META-ANALYSIS

Meta-analytic Evidence for Familial Coaggregation of Schizophrenia and Bipolar Disorder

Jared X. Van Snellenberg, MA, MPhil; Teresa de Candia, BA

Context: Several data sources suggest a link between schizophrenia and bipolar disorder (BD); however, family studies have not revealed coaggregation of these disorders.

Objectives: To systematically review family studies of probands with schizophrenia and BD and to determine whether these disorders coaggregate in families.

Data Sources: MEDLINE and PsycINFO databases.

Study Selection: All family studies published from January 1, 1980, to December 31, 2006, reporting morbid risk or raw counts of schizophrenia or BD in first-degree relatives (FDRs) of a proband group with DSM-III or later; International Classification of Diseases, Ninth or Tenth Revision; or research diagnostic criteria schizophrenia or BD were included. Of the original 2326 studies identified through the database search, 38 studies were used to investigate rates of BD in FDRs of probands with schizophrenia, while 39 studies were used to examine rates of schizophrenia in FDRs of BD probands.

Data Extraction: Data were analyzed with a novel random-effects bootstrapping technique that allows for the inclusion of studies lacking a patient or control group, which made up a substantial portion of the available data. Data were also blindly weighted for methodological quality.

Data Synthesis: The FDRs of probands with schizophrenia showed significantly ($P = .01$) increased rates of BD relative to control FDRs (odds ratio [OR] = 2.08). The FDRs of probands with BD showed marginally ($P = .06$) increased rates of schizophrenia relative to control FDRs (OR = 2.10); this analysis was significant ($P = .02$) when studies that did not report morbid risk estimates were excluded (in this case, OR = 3.49).

Conclusions: This meta-analysis provides direct evidence for familial coaggregation of schizophrenia and BD, a finding that argues against the view that these disorders are entirely discrete diagnostic entities. Rather, a continuum model is supported.

Arch Gen Psychiatry. 2009;66(7):748-755

AN ENDURING QUESTION IN the genetics of major mental illness is the extent of overlap between schizophrenia and bipolar disorder (BD). Genetic linkage studies have identified susceptibility loci that both disorders share, and the only twin study that has assigned diagnoses nonhierarchically (such that individuals meeting criteria for multiple disorders would be assigned all diagnoses) found significant genetic overlap among schizophrenia, BD, and schizoaffective disorder (SD). Furthermore, family studies have shown familial coaggregation of SD and unipolar depression in both schizophrenia and BD. However, evidence for direct familial coaggregation of schizophrenia and BD is scarce, and the replicability of linkage studies that show shared susceptibility genes between these disorders has been poor. An extensive literature search (see “Methods” section) failed to reveal any studies that demonstrated increased risk of BD in the families of probands with schizophrenia, as did an earlier meta-analysis of 3 family studies. While at least 2 studies of BD probands have revealed increased risk of schizophrenia, many more have not.

Given the shared genetic loading for SD and unipolar illness in both schizophrenia and BD, a lack of familial aggregation between these 2 disorders is somewhat puzzling. A potential explanation for negative findings is lack of power, because statistical tests for binary events with low expected frequencies suffer from low power. If the true risk of BD in the general population is 1% and the odds ratio (OR) of risk of BD in first-degree relatives (FDRs) of probands with schizophrenia over control probands is small (1.5 in this example), a family study would require assessment of

Author Affiliations:
Department of Psychology, Columbia University, New York, New York (Mr Van Snellenberg); and University of Colorado at Boulder, Boulder (Ms de Candia).

©2009 American Medical Association. All rights reserved.
roughly 7800 FDRs to achieve 50% power to detect an effect (assuming symmetrically distributed error variance); to our knowledge, no family studies of either schizophrenia or BD have approached this size. This power problem is due both to the low prevalence of these disorders in the population and, in this example, to the low OR. Thus, if the true risk of BD were instead 10%, one would need fewer than 900 FDRs to achieve 50% power with an OR of 1.5; whereas if the risk were 1% but the OR was 5, one would need fewer than 300 FDRs. Therefore, while the available literature rules out rates of BD in FDRs of patients with schizophrenia that are in excess of 3 or 4 times the normal rate, existing studies have not had sufficient power to detect effects much smaller than this (naturally, the same argument applies to rates of schizophrenia in the families of patients with BD).

