promising for epidemiologic studies of children. Although these scales are at varying stages of instrument development and most have not yet been tested in the community, they all reflect advances in the construction of standardized assessment tools suitable for survey research.

The many self-report scales and symptom checklist inventories available that are generally administered to parents, teachers, or
other adults about the child have been excluded from this review for several reasons. Self-reports are more vulnerable to response biases by subjects and less amenable to qualitative information gathering than interview techniques. They do not allow for standardized observational evaluations by trained professionals, so that assessments of observable behaviors such as motor activity, immaturity, and abnormal speech patterns become problematic. In addition, certain areas of functioning are more difficult to assess by self-report measures. This is particularly true for psychiatric symptomatology which may appear in the context of mania, depression, schizophrenia, or some other psychiatric disorder. The limitations of self-reports may be especially severe if a measure is not designed to be completed by the child, who may be an important source of information concerning subjective experiences such as anxiety, depression, guilt, and the like (Herjanic et al., 1975). Self-reports do offer a number of advantages (i.e., cost, time) and it is not our intention to argue against their use. However, interview techniques overcome many of the limitations of self-reports, while providing the flexibility of information gathering often needed when dealing with less-established areas of investigation such as child psychiatric epidemiology.

In selecting which interview schedules to present in this review, our guiding criterion was their suitability for or adaptability to epidemiologic research. The following criteria were also included in our selection decision: (1) coverage of a broad spectrum of diagnostic categories or behavioral dimensions, rather than concentration on a circumscribed number of symptoms or disorders (interviews concerned primarily or exclusively with depression, anxiety, or hyperactivity were excluded, as were other single focus interviews); (2) inclusion of milder disorders as well as more severe forms of pathology; (3) use of items that are relatively defined and allow for quantitative analysis; (4) utilization of the child as one of the sources of information; and (5) existence of some reported reliability data; most scales are in the process of additional testing and refinement.

Three of the interviews (the Diagnostic Interview for Children and Adolescents, the Kiddie-SADS, and the Isle of Wight Inventory) are structured according to a categorical approach. These interviews are designed to ascertain the presence of symptoms associated with a specific psychiatric disorder. The remaining two instruments (the Mental Health Assessment Form and the Screening Inventory) utilize primarily a dimensional approach to assess overall symptomatology without grouping symptoms according to typology. The dimensional approach may be advantageous in that it
allows the investigator to use linear multiple regression analysis for prediction. Multiple regression is a statistically powerful technique which does not employ the decision trees or a priori assumptions regarding typology that are applied to categorical instruments.

On the other hand, the use of multivariate statistical techniques (e.g., discriminant function analysis) will also provide predictive information and may be used with categorical data. The multivariate approach does not require that the data meet many of the assumptions necessary for linear regression (i.e., normality, linearity and additivity) and is superior in its handling of interactive effects between predictor variables (Paykel et al., 1973). A categorical orientation has the additional advantage of being conceptually similar to the framework used in psychiatry for research on etiology and treatment. Given that both approaches to assessment (categorical and dimensional) have advantages and limitations, the investigator must decide whether one or the other or some combination of the two meets the needs of the particular epidemiologic study in question.

The five interview schedules selected will be reviewed as to their primary purpose, scale properties, structure and content, prior use, and psychometric properties. A brief discussion of their limitations and present status is also provided.

**Diagnostic Interview for Children and Adolescents**

The Diagnostic Interview for Children and Adolescents (DICA) was developed by Herjanic and Welner. It is a structured psychiatric interview, with a categorical orientation, and is designed as a diagnostic instrument for children between the ages of 9 and 17 years. In a more preliminary version it was known as the Children's Psychiatric Interview (CPI) and was used with 6- to 16-year-old children. The current DICA is a considerable improvement over the CPI, but is still under revision and must be considered preliminary.

The DICA is administered directly to the child by an interviewer. A parallel form is available that is administered to the parent about the child. It requires approximately 1 to 1½ hours to complete and assesses psychiatric symptomatology during the child's lifetime. The interview is preceded and contains a complex scoring system. The DICA begins with basic demographic questions and inquiries about school functioning, relationships at home, and general interpersonal functioning. It covers the diagnostic categories of Conduct Disorder, Drug and Alcohol Abuse, Depression, Mania, Phobias, School Phobia, Obsessions and Compulsions, Anxiety Neurosis, Briquet's,
Depersonalization/Derealization, and Psychosis. Questions on enuresis and enoprosis are included, in addition to sections dealing with sexuality, insight, judgment, orientation, and memory.

