Family psychiatric screening instrument for epidemiologic studies: pilot testing and validation

Jennifer D. Lish^{*a}, Myrna M. Weissman^b, Philip B. Adams^b, Christina W. Hoven^b,
Hector Bird^b

^a*Clinical Neuroscience Research Unit, Medical College of Pennsylvania, 3200 Henry Avenue, Philadelphia, PA 19129, USA*

^b*Department of Psychiatry, Columbia University, College of Physicians and Surgeons, 722 West 168th Street, New York, NY 10032, USA*

Received 13 April 1994; revision received 29 November 1994; accepted 6 February 1995

Abstract

Family history, a risk factor for psychiatric disorders, is infrequently assessed in epidemiologic studies due to time and cost constraints. We designed a brief computer-scorable instrument, the Family History Screen for Epidemiologic Studies (FHE), which collects a pedigree and screens for 15 *DSM-III* diagnoses in an informant and in his family members. The FHE was administered to one informant in 77 families in which we had collected pedigrees, interviewed 77 informants and 239 relatives using the Lifetime Anxiety version of the Schedule for Affective Disorders and Schizophrenia or the Epidemiologic version of the Schedule for Affective Disorders and Schizophrenia for School-Aged Children, and performed best-estimate diagnoses. We evaluated the accuracy with which the FHE predicted best-estimate diagnoses. For adults reporting on themselves, the FHE demonstrated high levels of sensitivity and specificity for depression (67.4, 75.0) and panic (92.5, 89.2), and low sensitivity and high specificity for substance abuse (33.3, 93.6). For informants reporting on adult relatives, sensitivity was low and specificity was high for depression (35.2, 84.9), panic (20.0, 91.7), and substance abuse (42.1, 93.4). For informants reporting on children, perhaps due to lower prevalence, sensitivity and specificity were poor. The FHE is a good screen for psychiatric disorders in adult informants, but it is not useful for family history. It may be useful in primary care medical settings as a screen for psychiatric history.

Keywords: Family history; Diagnosis, psychiatric; Depression; Anxiety; Substance abuse

---

1. Introduction

Family psychiatric history has been demonstrated, in studies of both children and adults, to

be an important risk factor for the major psychiatric illnesses, including schizophrenia (Erlenmeyer-Kimling, 1968; Kety et al., 1968; Gottesman and Shields, 1982), bipolar disorder (Gershon et al., 1982; Strober et al., 1988; Kutcher and Marton, 1991), major depression (Orvaschel, 1983; Beardslee et al., 1985; Weissman et al., 1988;

---

* Corresponding author, Tel: +1 215-842-4235; Fax: +1 215-843-4321.

Puig-Antich et al., 1989), alcohol abuse (Goodwin, 1979; Famularo et al., 1985; Tarter and Edwards, 1986), obsessive-compulsive disorder (Swedo et al., 1989), and other anxiety disorders (Crowe et al., 1983; Last et al., 1987; Turner et al., 1987; Biederman et al., 1991). However, some epidemiologic and clinical studies that have attempted to identify risk factors for psychiatric disorders, such as the Epidemiologic Catchment Area (ECA) study, have excluded assessment of family psychiatric history. Omission of family history as a risk factor may obscure or confound results.

The Family History Research Diagnostic Criteria (FH-RDC; Andreasen et al., 1977) cover most of the psychiatric disorders that have been shown to be familial. This diagnostic instrument is usually used in connection with other methods for systematically enumerating the subject's biological relatives (e.g., Thompson et al., 1979). The FH-RDC has excellent test-retest reliability (0.90; Zimmerman et al., 1988) and an acceptable level of validity, or ability to predict the diagnoses that the subject will receive on a direct semistructured diagnostic interview (Thompson et al., 1982; Andreasen et al., 1986; Lish et al., 1993). The specificity of the FH-RDC (i.e., ability to detect true noncases) is uniformly excellent, but the sensitivity (i.e., ability to detect true cases) varies from good to fair, depending upon the specific diagnosis, and the relationship of the informant to the subject (Thompson et al., 1982; Andreasen et al., 1986). Thus, there is generally underreporting of *DSM-III* Axis I disorders on the FH-RDC, as compared with diagnoses obtained in direct interviews.

The FH-RDC is used widely in clinical family studies. Including it in epidemiological studies, in which family history is not the major focus, is problematic, because the FH-RDC requires more administration time, professional expertise, or training than can realistically be made available, especially in a large community survey. The FH-RDC takes between 10 and 50 min per family member whose psychiatric history is reported. Furthermore, the FH-RDC is designed to be administered by experienced clinicians. For that reason, some items simply cue the interviewer to determine the presence in the family member of a

symptom described in technical terms (e.g., nihilistic delusions) without providing a question in ordinary English to be addressed to the informant. Interviewers are expected to use their knowledge of clinical psychiatry to administer and score the FH-RDC. The FH-RDC can be administered to multiple family members, and the resulting information can be combined.

It should be possible to develop a valid measure of family history that can be administered quickly. An internally consistent scale whose reliability is in the range of 0.90 can shed up to three-quarters of its items without losing much of its criterion validity, or sensitivity and specificity (Shrout and Yager, 1989). The questions of the FH-RDC may be considered to be a set of scales, and the test-retest reliability of several FH-RDC diagnoses is in the specified range.