One way to increase power is to combine the results of multiple studies using meta-analysis. However, many family studies lack an appropriate control group and so cannot be used in traditional meta-analyses that calculate a measure of association (eg, $\Phi$, a correlation coefficient for contingency tables) for each study and assess significance across studies, because these measures require a control group. Consequently, we developed a novel meta-analytic procedure that allows inclusion of all studies reporting rates of schizophrenia or BD in family members of patient or control probands (and not necessarily both) to determine whether there is evidence of familial coaggregation of schizophrenia and BD.

**METHODS**

**STUDY ASCERTAINMENT**

We attempted to identify all family studies that (1) were published in English between January 1, 1980, and December 31, 2006; (2) included schizophrenia or BD proband groups; (3) reported lifetime counts or morbid risks of schizophrenia or BD in probands’ FDRs; and (4) used 1 of the following diagnostic criteria for schizophrenia or BD: the research diagnostic criteria of Spitzer et al,† DSM-III, DSM-III-R, or DSM-IV-TR; or International Classification of Diseases, Ninth (ICD-9) or Tenth (ICD-10) Revision. We identified candidate studies through a search of the MEDLINE and PsyCINFO databases using the keywords listed in the eTable (available at http://www.archgenspsychiatry.com), resulting in 2326 hits.

When 2 or more studies investigated the same proband group, we compared the quality of data from each study and used only the best available data (see below), such that data were only included in the analysis once. In a few cases, we used 1 study’s data on a proband group’s familial rates of schizophrenia and another study’s data on the same proband group’s familial rates of BD. None of the probands included in the meta-analysis had a diagnosis of BD, as it is known to coaggregate with both schizophrenia and BD. When multiple studies reported on the same probands, the study we included was based on the following criteria (in decreasing order of importance): (1) recency of diagnostic criteria used; (2) diagnostic methods used (eg, direct interviews preferred to questionnaires); (3) number of data sources used for diagnosis; (4) size of proband and relative samples; (5) availability of morbid risk counts in preference to lifetime counts; and (6) morbid risk methods, preferring methods that given lower penalty scores in the weighting procedure described below.

A total of 66 studies were retained for meta-analysis; of these, 12 reported rates of BD in relatives of schizophrenic probands, 6,7,16-25 13 reported rates of schizophrenia in relatives of BD probands, 6,10,14,26-31 and 6 reported both rates of BD in relatives of schizophrenic probands and rates of schizophrenia in relatives of BD probands.‡§ This resulted in a total of 18 studies reporting rates of BD in at least (exact numbers were not available for all studies) 6896 relatives of 2084 schizophrenic probands and 19 studies reporting rates of schizophrenia in at least 7066 relatives of 1312 BD probands. These numbers of relatives are minimums, because some studies report only the number of relatives at risk after estimation of morbid risk—commonly known as Bezugsziffer—and not the total number of relatives assessed.

Therefore, 31 of the 66 studies included the data of primary interest, while the remaining 35 studies reported rates of either BD or schizophrenia in relatives of probands with the same disorder. Of all of the 66 studies, any that included data on rates of schizophrenia or BD in relatives of control probands were used in our primary analyses. Thus, from the full set of 66 studies, in addition to the data already described, 38 studies* reported rates of schizophrenia in 24 106 relatives of 5366 probands with schizophrenia, 32 studies† reported rates of BD in 12 598 relatives of 2576 probands with BD, 20 studies* reported rates of schizophrenia in 8536 relatives of 1571 control probands, and 18 studies‡ reported rates of BD in 8640 relatives of 1777 control probands.