Analysis of this interview has thus far been reported only for the preliminary CPI version. The CPI was administered to 30 boys and 20 girls (aged 6 to 16 years), selected at random from a children's mental health clinic (Herjanic et al., 1975). Interrater reliability scores ranged from .80 to .95 with a mean of .89. Co-ratings showed an average agreement of .84. Comparisons of parent and child responses (for parallel instruments) showed an average agreement between items of 80%. No significant differences were found for content by age group, although, by sex, girls showed a consistently higher level of agreement with their parents. Testing to establish discriminant validity has been promising and attempts at obtaining diagnostic validity are in process (Herjanic and Campbell, 1976).

The DICA could be considered a very promising diagnostic instrument with considerable epidemiologic potential. Questions are well phrased and diagnostic categories appear theoretically sound. The authors are aware that further testing and refinement are necessary and are in the process of collecting the necessary data (Herjanic, 1979).

Mental Health Assessment Form

The Mental Health Assessment Form (MHAF) was developed by Kestenbaum and Bird (1979) as a screening instrument for the clinical assessment of school aged children (7 to 12 years old). It has a dimensional orientation to symptomatology and was originally designed as a tool for collecting standardized data on a sample of children with schizophrenic parents, in order to identify vulnerability to pathology.

The MHAF is not actually a psychiatric interview but, rather, a rating form that is used as an outline, around which a semistructured interview is conducted. The length of the interview may vary considerably although it is suggested that approximately 45 minutes are adequate. There are about 180 defined items to be rated on a variable scale with most items scored from 1 to 5 (ranging from no deviation to marked deviation). Specific instruction regarding the time period assessed is left unclear, but the items appear to refer to "current" functioning. The form is precoded and contains a separate score sheet.

The MHAF contains two major sections which are further subdivided into a number of areas. The first section is rated on the basis
of the interviewer's observations of the child and includes items in
the areas of physical characteristics, motoric behavior and speech,
interpersonal relatedness, affect, and language and thinking. The
second section of the form is rated on the basis of material the
interviewer has elicited from the child during their interactions, and
includes categories dealing with feeling states, interpersonal rela-
tions, dreams and fantasies, self-concept, moral judgment, and
general level of adaptation.

A trained mental health professional, skilled in child interviewing
techniques, is assumed to be required to do the interview and
ratings. The ratings are based on "accepted clinical impressions
about normality." A half-hour videotape demonstrating the use of
the MHAF is available. A form for use with adolescents (ages 13 to
19) is also being developed by the authors (Bird, 1979). It contains
additional items in such areas as drug abuse, alcoholism, and sexual
behavior. This form is still in the preliminary phase of development.

Results of the MHAF have been reported for 35 children in the
following three groups: 7 children with a schizophrenic parent; 2
children with a manic-depressive parent; and 26 children with a
"normal" parent (Kestenbaum and Bird, 1978). Interviews were
videotaped and independently rated by three child psychiatrists.
Reliabilities among raters ranged from .43 to .94 for items on which
some variance was present. High reliability was reported for items in
the Feeling States, Interpersonal Relationships, Self-Concept, and
Moral Judgment sections. Poor reliability was found for General
Level of Adaptation and Use of Defense Mechanisms.

The items of depression, anger, anxiety, disturbed relationships
with mother, aggressive behavior, covert aggression, bizarreness of
dreams, and discrepancy between perceived and ideal self were
found to discriminate between the children with psychiatrically
disturbed parents and the children with normal parents. A validity
study comparing the interview results of 30 clinic children to cli-
cian ratings has also been completed. Preliminary findings suggest
some validity on 35 items, particularly anxiety and depression.
Many of the items could not be assessed because of a lack of
variance in the ratings (Kestenbaum and Bird, 1979).

The MHAF requires a good deal more testing. Many of the items
are unclear in meaning and have questionable placement in the text
of the interview (I.e., masturbatory activity is rated as a motor
behavior along with items such as tics and tremors). The ratings of
many items require highly subjective interpretations of behavior by
the interviewer (I.e., egocentricity, character of fantasy). These
ratings could be improved if confined to the evaluation of specified
and observable behaviors. Reliability scores reported thus far on the MHAf have been poor on many items. In addition, a better test of reliability for the interview as a whole would be generated by a test-retest situation, using different interviewers. Discriminant validity has yet to be demonstrated, since initial data were reported for a very small sample of children. The authors have indicated that the MHAf is still in revision and that they are attempting to improve the specificity of their scoring system.

**Screening Inventory**

The Screening Inventory (SI) was developed by Langner, Gersten, and Eisenberg to assess psychiatric impairment in children from 6 to 18 years of age. It is designed for use in community samples as a general screening instrument rather than as a diagnostic tool. The scale has a dimensional orientation and was intended to be descriptive and to represent a wide range of behaviors with predictive value for psychiatric impairment (Greene et al., 1973).