As part of the development of methodology for a community-based epidemiologic study of the rates and risk factors for psychiatric disorders in children ages 9–17 planned by the National Institute of Mental Health, we undertook development of a brief instrument to assess the history of psychiatric disorders in a child's family members. The screen performs three functions: (1) identification of the biological relatives to be assessed (pedigree collection); (2) screen of lifetime history of psychiatric disorders in the informant (who is one of the biological relatives); and (3) screen of lifetime history of psychiatric disorders, as reported by the informant, in the other biological relatives identified on the pedigree. This article is the first report of the development and testing of this instrument.

2. Methods

2.1. Instrument design

The Family Psychiatric History Screen for Epidemiologic Studies (FHE) is administered to a single individual in each family (the informant), who reports on himself or herself as well as on the other relatives. It can be administered by telephone. The FHE was designed for use with a child's mother or other adult guardian. The FHE interviewer first collects the relevant pedigree by identifying the child's immediate biological

Table 1
Example of pedigree collection^a

Relationship	First name	Sex	Age	Deceased?	ID
Biological mother:	Frances	F	41	No	01
Biological father:	Gordon	M	38	No	02
First-born child:	Rebecca	F	13	No	03
Second-born child:	Ethan	M	10	No	04

^a“Please tell me the first names and ages of [Rebecca’s] biological parents. Now tell me the names, ages and sexes of all the biological children that [Frances] had with [Gordon], including [Rebecca]. Please start with the first-born and continue until you get to the last child they had together.”

relatives (his parents and full biological siblings) and assigning a unique identification (ID) number to each family member. Table 1 shows how this information is elicited and recorded.

The interviewer then asks the informant questions about the presence or absence of psychopathology, over the course of the entire lifetime, in the family members, including the informant. If the informant answers affirmatively to any question, the interviewer determines specifically which family member had the symptom, and records the name and ID number. Thus, in the FHE, the psychopathology questions are each posed once, about all of the family members as a group. This procedure differs from that used in the FH-RDC, in which the psychopathology questions are asked multiple times, once about each family member. We expected that inquiring about all family members at once would decrease the amount and quality of information available about each individual, but we felt that this sacrifice was required by the time constraints, and would be minimal in the typically small families of contemporary 9- to 17-year-olds. Information about age of onset, although of interest, was also sacrificed in the interest of time.

The FHE items are modeled on the “gate questions” in the Lifetime Anxiety version of the Schedule for Affective Disorders and Schizophrenia (SADS-LA; Mannuzza et al., 1989); those questions which determine whether a set of additional questions should be administered. The gate questions in the SADS-LA usually correspond to the *DSM-III* Criterion A for the diagnosis. For most of the diagnoses, the SADS-LA gate

question has been modified by the addition of another symptom or two, to reduce the frequency of false-negative cases, and by mentioning exclusion criteria, or emphasizing the frequency, persistence, duration, or deviance of the symptoms, or the functional impairment caused by them, to reduce the frequency of false-positive cases. Since inclusion of this material lengthened the questions, 11 items were split into two parts. The second part is administered only if the response to the first part is positive and the response to the question is considered to be in the affirmative; the identity of the individual for whom the question is positive is recorded only if the answer remains affirmative after both parts have been administered. Two or three separate questions (some of which are in two parts) were included for seven diagnoses. For these diagnoses, an individual is considered “screen positive” if the answer to any of the questions is affirmative.

For example, the mania/hypomania question (see Table 2) is in two parts, specifies a minimum duration, excludes behavior attributed to substance use, and requires either that the person’s functioning be impaired, or that others become concerned about the person. Also, for example, the two panic attack/limited symptom attack questions shown in Table 2 mention three symptoms (anxiety, dyspnea, and tachycardia), but also mention some exclusion criteria (medical illness, exercise, and frightening events) and specify that the symptoms must recur frequently.

2.2. Scoring

Fifteen diagnostic categories are assessed with

Table 2

Sample questions from the Family History Screen for Epidemiologic Studies

Mania/hypomania

Question 13: Has anyone on the list [of family members] ever had a period of at least 2 days in which they were extremely happy or high, so much so that people thought it was unusual? This does not include behavior that was caused by taking drugs or alcohol or that occurred at a time when it would be normal to be very happy.

If no, skip to next question. If yes, say:

Question 13A: I mean someone being "high as a kite," in a mood that was really abnormal, and that would make you worry about them, or interfere with their ability to carry out their normal responsibilities. Has anybody been like that for 2 days or more?

If no, skip to next question. If yes, say: "Who?" and record name and ID number only of those persons for whom answers to both Questions 13 and 13A were "yes." Then say, "anyone else?"... repeat until no additional names are given.