**META-ANalytic PROCEDURES**

**Statistical Procedures**

Unfortunately, 8 of the studies that reported familial rates of BD in probands with schizophrenia and 15 of the studies that reported rates of schizophrenia in the families of BD probands did not report data on a control group; if we used a procedure that could not accommodate this missing data, there would only be 10 and 4 studies, respectively, remaining in these analyses. Thus, to include an adequate number of studies in our analyses, we had to use a procedure that would allow us to treat the absence of a control group as missing data—an approach that is valid, provided that data missingness is not systematically related to the outcome of interest (ie, studies lacking a control group do not have systematically higher or lower diagnostic rates than other studies).

An intuitive approach is to simply add up the raw counts across all of the available studies into a single contingency table and carry out a $\chi^2$ test. However, this is a fixed-effects approach that is not appropriate in cases in which there is meaningful between-study variance. Because the studies included in our analysis vary on a large number of potentially relevant dimensions (eg, diagnostic criteria, methods of proband ascertainment, and diagnostic methods in family members), it is reasonable to expect there to be meaningful sources of variance between studies, in which case fixed-effects models inflate type I error rates beyond the nominal value of $\alpha$, sometimes quite substantially.¶ Consequently, we developed a meta-analytic technique that allows missing studies to be included in a random-effects model and used bootstrapping procedures to assess the statistical significance of the results. To our knowledge, this is the only valid random-effects approach that allows for the inclusion of studies with missing data. All analyses were carried out using Matlab (The MathWorks Inc, Natick, Massachusetts).

†References 6, 7, 16-18, 20-25, 33, 36-61.
§References 6-8, 10, 16-18, 21, 23-25, 31, 40, 63-65, 69, 76.
First, the table $P$, with proband diagnoses (patient or control) in rows and relative diagnoses (including no diagnosis) in columns, is defined as

$$P_{ij} = \frac{\sum_{k=1}^{K} \frac{d_{ijk}}{n_{i,j,k}} \left( \sum_{k=1}^{K} \frac{w_k}{n_{i,j,k}} \right)}{\sum_{k=1}^{K} \frac{w_k}{n_{i,j,k}}}$$

where $P_{ij}$ is the value in the $i$th row and the $j$th column of $P$, there are $K$ studies in the analysis that report complete data for the $i$th row, $w_k$ is the study quality weight for the $k$th study (see “Study Weighting” section), $d_{ijk}$ is the number of observations in the $i$th row and the $j$th column for the $k$th study, and $n_{i,j,k}$ is the number of observations in the $i$th row of the $k$th study (the marginal value for the $i$th row). Thus, $P_{ij}$ gives the weighted proportion of relatives of probands with the $j$th diagnosis who are expected to receive the diagnosis in the $i$th column across the $K$ studies, with the weight given to each study determined by the product of the study quality weight, $w_k$, and the square root of the sample size (which is inversely proportional to the variance of each study).

Once the table $P$ is calculated, the correlation coefficient $\Phi$ and the OR can be computed; however, because only 1 contingency table (with all values bounded by 0 and 1) is calculated, it is not possible to obtain $P$ values or confidence intervals (CIs) using parametric statistics such as $\chi^2$. However, bootstrapping can be used to obtain the empirical distribution of $\Phi$ and the OR, and bootstrapped CIs and $P$ values can be calculated. For each of the analyses described below, we used 100,000 bootstrap resamplings to calculate a bias-corrected bootstrap estimate. The full procedure for calculating the bias-corrected bootstrap for the table $P$ is described in the eAppendix.

In the course of analyzing the data, it became clear that the OR produces a very large bias estimate in this procedure owing to the extreme skewedness of its distribution. That is, because the OR is bounded by 0 and infinity but has an expected value of 1 under the null hypothesis, the distribution of the statistic has a strong positive skew. The bias estimate correctly detects this as a strong positive bias, and so the bootstrapped values of the statistic are substantially reduced, resulting in a loss of power relative to an unbiased statistic. Because $\Phi$ is approximately normally distributed when its value is close to 0, its estimated bias is very small. Consequently, for each of our analyses, we report bootstrapped $P$ values and CIs only for $\Phi$ and use the OR simply as a descriptive statistic.