The SI is composed of three sections: (1) the background section consisting of 11 demographic items; (2) the child section consisting of 40 child behavior items; and (3) the parental section consisting of 32 parental behavior items. There are also two forms of the SI, an M-form designed to be administered to the child’s parent, and a C-form designed to be administered to the child. Direct administration of the C-form, however, is limited to children 14 years of age or older. The inventory items are specific, defined, and have precoded corresponding ratings. While the time period assessed is not clearly stated, the instrument is intended to correlate with “current impairment.” The questionnaire is generally administered as a structured interview (but probably can be adapted for use as a self-report). Completion time is approximately 20 minutes for a trained interviewer. The SI is accompanied by instructions and scoring sheets, age- and sex-specific forms, and a scoring procedure.

The original SI consisted of 654 items measuring the child’s functioning with parents, siblings, peers, and in school. The mothers of 1,034 randomly-selected children and 1,000 welfare-supported children were interviewed for 2 to 3 hours each. Male and female children were included in the sample which represented all age groups between 6 and 18 years. The mother’s responses to the questionnaire were then evaluated by two psychiatrists who rated the child's level of impairment on a 5-point scale. Interrater reliability for total impairment was .84 (Langner et al., 1976). Of the 2,034 children in the total sample, 357 were randomly selected for an
individual psychiatric interview lasting approximately 1½ hours. Psychiatrists rated these children on the same 5-point psychiatric impairment scale. Ratings based on the children's interview correlated from .33 to .48 with ratings based on the mother's interview.

The questionnaire was reduced from its original 634 items to 287 items by eliminating items with low frequencies and condensing or combining other items into composites. A principal component factor analysis with orthogonal varimax rotation was performed on the remaining items. Eventually, 7 factors were identified which correlated .73 with the total impairment scores (Gentry et al., 1976). The five items with the highest loadings on each factor were selected to represent their subscale. The seven factors were: Self-Destructive Tendencies, Mistrust Problems, Conflict with Parents, Regressive Anxiety, Fighting, Delinquency, and Isolation. The multiple correlation of the subscales with overall impairment was .64. An additional five items were added to the child behavior scale bringing it to a total of 40 items, with an internal consistency of .76 and interrater reliability from .43 to .66. Five factors for the parental items are also reported: Mother Rejecting, Parental Coldness, Mother's Physical and Emotional Illness, Parents Quarreling, and Child's Mild Chronic Illness.

An additional form of the SI has recently become available. This form is known as the Family Research Project Questionnaire and is intended as a cross-cultural version of the SI C-form. It is expected to be used in a number of countries in the hopes of developing cross-cultural norms of child psychopathology (Katz, 1979).

The SI has a number of valuable features as a screening inventory. It contains a relatively small number of well-defined items which can be administered with adequate reliability. It also enables investigators to compare data with age- and sex-appropriate norms which have been derived from a large sample. The inventory has some limitations in that it is not intended for direct administration to children under the age of 14 and therefore relies on the parent as its primary information source. In addition, interrater reliability for the 40 child behavior items was acceptable but modest (.43 to .66) and the correlations between psychiatrist's impairment ratings for mother vs. child informants were even lower (.33 to .48).

Kiddie-SADS

The Kiddie-SADS (K-SADS) was developed by Puig-Antich and Chambers as a children's version of the Schedule for Affective Disorders and Schizophrenia developed for adults by Spitzer and
Endicott. It is a structured psychiatric interview with a categorical orientation, designed primarily as a diagnostic instrument for children between the ages of 6 and 17 years. It is administered directly to children, as well as to parents about their children, and necessitates the use of a trained interviewer. It requires approximately 45 minutes to 1½ hours to complete and assesses psychiatric symptomatology during the past month or since the onset of the illness episode in question.

The first few minutes of the K-SADS involve a semistructured interaction between the interviewer and the child, during which the nature of the problem(s) and the duration are ascertained. This discussion is followed by the structured portion of the interview, which systematically investigates the child's overall symptom picture. The K-SADS covers the following diagnostic categories: Major Depression (including subtypes), Manic Disorder, Separation Anxiety (with or without school phobia), Generalized Anxiety, Phobias, Depersonalization/Derealization, Obsessive-Compulsive Disorder, Conduct Disorder, and Schizophrenia. The coding system is defined and diagnostic decisions are based on DSM-III criteria for the particular syndrome.

Preliminary testing of the K-SADS has been reported, although the number of children involved was quite small (Puig-Antich et al., 1978). Interrater reliability was excellent for the major diagnostic syndromes and ranged from .63 to .96 for individual symptoms. Test-retest were conducted following imipramine treatment of depressed children and demonstrated that K-SADS ratings were sensitive to changes due to drug treatment. Comparisons between mother and child interviews have also been encouraging and suggest the necessity of ascertaining diagnostic information from both sources (Chambers et al., 1979).