Panic attack/limited symptom attack

Question 15: Has anyone on the list [of family members] had many periods or spells of a half-hour or less in which they suddenly felt extremely frightened or terrified even though there was really nothing to be afraid of? (*Instructions as above.*)

Question 16: Has anyone on the list [of family members] had many periods or spells of a half-hour or less in which they suddenly had difficulty breathing and had a rapid, pounding heartbeat that weren't caused by heart problems, exercise, or something that would have terrified most people? (*Instructions as above.*)

25 questions, 11 of which are in two parts. A subject is considered to be screen positive for a diagnosis if the informant answers affirmatively for him to any of the questions for that diagnosis (one or more of which could be a two-part question). The conditions assessed include (number of interview questions per diagnosis shown in parentheses); major depression/dysthymia (3), mania/hypomania (2), panic attack/limited symptom attack (2), generalized anxiety disorder/overanxious disorder (1), agoraphobia/simple phobia (1), other simple phobia (1), social phobia (1), obsessive-compulsive disorder (2), psychosis (2), conduct disorder (2), antisocial personality disorder (2), attention deficit hyperactivity disorder (1), alcohol abuse/dependence (1), drug abuse/dependence (1), and suicide attempt (1). The four diagnostic categories that are referred to by compound names (e.g. mania/hypomania) are assumed not to distinguish between the two disorders. The FHE generates lifetime diagnoses. Thus, each family member is rated as either lifetime screen positive or screen negative for each of the 15 diagnostic categories. Each individual is also scored as positive or negative on five items regarding treatment

and functional impairment that precede the symptom questions; however, these questions were not used in any of the analyses reported herein. Family members who are reported to have been treated or hospitalized, but for whom the informant does not answer affirmatively to one of the screens for specific disorders, are considered not to have any psychiatric illness.

2.3. Validity study: selection of families

To test the validity of the FHE quickly and economically, we selected families from among those enrolled in a family aggregation study in which semistructured interviews had been performed by clinically trained interviewers between 1 and 4 years previously. These families had been selected from among the New Haven area participants in the ECA survey (Robins and Regier, 1991) or from among patients at specialty clinics for anxiety and depression at Yale University (Weissman et al., 1993). Although the long variable interval between the SADS-LA or K-SADS-E (Epidemiologic version of the Schedule for Affective Disorders and Schizophrenia for School-Aged Children; Orvaschel et al., 1982) in-

terviews and the FHE interviews compromises the ability of the validity study of the FHE to attain its goals, the availability of the SADS-LA and K-SADS-E data collected previously enabled us to pilot the FHE rapidly.

To maximize the statistical power of the validity study, we selected families in which (a) the proband had a large number of first-degree relatives; (b) a relatively high proportion of the proband's first-degree relatives had been interviewed; (c) some of the proband's first-degree relatives were minor children; (d) a relatively large number of diagnoses had been made in the family; or (e) the proband or his relatives had been diagnosed with disorders that were uncommon in this study. Using these guidelines, we selected 111 families, from among the approximately 300 that were enrolled in the family aggregation study on that date, as candidates for participation in the FHE validity study.

This study was approved by the institution's ethical review board, and informed consent was obtained from informants. Of the 111 probands selected for the FHE validity study, three could not be reached by telephone, and three refused to participate. After a quota of 77 probands had participated in the study, 11 consenting probands were not interviewed, and the remaining 17 probands were not contacted. Thus, 77 informants provided screening information. In these families, we had previously and independently obtained direct interviews from 316 individuals (77 probands; 53 spouses; 13 parents; 45 siblings; 49 adult offspring; and 76 minor children, ages 6–17).

2.4. FHE training and administration

This study used recently hired B.A.-level lay interviewers (R.H. and K.M.) who were unaware of the family selection procedure and of previous diagnostic results. The interviewers received 1 h of training in FHE administration. The interviewers then administered the FHE by telephone to one adult informant in each family. That is, a single informant in each family reported on the psychopathology of all of the biological first-degree relatives identified in the pedigree, including the informant. This single informant was the proband in the family aggregation study in 67 families, and

the proband's spouse in 10 families in which the proband was not available for interview. Administration of the FHE, including pedigree collection and psychiatric history on informant and relatives, took an average of 23 min (range 8–34 min).

2.5. Best-estimate diagnoses

The diagnoses used as the validity criteria for the FHE were best-estimate lifetime diagnoses of *DSM-III* Axis I disorders made as part of a family aggregation study. The best-estimate diagnoses were based on all available information from multiple informants (excluding the FHE). The principal source of information was a direct interview, by a clinically trained person (e.g., M.S.W., M.A. in psychology) who used the SADS-LA for adults or the K-SADS-E for children.

Training in administering SADS-LA and K-SADS-E interviews included a 1-week course, observation of interviews, supervised interviewing, and ongoing supervision by the field director (J.D.L.). Good to excellent test-retest reliability was demonstrated for the SADS-LA on a separate sample of 38 adults (panic disorder: $\kappa = 0.65$; major depression: $\kappa = 0.92$). All K-SADS-E interviews were conducted in person. Approximately two-thirds of the SADS-LA interviews with adults were conducted by telephone, and the remainder were conducted in person; rates of disorders and interview characteristics did not differ by mode of interview administration (Sobin et al., 1993).