We performed separate analyses to determine whether FDRs of patients with schizophrenia exhibit increased rates of BD and whether FDRs of patients with BD exhibit increased rates of schizophrenia. In addition, we also report on analyses of the rate of schizophrenia in FDRs of patients with schizophrenia and the rate of BD in FDRs of patients with BD. While it is already well known that both of these disorders show strong familial aggregation, a report of the average $\Phi$ or OR across studies could be valuable to some readers, and so we report them herein.

Another issue with combining studies in this literature is that some studies report morbid risk estimates (in which the number of relatives assessed for a disorder is reduced by some proportion to account for the fact that some are expected to develop a disorder later in life, after the study is carried out), while others report raw counts of the number of ill and healthy relatives. Including both types of data in a single analysis results in substantial variance between the 2 types of study, because raw counts will produce systematically lower estimates of risk than morbid risk estimates. Consequently, we used the entire set of studies reporting both raw counts and morbid risk of schizophrenia and BD and calculated the mean reduction in the number of relatives at risk after morbid risk estimation (ie, Bezugsziffer divided by the raw number of relatives assessed) for each disorder (for schizophrenia, mean=0.717 [standard deviation=0.149]; for BD, mean=0.634 [standard deviation=0.166]). We then reduced the number of relatives used for calculating $P$ (ie, $n_{i,j}$) from studies that reported only raw counts by this proportion, so that risk estimates from both study types would be on par. We additionally reduced the weight given to studies using raw counts (see “Study Weighting” section) to reflect the fact that the morbid risk adjustment we used reflects an estimate, rather than actual morbid risk estimation in the study in question. Finally, to evaluate the extent to which inclusion of studies reporting only raw counts influenced the results of our analyses, we carried out an additional set of analyses that excluded these studies altogether.

### Study Weighting

To give greater weight to studies of higher methodological quality, we calculated a penalty score, $Q_k$, for each included study. Various methodological qualities were assigned a penalty value, all of which were summed to generate the penalty score for that study. We assigned penalty values blind to study outcomes. Weights were then calculated as $w_k=1/(2+Q_k)$. The mean of the weights was 0.276 (standard deviation=0.074) and the range was 0.150 to 0.447. In addition to giving greater weight to higher quality studies, this procedure is used to minimize any between-study variance due to higher or lower diagnostic rates being systematically reported by studies of poorer methodological quality. To ensure that weighting has not unduly influenced the results, we report results for both weighted and unweighted data (in which all values of $w$ are set to 1) for each of our analyses. To the extent that the weighted and unweighted analyses produce comparable findings, one can be confident that the weighting procedure is not driving the results.

Higher penalty scores were given for the following methodological features: (1) diagnostic criteria differing from those of the DSM-IV; (2) direct clinical interviews not used for proband diagnosis; (3) relative diagnoses assigned without a blinding procedure; (4) relative diagnoses based solely on medical or other records; (5) use of family history method (as opposed to direct interviews of relatives), with lower penalties for using more family members as informants; (6) no assessment of interrater reliability or use of consensus diagnosis for relatives’ diagnoses; (7) no morbid risk estimation; (8) use of morbid risk procedures that do not directly use age-of-onset information in the data (eg, the Weinberg method); (9) use of the modified Stromgren procedure, as it has been shown to be heavily biased; (10) inclusion of probands with other comorbid psychiatric disorders; (11) use of a special population (eg, adolescent-onset patients, adoptees, and twins); (12) use of only drug responders or nonresponders, or only male or female probands; (13) use of an unrepresentative control group (students and nurses); (14) relaxed diagnostic criteria (eg, inclusion of individuals with schizophreniform disorder); (15) Bezugsziffer had to be estimated based on the reported age-of-onset distribution and number of relatives assessed, because it was not reported directly; and (16) only probands with type I or type II BD were included, not both, because of the possibility that one of these types could show more or less coaggregation with schizophrenia than BD as a whole (too few studies reported both separately to adequately assess this possibility). In cases in which a penalty applied only to some proportion of the relatives assessed (eg, when some number of assessed relatives were interviewed directly...
RESULTS

RATES OF BD IN FDRs OF PROBANDS WITH SCHIZOPHRENIA

The risk of BD in FDRs of probands with schizophrenia was 0.99% compared with 0.48% in control probands (OR=2.08; Φ=0.030). This difference was significant (P=.02; OR=2.02; Φ=0.030). As with the more inclusive analysis, these results were significant (P=.02; Φ=0.030). Again, little change was seen in the unweighted analysis, with risks at 1.04% for FDRs of probands with schizophrenia and 0.54% for control probands (P=.03; OR=1.95; Φ=0.029; 95% CI=0.003-0.052).