Work on a lifetime version and an epidemiologic version of the K-SADS (K-SADS-L and K-SADS-E) are currently underway and revision of the K-SADS (current version) is also in process (Puig-Antich, 1979). Diagnostic categories are being added such as Attention Deficit Disorder, Drug and Alcohol Abuse, and Panic Disorder. Previous categories have been further refined or augmented so that diagnostic decisions may be more easily determined. Symptoms on the K-SADS-L and the K-SADS-E are rated only for their presence or absence since these versions are not concerned with treatment effects but rather with a determination of psychiatric disorders (past or present). When completed, the K-SADS-L and K-SADS-E will be useful for epidemiologic research of psychiatric disorders in children. Reliability and validity studies of the K-SADS-L and K-SADS-E are underway.
Orvaschel, Sholomskas, Weissman  

Isle of Wight Survey

The Isle of Wight Survey (IWS) was developed by Rutter and Graham in an effort to obtain systematic information about psychiatric disorders in children. Psychiatric disorder was defined by the authors as "... abnormalities of emotions, behaviour or relationships which are developmentally inappropriate and of sufficient duration and severity to cause persistent suffering or handicap to the child and/or distress or disturbance of the family or community..." (Rutter and Graham, 1968, p. 153). The instrument was intended to be of diagnostic utility and was designed to elicit information directly from children between the ages of 7 and 12 years.

The IWS is a semistructured interview of approximately a half-hour's duration which utilizes both a categorical and dimensional approach. It contains over 100 items dealing with interpersonal and behavioral aspects of the child's life. Initial questions focus on school, peer contacts, family interactions, and spare-time activities. Following this, the child is assessed on behavior more specifically related to symptomatology, such as antisocial activities, hypochondriasis, fears, worries, and depression. Items are rated on a scale which varies from yes/no to a 4- or 5-point range. The time perspective of the interview is unclear, but appears to emphasize "current" behavior. When the interview is completed the interviewer rates the child for probable, definite, or no psychiatric disorder, and also indicates the nature of the disorder, if one is present.

To examine test-retest reliability, the IWS was administered to 89 children at two points in time using different interviewers. Agreement between interviewers for the rating of definite psychiatric disorder was .84 and for probable disorder was .51. Reliability scores for individual items were much lower, however, particularly for behaviors such as mood, attention, and activity. Additional examination of test-retest reliability showed good agreement for the extremes of behavior and was less favorable on more moderate item variation. Interrater reliability was tested on 25 children between the ages of 7 and 12 years and produced item agreement which ranged from .63 to .95. Discriminant validity was examined by comparing the interview findings for a normal population of 159 children with those of a population of 108 children considered psychiatrically disturbed on the basis of information obtained from parents and teachers. Interviewers were blind to the child's psychiatric status prior to the interview. Differentiation between the groups was good with psychiatric disorder ratings occurring far more frequently in the abnormal group than in the control group.

The IWS was an important first step in the development of a
psychiatric interview for use with preadolescent children. It achieved acceptable reliability scores on overall impairment ratings and was a useful tool for the initial discrimination between children with and without a psychiatric disorder. The interview did have difficulty obtaining good interrater agreement for many items and areas of functioning. This was due probably to shortcomings in the manner in which many questions were worded and the need for a more clearly defined rating system. The IWS was not always systematic in obtaining information for a particular diagnostic entity and the criteria for arriving at a diagnostic determination were not always clearly defined.

Conclusions

The five interviews reviewed do not exhaust the assessment instruments available for use with children. They are presented as an overview of the state of the art of interview techniques suitable for psychiatric epidemiology with children. All five interviews differ in basic structure, theoretical premise, and underlying diagnostic criteria (Feighner Criteria, RDC, DSM-III, ICD-9). The Screening Inventory, The Mental Health Assessment Form and to a lesser extent the Isle of Wight Survey are designed to assess overall functioning and impairment. They may be viewed as first line screening tools and would more than likely require follow-up assessments of identified cases in order to determine rates of specific disorders. The Diagnostic Interview for Children and Adolescents and the Kiddie-SADS are designed to function as basic screening devices which at the same time assign cases to discrete diagnostic categories. All the instruments require further testing and refinement and most of this is already underway. The advantages and disadvantages of a particular instrument have been discussed previously. The selection of one instrument over another must be based on the investigator's own theoretical orientation and the needs of the study for which it is intended.

Acknowledgments

This work was supported in part by Contract ADM 42-74-83 (DBE) from the Center for Epidemiologic Studies, National Institute of Mental Health, Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA), Rockville, Maryland, and USPHS Research Grant #1 R01 MH28274.
References

Robins, L.N.: Diagnostic Interview Schedule. Developed under contract from the Center for Epidemiologic Studies, National Institute of Mental Health, ADAMHA, Rockville, Maryland, 1979.