Best-estimate diagnoses were made separately for each relative by a psychiatrist or Ph.D. clinical psychologist who had not been involved in FHE or SADS-LA/K-SADS-E data collection and was unaware of the diagnoses of other family members. The exception to this was the field director (J.D.L.), who also functioned as a "best estimator" and was therefore excused from involvement in best-estimate diagnoses of cases in which she knew the interview diagnoses of other family members. Every case was reviewed independently by at least one other psychiatrist or Ph.D. clinical psychologist. If there was no disagreement for any diagnosis, the first best-estimate diagnosis was considered final. In cases where there was disagreement, a consensus diagnosis was made by a panel including at least the original two diagnosti-

cians. Diagnoses were made at the definite and probable levels; only definite diagnoses are used as the validity criterion in the present study. In a separate reliability study with all 11 best estimators, using both definite and probable diagnoses, on a sample of 83 interviewed probands and relatives, the interrater reliability κ coefficient for a combined category of *DSM-III* panic disorder and agoraphobia with panic attacks was 0.83, and the κ coefficient for RDC major depression was 0.66. The κ value for agreement between the two best estimators who completed the majority of the cases (J.D.L. and E.H.) was 0.86 for panic disorder and 0.82 for major depression. For further details of procedures, see Weissman et al. (1993) and Goldstein et al. (1994).

2.6. Pedigree collection

The pedigrees used as the validity criterion for the FHE pedigrees were collected by the SADS-LA/K-SADS-E interviewers using the method described by Thompson et al. (1979).

2.7. Data analysis

Diagnoses generated by the FHE, administered to one informant in each family, were compared to *DSM-III* best-estimate diagnoses based on direct SADS-LA or K-SADS-E interviews of each member of the family. Results are reported for individual diagnoses, in which instance a true-positive case is a subject who receives the same specific diagnosis on both FHE and best-estimate evaluations. Results are also reported for categories of diagnoses (i.e., any anxiety disorder), in which instance a true-positive case is a subject who receives any diagnosis in the category on the FHE, and any diagnosis in the category on best-estimate assessment (not necessarily the same diagnosis). We also report results for the overall summary category of "any FHE diagnosis." No results are reported for diagnoses received by fewer than five persons. The following conventions for descriptions of validity results were adopted: for *specificity*, excellent = 0.80–1.00, good = 0.60–0.79, fair = 0.40–0.59, poor = 0.20–0.39, and very poor = 0–0.19; for *sensitivity*, excellent = 0.70–1.00, good = 0.50–0.69, fair = 0.30–0.49, and poor = 0–0.29. These guidelines are not absolute; the desired sensitivity

and specificity depend on the intended use of the instrument (see Discussion). However, these evaluative guidelines are useful for the present study.

2.8. Definitions of terms

The *sensitivity* of a test is the percentage of persons with a disorder who are correctly identified by that test as having that disorder. *Specificity* is the percentage of persons without the disorder who are correctly identified by a test as not having the disorder. *Positive predictive value* (PPV) is the proportion of the actual cases of the disorder in the sample that are identified by the screen, or the likelihood that a person with a positive test actually has the disorder. *Negative predictive value* (NPV) is the proportion of actual noncases among respondents found to be screen negative, or the likelihood that a person who is screen negative is actually free of the disorder. PPV and NPV are affected by the prevalence of the disorder, and by the reliability of measurement of both cases and noncases. The κ values for sensitivity and specificity are measures of sensitivity and specificity that are corrected for chance agreement. The κ values are scaled from 0 to 1, with 0.00 indicating lowest sensitivity/specificity and 1.0 indicating perfect sensitivity/specificity (Glaros and Kline, 1988; Kraemer, 1992).

3. Results

3.1. Accuracy of pedigree collection

Twenty-seven pedigrees collected by the FHE were compared with those collected in the family aggregation study. Of 226 persons who had been identified as relatives of these 27 probands by the standard technique, the name and relationship to the proband were correctly ascertained by the FHE in 224/226 (99%) of the cases.

3.2. FHE validity

Sensitivity and specificity. Table 3 shows the sensitivity and specificity of the FHE for informants reporting on (a) themselves, (b) other adults, and (c) minors. The number of persons who received each diagnosis is shown in parentheses.

For adult informants reporting on themselves,

Table 3

Sensitivity (SE) and specificity (SP) of the Family History Screen for Epidemiologic Studies (FHE) compared with definite best-estimate *DSM-III* diagnoses

Diagnoses	Informants reporting on								
	Selves (<i>n</i> = 77)			Adult relatives (<i>n</i> = 160)			Minors (<i>n</i> = 79)		
	SE	(<i>n</i>)	SP	SE	(<i>n</i>)	SP	SE	(<i>n</i>)	SP
Major depression/dysthymia	67.4	(49)	75.0	35.2	(54)	84.9	0	(6)	97.3
Any affective disorder	67.4	(49)	75.0	36.4	(55)	84.8	0	(6)	97.3
Generalized anxiety disorder/ overanxious disorder	57.1	(14)	88.9	60.0	(5)	94.8			
Panic attack/limited symptom attack	92.5	(40)	89.2	20.0	(15)	91.7			
Agoraphobia	88.0	(25)	76.9	16.7	(6)	97.4			
Simple phobia									
Social phobia									
Separation anxiety disorder							20.0	(5)	89.2
Any anxiety disorder	90.5	(42)	68.6	52.2	(23)	81.8	0.0	(6)	91.8
Conduct disorder				42.9	(7)	92.2			
Suicide attempts	50.0	(8)	98.5	20.0	(10)	98.7			
Alcohol abuse/dependence	22.2	(9)	94.1	36.7	(30)	93.9			
Drug abuse/dependence	44.4	(9)	98.5	38.9	(18)	97.2			
Any substance abuse/ dependence	33.3	(15)	93.6	42.1	(38)	93.4			
Any FHE diagnosis	88.7	(62)	73.3	62.2	(82)	71.8	36.4	(11)	73.5

Note. Results are not reported for cells with fewer than five subjects.

specificity is excellent to good for each individual diagnosis and for all summary categories. Sensitivity is excellent for any FHE diagnosis, any anxiety disorder, panic attack/limited symptom attack, and agoraphobia/simple phobia; good for major depression/dysthymia, suicide attempt, and generalized anxiety disorder; fair for drug abuse/dependence and any substance abuse; and poor for alcohol abuse/dependence.