RATES OF SCHIZOPHRENIA IN THE FDRs OF PROBANDS WITH BD

The risk of schizophrenia in the FDRs of probands with BD was 1.77% compared with 0.85% in control probands (OR=2.10; Φ=0.041). This difference was marginal (P=.06; 95% CI for Φ=−0.001 to 0.077). The distribution and 95% CI of Φ for these data are shown in Figure 2. Results in the unweighted analysis were similar, with risk of schizophrenia in the FDRs of BD probands at 1.74% and at 1.00% for control probands. This difference was not significant (P=.13; OR=1.77; Φ=0.032; 95% CI=0.010 to 0.072).

When studies that did not report morbid risk data were excluded from the analysis, the estimated risk of schizophrenia was 2.04% in the FDRs of BD probands and 0.59% in the FDRs of control probands (P=.02; OR=3.49; Φ=0.063; 95% CI=0.013-0.108). This result was also significant in the unweighted analysis, with risk in the FDRs of BD probands at 2.10% and 0.60% in control probands (P=.01; OR=3.57; Φ=0.065; 95% CI=0.014-0.110).
The findings of the present study are inconsistent with categorical models of psychiatric disorder that treat schizophrenia and BD as separate, unrelated illnesses. While there are obviously important clinical distinctions between these disorders, the data are consistent either with a continuum from psychotic to affective illness on which diagnostic categories are imposed or with a set of distinct disease entities with overlapping symptomatology that are not distinguished effectively in current psychiatric nosology. Either model fits well with findings that FDRs have a very high risk of developing the same disorder as the parents of patients with schizophrenia and bipolar disorder, respectively. While this does not unequivocally demonstrate that misdiagnosis does not account for a portion of the observed effects, the rates of misdiagnosis required to account for our findings seem much too large. Similarly, given that SD aggregates in the families of individuals with both schizophrenia and BD, it is also possible that misdiagnosis of schizoaffective family members as having either schizophrenia or BD could drive our results. However, based on observed rates of SD in the families of patients with schizophrenia and BD, 15% of FDRs of patients with SD and 58% of FDRs of patients with BD would have to be misdiagnosed to account for our data.

Another potential explanation for our findings is that elevated rates of disorder are reported in the FDRs of patient probands owing to a surveillance or reporting bias in family history studies. This is unlikely for a number of reasons. First, in most of the studies in our analysis, relative diagnoses were assigned blind to proband diagnosis (and the influence of unblinded studies was reduced with the weighting procedure); so, in these studies there is no reason to expect the overreporting of illness to differentially affect diagnostic rates in patient and control FDRs. Indeed, the rates of disorder in control FDRs are at the low end of the lifetime prevalence generally found for these disorders (0.5%-1.3% for schizophrenia and 0.4%-1.6% for BD), suggesting that the occurrence of disorder is not being overreported. Additionally, even if there were some degree of bias acting on FDRs of patient probands but not control probands, it seems unlikely that such bias could be of a sufficient magnitude to account for our data—that is, diagnoses of the nonindex disorder would have to be twice as likely (given ORs > 2.0) for FDRs of patient than control probands, a level of bias that seems unreasonably high.

Finally, a concern with many meta-analyses is publication bias, in which studies reporting null or small effects may be underrepresented in the literature because of difficulty in publishing null results, thereby resulting in overestimation of the true effect size. However, there is no reason to suspect that publication bias has influenced our results, because only 2 studies in our analysis reported significant aggregation of schizophrenia in the families of probands with BD, and none reported significant aggregation of BD in the families of probands with schizophrenia. Thus, nearly all of the included studies had null results. Furthermore, many of the studies in our analyses were not specifically interested in detecting coaggregation of these disorders, so no reporting bias would be expected in these studies.