For informants reporting on adult relatives (offspring, siblings, and parents), specificity is excellent for the individual diagnoses and summary categories, except that it is only good for any FHE diagnosis. Sensitivity is good for generalized anxiety disorder, any anxiety disorder, and any FHE diagnosis; fair for major depression/dysthymia, any affective disorder, conduct disorder, alcohol abuse/dependence, drug abuse/dependence, and any substance abuse/dependence; and poor for panic attack/limited symptom attack, agoraphobia/simple phobia, and suicide attempt.

For informants reporting on their minor children, the number of subjects with the target diagnosis was too small to evaluate FHE validity for most categories. Specificity was excellent for major depression/dysthymia, any affective disorder, separation anxiety disorder, and any anxiety disorder; and good for any FHE diagnosis. Sensitivity was fair for any FHE diagnosis; and poor for separation anxiety disorder, any anxiety disorder, major depression/dysthymia, and any affective disorder.

Positive and negative predictive value. For the informant, PPV is poor for alcohol abuse/dependence, fair for any substance abuse and generalized anxiety disorder, good for agoraphobia, and excellent for major depression/dysthymia, any affective disorder, panic attack/limited symptom attack, any anxiety disorder, suicide attempt, drug abuse/dependence, and any FHE diagnosis. NPV is excellent to good for all diagnoses (Table 4). For the informant reporting on adult relatives, PPV is

Table 4

Positive and negative predictive values (PPV, NPV) and κ values for sensitivity (SE) and specificity (SP) of the Family History Screen for Epidemiologic Studies (FHE) compared with definite best-estimate *DSM-III* diagnoses

Diagnoses	Informants reporting on											
	Selves (<i>n</i> = 77)				Adult relatives (<i>n</i> = 160)				Minor relatives (<i>n</i> = 79)			
	PPV	NPV	κ		PPV	NPV	κ		PPV	NPV	κ	
			SE	SP			SE	SP			SE	SP
Major depression/ dysthymia	0.83	0.57	0.32	0.52	0.54	0.72	0.17	0.31	0.00	0.92	0.03	0.08
Any affective disorder	0.83	0.57	0.32	0.52	0.56	0.72	0.18	0.32	0.00	0.92	0.03	0.08
Generalized anxiety disorder/overanxious disorder	0.53	0.90	0.47	0.43	0.27	0.99	0.57	0.25				
Panic attack/limited symptom attack	0.90	0.92	0.84	0.80	0.20	0.92	0.12	0.12				
Agoraphobia	0.65	0.93	0.79	0.48	0.20	0.97	0.14	0.17				
Simple phobia												
Social phobia												
Separation anxiety disorder									0.11	0.94	0.10	0.05
Any anxiety disorder	0.78	0.86	0.74	0.51	0.32	0.91	0.38	0.21	0.00	0.92	0.08	0.08
Conduct disorder					0.20	0.97	0.37	0.16				
Suicide attempts	0.80	0.94	0.47	0.78	0.50	0.95	0.18	0.47				
Alcohol abuse/depend- ence	0.33	0.90	0.16	0.25	0.58	0.87	0.28	0.48				
Drug abuse/dependence	0.80	0.93	0.41	0.45	0.64	0.93	0.34	0.59				
Any substance abuse/ dependence	0.56	0.85	0.25	0.45	0.67	0.84	0.32	0.56				
Any FHE diagnosis	0.93	0.61	0.52	0.65	0.70	0.64	0.30	0.38	0.18	0.88	0.12	0.05

Note. Results are not reported for cells with fewer than five subjects.

poor for the anxiety disorders and conduct disorder, and NPV is excellent to good for most disorders. PPV for minor children for all disorders is poor, and NPV for minor children for all disorders is excellent.

Sensitivity and specificity. The κ values for sensitivity and specificity are presented for completeness and, in general, are lower, reflecting in part the base rates in the sample.

Results by source of informant. It is of interest to determine the psychometric characteristics of the FHE in community samples in epidemiologic studies. For the 77 informants reporting on themselves only, for the category any FHE diagnosis, sensitivity was 95.0 for the 51 clinic-derived subjects and

88.9 for the 26 ECA-derived subjects. Specificity was 100.0 for clinic subjects and 87.5 for ECA subjects. PPV was 1.00 for clinic subjects and 0.94 for ECA subjects, and NPV was 0.33 for clinic and 0.78 for ECA subjects. Presumably, the decrement in NPV for clinic subjects reporting on themselves occurred because of the small number of true non-cases in the clinic informant sample (namely, only spouses of clinic probands acting as informants).