The findings of the present study are inconsistent with categorical models of psychiatric disorder that treat schizophrenia and BD as separate, unrelated illnesses. While there are obviously important clinical distinctions between these disorders, the data are consistent either with a continuum from psychotic to affective illness on which diagnostic categories are imposed or with a set of distinct disease entities with overlapping symptomatology that are not distinguished effectively in current psychiatric nosology. Either model fits well with findings that FDRs have a very high risk of developing the same disorder as...
index probands, an intermediate risk of developing schizoaffective disorder regardless of whether the index proband has schizophrenia or BD,\textsuperscript{8,10} and a low, but still increased, risk of developing BD for FDRs of probands with schizophrenia and vice versa. This pattern of results would be expected if all 3 disorders share genetic liability, but FDRs who develop one of these disorders are more likely to develop a disorder at the same point on the psychotic-affective continuum as the index case. Alternatively, it could be that some individuals with a familial disorder whose symptomatology overlaps with schizophrenia and BD (such as SD) meet current diagnostic criteria for schizophrenia or BD,\textsuperscript{81} with familial aggregation leading to an apparent continuum. Unfortunately, we did not systematically obtain data pertaining to SD, so we do not have meta-analytic estimates of coaggregation among SD, schizophrenia, and BD, which could help differentiate these 2 possibilities.

An important implication of these results is that genetic linkage studies could increase their power to identify susceptibility genes by treating all of the disorders along this spectrum as a single phenotype. Given that there is shared genetic risk of developing schizophrenia, BD, and SD, a plausible model is that some number of genes confer risk of psychotic-affective disorder in general, while another small set of genes—or epigenetic or environmental factors—determine where on the continuum the disorder will fall if it develops. If this is the case, treating 1 of the disorders as a different phenotype than the disorder of interest would substantially reduce the power to detect susceptibility genes that confer risk to disorders across the psychotic-affective continuum.

Our results, while novel, fit well within a set of findings in the literature that suggest a link between schizophrenia and BD. In addition to the previously discussed findings of shared susceptibility loci and high rates of comorbidity when using nonhierarchical diagnoses, it is worth noting that these disorders share a number of epidemiological\textsuperscript{83} and perinatal\textsuperscript{86} risk factors. Indeed, in recent years a number of authors have argued for a degree of continuity between these disorders, in opposition to the traditional Kraepelinian dichotomy.\textsuperscript{1,2,8,11,21,86,87} In light of the findings presented herein that schizophrenia and BD coaggregate, it is unlikely that these 2 disorders represent truly discrete and unrelated disease entities. While there are obviously important distinctions between schizophrenia and BD, some degree of relationship between them clearly exists—at a minimum there is a shared genetic risk factor between at least a subset of individuals with either disorder (suggesting a shared etiology), while at most they may be different behavioral expressions of a common, or at least substantively overlapping, pathology.

Submitted for Publication: August 14, 2008; final revision received December 15, 2008; accepted January 2, 2009.

Correspondence: Jared X. Van Snellenberg, MA, MPhil, Department of Psychology, Columbia University, 406 Schermerhorn Hall, 1190 Amsterdam Ave, Mail Code 5501, New York, NY 10027 (jared@psych.columbia.edu).

Author Contributions: Mr Van Snellenberg takes full responsibility for the integrity of the data and the accuracy of the data analysis; Mr Snellenberg and Ms Candia had full access to the data throughout the study.

Financial Disclosure: None reported.

Additional Information: The eTable and eAppendix are available at http://www.archgenpsychiatry.com.

Additional Contributions: Niall Bolger, PhD, provided comments on the bootstrapping procedures used for the analyses; Joshua Weinberg, BS, assisted in developing Matlab scripts for data analysis; and Lucas Sjulson, MD, PhD, provided valuable comments on an earlier version of the manuscript.

REFERENCES