4. Discussion

The prerequisites for a family history instrument to be used in an epidemiological population study are that it (a) be capable of rapid administration

by a layperson after minimal training, (b) accurately collect the pedigree, and (c) generate diagnoses with acceptable reliability and criterion validity (sensitivity and specificity). The present study was intended to determine whether the FHE met those requirements. We report a comparison of the FHE, a brief screen for 15 psychiatric diagnoses administered by lay interviewers, against best-estimate diagnoses made independently and blindly by doctoral-level persons based on all available information except the FHE, informants drawn from clinic and epidemiologic samples, and their first-degree relatives. The results showed the following:

- The FHE was used successfully by two lay interviewers after minimal training. Although administration time was long (8–23 min), this is attributable to our deliberate selection of large families with many ill members. In a pilot study for the child epidemiologic study with community residents, FHE administration time averaged 8 min (C.W. Hoven et al., unpublished data).
- The FHE was highly accurate in pedigree collection (i.e. identifying first-degree relatives by name and relationship).
- FHE specificity, for the informant reporting on himself, was excellent to good for all diagnoses, whereas sensitivity was excellent to good for panic, agoraphobia, depression, suicide attempt, generalized anxiety disorder, any anxiety disorder, and any FHE diagnosis, and fair to poor for drug and alcohol disorders. The PPV reflected these results. The overall κ coefficients of agreement were usually lower, reflecting in part the base rates in the sample. The results for the overall diagnoses did not vary by whether the informant came from a clinic or the ECA.
- As with all family history measures, the results were considerably poorer for reports by informants on either adult or child relatives. Specificity was excellent for individual diagnoses and many summary categories, and it was good for any diagnosis for adults and minors. For adults, sensitivity was good only for generalized anxiety disorder, any anxiety, and

any FHE diagnosis; it was fair to poor for other diagnoses. For minors, sensitivity for individual diagnoses either could not be evaluated or was poor.

The criterion validity of the FHE must be evaluated in light of the potential uses of the instrument. There are two ways in which data on the family psychiatric history collected in an epidemiologic study of children may be used: (a) as a measure of a risk factor to determine how family psychiatric history predicts childhood psychiatric illness in the population or (b) as a screen in a two-stage study to select children with particular profiles of family psychiatric history for further study with some more definitive techniques (Newman et al., 1990).

The two potential uses for a family history screen carry with them specific requirements with respect to predictive validity. In a two-stage study, one wishes to avoid spending limited resources to conduct the second-stage procedure on false-positive cases generated by the screen, even at the expense of missing some cases. The optimal screening measure thus has excellent PPV; it maximizes specificity, or ability to identify noncases, in exchange for some loss of sensitivity, or ability to identify cases. Thus, the ideal measure, for use to select children with particular family histories for further study, would be *highly specific and somewhat insensitive*.

For the purpose of using the results of the measure in statistical analyses of the role of family history in the etiology of childhood psychiatric disorder, it is desirable to have as valid a measurement as possible, or to have an instrument that is *high on both sensitivity and specificity*. This is most desirable for disorders of high prevalence (e.g., depression or alcoholism). However, for disorders whose base rate in the population is quite low (e.g., schizophrenia, mania), small drops in specificity can increase the number of false-positive cases so greatly that, given less than perfect sensitivity, random prediction based on the base rate of the disorder would detect a larger number of true cases than would use of the screen (Glaros and Kline, 1988). Therefore, for low base rate disorders, for the purpose of use to predict

child psychopathology from parent disorders of low prevalence, the ideal instrument would also be *highly specific and somewhat insensitive*.

For adult informants, the levels of specificity and sensitivity achieved by the FHE are reasonably appropriate for use in predicting child disorder from any parental disorder. However, it would be preferable, for this purpose, for the FHE to achieve higher levels of sensitivity, at the expense of some specificity, for most disorders. This is especially true for the disorders of highest prevalence in the population (i.e., major depression and alcohol abuse) because there is less danger, in those cases, that the number of false-positive cases generated by a measure with low specificity will exceed the number of additional true-positive cases resulting from an increase in sensitivity. For adult informants reporting on their own psychiatric history, the FHE has acceptable degrees of sensitivity and specificity, for most disorders, for the purpose of a two-stage study. Indeed, the level of validity obtained for informants is so high that the FHE has promise as a very brief screen for psychiatric illness in adult respondents in studies or clinical settings that do not require a measure of family psychiatric history. In fact, an investigator interested in collecting information on the psychiatric status of the biological parents of offspring might best use the FHE as a brief telephone interview directly with each parent. Another possible use of the FHE is as a screen for psychiatric history in primary care medical settings.

For related adults reported on by an informant, the degree of sensitivity and specificity achieved by the FHE may be acceptable for use in a two-stage study, although it would again be preferable to sacrifice some of the specificity for additional sensitivity.

For uninterviewed minors, since FHE validity is either poor or could not be evaluated in this study, we cannot recommend its use as a measure of pathology in siblings either for prediction of target subject pathology or for selection of families with uninterviewed ill children.

4.1. Limitations

The present estimates of the validity of the FHE were potentially susceptible to two kinds of bias.

First, the long interval between the SADS-LA or K-SADS-E interview and the FHE may have made it appear to have lower specificity because of new onsets of disorders in persons previously free of those disorders. Second, inclusion of probands selected from treatment clinics could have resulted in overestimates of the FHE's sensitivity and specificity, because test-retest reliability and diagnostic stability are higher for patients than community residents (Bromet et al., 1986; Williams et al., 1992). However, we did not find differences in FHE accuracy by source of subject. The low base rate of diagnoses in minor children in this sample is a major limitation to the study. Also, we expected that inquiring about all family members at once would decrease the amount and quality of information available about each individual, but we felt that this sacrifice was required by the time constraints and would be minimal in the typically small families of contemporary 9- to 17-year-olds. Finally, information about age of onset, although of great interest, was sacrificed in the interest of time.

We recommend use of the FHE for rapid screening for psychiatric disorders in an adult informant. However, we do not recommend use of the FHE for collecting family history information, because of its demonstrated low sensitivity and specificity for persons other than the informant. This instrument, whether administered by a lay person or in written self-report format, may have utility as a screen for psychiatric disorders in adult respondents in settings other than epidemiologic studies — for example, in primary care medical clinics.

Acknowledgments

This study was supported by grants MH-43878, MH-28274, MH-36197, MH-46718, and MH-30906 from the National Institute of Mental Health. The authors thank Patricia Cohen, Ph.D., Michael Fendrich, Ph.D., Prudence Fisher, M.S., Virginia Warner, M.P.H., and Helena Kraemer, Ph.D. for advice; Robert Hammond and Karen Morrison for conducting FHE interviews; and Ewald Horwath, M.D., for participating in the best-estimate diagnostic assessments.

References

- Andreasen, N.C., Endicott, J., Spitzer, R.L. and Winokur, G. (1977) The family history method using diagnostic criteria: reliability and validity. *Arch Gen Psychiatry* 34, 1229–1235.
- Andreasen, N.C., Rice, J., Endicott, J., Reich, T. and Coryell, W. (1986) The family history approach to diagnosis: how useful is it? *Arch Gen Psychiatry* 43, 421–429.
- Beardslee, W.R., Keller, M.B. and Klerman, G.L. (1985) Children of parents with affective disorders. *Int J Fam Psychiatry* 6, 283–299.
- Biederman, J., Rosenbaum, J.F., Bolduc, E.A., Faraone, S.V. and Hirshfeld, D.R. (1991) A high-risk study of young children of parents with panic disorder and agoraphobia with and without comorbid major depression. *Psychiatry Res* 37, 333–348.
- Bromet, E.J., Dunn, L.O., Connell, M.M., Dew, M.A. and Schulberg, H.C. Long-term reliability of diagnosing lifetime major depression in a community sample. *Arch Gen Psychiatry* 43, 435–440.
- Crowe, R.R., Noyes, R., Pauls, D.L. and Slymen, D. (1983) A family study of panic disorder. *Arch Gen Psychiatry* 37, 77–79.
- Erlenmeyer-Kimling, L. (1968) Studies on the offspring of two schizophrenic parents. In: Rosenthal, D. and Kety, S. (Eds.), *The Transmission of Schizophrenia*. Pergamon Press, Oxford, UK.
- Famularo, R., Stone, K. and Popper, C. (1985) Preadolescent alcohol abuse and dependence. *Am J Psychiatry* 142, 1187–1189.
- Gershon, E.S., Hamovit, J., Guroff, J.J., Dibble, E., Leckman, J.F., Sceery, W., Targum, S.D., Nurnberger, J.I., Goldin, L.R. and Bunney, W.E. (1982) A family study of schizoaffective, bipolar I, bipolar II, unipolar and normal control probands. *Arch Gen Psychiatry* 39, 1157–1167.
- Glaros, A.G. and Kline, R.B. (1988) Understanding the accuracy of tests with cutting scores: the sensitivity, specificity and predictive value model. *J Clin Psychol* 44, 1013–1023.
- Goldstein, R.B., Weissman, M.M., Adams, P.B., Horwath, E., Lish, J.D., Charney, D., Woods, S.W., Sobin, C. and Wickramaratne, P.J. (1994) Psychiatric disorders in relatives of probands with panic disorder and/or major depression. *Arch Gen Psychiatry* 51, 383–394.
- Goodwin, D.W. (1979) Alcoholism and heredity. *Arch Gen Psychiatry* 36, 57–61.
- Gottesman, I.I. and Shields, J. (1982) *Schizophrenia: The Epigenetic Puzzle*. Cambridge University Press, New York.
- Kety, S.S., Rosenthal, D., Wender, P.H. and Schulsinger, F. (1968) The types of prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. In: Rosenthal, D. and Kety, S. (Eds.), *The Transmission of Schizophrenia*. Pergamon Press, Oxford, UK.
- Kraemer, H. (1992) *Evaluating Medical Testings: Objective and Quantitative Guidelines*. Sage Publication, Newberry Park.
- Kutcher, S. and Marton, P. (1991) Affective disorders in first-degree relatives of adolescent-onset bipolars, unipolars and normal controls. *J Am Acad Child Adolesc Psychiatry* 30, 75–79.
- Last, C.G., Phillips, J.E. and Statfeld, A. (1987) Childhood anxiety disorders in mothers and their children. *Child Psychiatry Hum Dev* 18, 103–112.
- Lish, J.D., Gyulai, L., Resnick, S.M., Kirtland, A., Amsterdam, J., Whybrow, P.C. and Price, R.A. (1993) A family history study of rapid-cycling bipolar disorder. *Psychiatry Res* 48, 37–45.
- Mannuzza, S., Fyer, A.J., Martin, L.Y., Gallops, M.S., Endicott, J., Gorman, J., Liebowitz, M.R. and Klein, D.F. (1989) Reliability of anxiety assessment: I. Diagnostic agreement. *Arch Gen Psychiatry* 46, 1093–1101.
- Newman, S.C., Shrout, P.E. and Bland, R.C. (1990) The efficiency of two-phase designs in prevalence surveys of mental disorders. *Psychol Med* 20, 183–193.
- Orvaschel, H. (1983) Parental depression and child psychopathology. In: Guze, S.B., Earls, F.J. and Barrett, J.E. (Eds.), *Childhood Psychopathology and Depression*. Raven Press, New York, pp. 53–56.
- Orvaschel, H., Puig-Antich, J., Chambers, W., Tabrizi, M.A. and Johnson, R.A. (1982) Retrospective assessment of prepubertal major depression with Kiddie SADS-E. *J Am Acad Child Adolesc Psychiatry* 21, 392–397.
- Puig-Antich, J., Goetz, D., Davies, M., Kaplan, T., Davies, S., Ostrow, L., Asnis, L., Twomey, J., Iyengar, S. and Ryan, N.D. (1989) A controlled family history study of prepubertal major depressive disorder. *Arch Gen Psychiatry* 46, 406–418.
- Robins, L.N. and Regier, D.A. (Eds.) (1991) *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study*. Macmillan Publishing Co., Inc., New York.
- Shrout, P.E. and Yager, T.J. (1989) Reliability and validity of screening scales: effect of reducing scale length. *J Clin Epidemiol* 42, 69–78.
- Sobin, C., Weissman, M.M., Goldstein, R.B., Adams, P.B., Wickramaratne, P.J., Warner, V. and Lish, J.D. (1993) Diagnostic interviewing for family studies: comparing telephone and face-to-face interview methods for the diagnosis of lifetime psychiatric disorders. *Psychiatr Genet* 3, 227–233.
- Strober, M., Morrell, W., Burroughs, J., Lampert, C., Danforth, H. and Freeman, R. (1988) A family study of bipolar disorder in adolescence: early onset of symptoms linked to increased familial loading and lithium resistance. *J Affect Disord* 15, 255–268.
- Swedo, S., Rapoport, J., Leonard, H., Lenane, M. and Cheslow, D. (1989) Obsessive-compulsive disorder in children and adolescents. *Arch Gen Psychiatry* 46, 335–344.
- Tarter, R.E. and Edwards, K. (1986) Antecedents to alcoholism: implications for prevention and treatment. *Behav Ther* 17, 346–361.
- Thompson, W., Kidd, K. and Weissman, M.M. (1979) A procedure for the efficient collection and processing of pedigree data suitable for genetic analysis. *J Psychiatr Res* 5, 291–303.
- Thompson, W., Orvaschel, H., Prusoff, B. and Kidd, K.K. (1982) An evaluation of the family history method for ascertaining psychiatric disorders. *Arch Gen Psychiatry* 39, 53–58.
- Turner, S.M., Beidel, D.C. and Costello, A. (1987) Psychopathology in the offspring of anxiety disorder patients. *J Consult Clin Psychol* 55, 229–236.

- Weissman, M.M., Adams, P.B., Lish, J.D., Wickramaratne, P.J., Horwath, E., Charney, D., Wood, S.W., Leeman, E. and Frosch, E. (1993) The relationship between panic disorder and major depression: a new family study. *Arch Gen Psychiatry* 50, 767–780.
- Weissman, M.M., Gammon, G.D., John, K., Merikangas, K.R., Warner, V., Prusoff, B.A. and Sholomskas, D. (1987) Children of depressed parents: increased psychopathology and early onset of major depression. *Arch Gen Psychiatry* 44, 847–853.
- Weissman, M.M., Merikangas, K.R., John, K., Wickramaratne, P., Prusoff, B.A. and Kidd, K.K. (1986) Family-genetic studies of psychiatric disorders: developing technologies. *Arch Gen Psychiatry* 43, 1104–1116.
- Weissman, M.M., Warner, V., Wickramaratne, P. and Prusoff, B.A. (1988) Early-onset major depression in parents and their children. *J Affect Disord* 15, 269–278.
- Williams, J.B.W., Gibbon, M., First, M.B., Spitzer, R.L., Davies, M., Borus, J., Howes, M.J., Kane, J., Pope, H.G., Jr. and Rounsaville, B. (1992) The Structured Clinical Interview for *DSM-III-R* (SCID): II. Multi-site test-retest reliability. *Arch Gen Psychiatry* 49, 630–636.
- Zimmerman, M., Coryell, W., Pfuhl, B. and Stangl, D. (1988) The reliability of the family history method for psychiatric diagnoses. *Arch Gen Psychiatry* 45, 320–322.