Comment on Lidz and Associates' Critique of the Danish-American Studies of the Offspring of Schizophrenic Parents

William M. Grove, M.A.
Comment on Lidz and Associates’ Critique of the Danish-American Studies of the Offspring of Schizophrenic Parents

William M. Grove, M.A.

The author replies to a critique by Lidz and associates of the Greater Copenhagen Adoption Study conducted by Rosenthal and colleagues. A detailed examination of the reanalysis by Lidz and associates of tabulated data from the original adoption study reports substantiates only one of their objections. The author argues that, because of the way in which the original data were gathered, the reanalysis is without substantial merit in evaluating the adoption study.


In an article in the August 1981 issue of this journal, Lidz and associates (1) criticized the study by Rosenthal and associates (2, 3) on the prevalence of schizophrenia spectrum disorders in the adopted-away offspring of schizophrenic parents. Rosenthal and associates (3) found spectrum disorders in 31.6% of offspring with a schizophrenic parent and in 17.9% of adopted-away offspring of control parents. Although the original investigators concluded that “the evidence supports the theory that heredity plays a significant role in the etiology of schizophrenia spectrum disorders” (3, p. 310), Lidz and associates stated, “Their study of the adopted-away offspring of schizophrenic parents fails to provide definite or statistically significant evidence of a genetic factor in the etiology of schizophrenic disorders” (1, p. 1067). It is disquieting to the reader to see such varied conclusions drawn from the same body of data. The purpose of this comment is to show that the criticisms of Lidz and associates have little force.

DESIGN OF THE ORIGINAL STUDY

It is necessary to recapitulate the design of the Rosenthal and associates study in order to understand

Received Feb. 1, 1982; revised July 19, 1982; accepted Aug. 23, 1982. From the Department of Psychology, University of Minnesota, Minneapolis. Address reprint requests to Mr. Grove, N438 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455.

Copyright © 1983 American Psychiatric Association.

the controversy. Working in Copenhagen, the investigators culled more than 10,000 records of biological parents of over 5,000 adoptees in order to locate a number of parents who had been registered as hospitalized for a psychiatric disorder. Adoptees were restricted to those legally adopted between 1924 and 1947 by a nonrelative early in life. The former restriction was evidently due to a paucity of central records on informal adoptions, and the latter were necessary to separate the effects of shared genes and shared environments. The hospital records for registered biological parents were abstracted in a standardized way, translated into English, and independently diagnosed by at least two diagnosticians. A diagnostic scheme consisting of the following categories was used: 1) not in the schizophrenia spectrum; 2) definite chronic or process schizophrenia, definite reactive schizophrenia, and definite latent or borderline schizophrenia; 3) schizoid or inadequate personality; 4) uncertain chronic or process schizophrenia, uncertain reactive schizophrenia, and uncertain latent or borderline schizophrenia; and 5) manic-depressive illness (3, p. 309). Meanings of these designations were generally taken from DSM-II.

All those adoptees whose parents were judged by consensus as having definite schizophrenia were defined as index cases. A few offspring of parents judged by consensus as having manic-depressive illness were also included, both because “there were periods when we simply did not have enough schizophrenic parents processed and the staff in Copenhagen had no subjects to examine” and because “we might also learn something about the possible genetic relationship between schizophrenia and manic-depressive psychosis” (2, p. 382). Kety and associates (4) pointed out that “manic-depressive illness was never thought to be in the schizophrenia spectrum by us” (p. 417). However, statements by Rosenthal and associates that their “data suggest that the inherited core diathesis is the same for both schizophrenia and manic-depressive psychosis” (2, p. 387) and that “we include both types of subjects in the spectrum tentatively” (3, p. 309), stripped of qualification, may have led even the most careful readers to the opposite conclusion. The balance of the index group comprised offspring of parents who were judged by at least one diagnostian as being in
the spectrum but for whom there was no diagnostic consensus.

Seventy-nine such adoptees were eventually selected. The investigators chose, for each index adoptee, a control adoptee who was listed on the adoption register close to the index case and closely matched in sex, age, age at placement, and social class of adoptive family. Although the adoptees were thus paired, there was some attrition, and they have always been reported in published analyses in unpaired fashion. Thus in 1968, data on 39 index adoptees and 47 control adoptees were analyzed, but in 1971 the comparable figures were 76 and 67, respectively.

Each adoptee was contacted, and cooperating adoptees were given a semistructured psychiatric interview lasting 3–5 hours, conducted by Dr. Welner (or, in some cases, Dr. Schulsinger). In every case the interviewer made diagnoses essentially according to the same scheme as that of the parental diagnoses, but more idiosyncratic terms were allowed (e.g., “prepsychotic,” “vague schizoid tendencies”). These diagnoses are the ones analyzed by Rosenthal and associates (2, 3) and reanalyzed by Lidz and associates (1).

LIDZ AND ASSOCIATES’ CRITIQUE

Lidz and his colleagues criticized the Rosenthal reports on several grounds. These are 1) an alleged discrepancy between that portion of the data reported in 1968 and that added in 1971, 2) what Lidz and associates regarded as a remarkably small amount of frank schizophrenia in the whole sample, and 3) problems in selecting the index adoptees. Trimming the sample of index adoptees who they believed should never have been included, Lidz and associates reanalyzed the data and found no significant difference between index and control adoptees in frequency of spectrum disorders. I will take these points in order in separating out what I consider well-taken from what I think are overstated criticisms.

The first point is that the illness rates from the 1971 portion of the series did not much resemble those reported in 1968. Lidz and associates wrote, "The failure of the 1971 series to replicate the 1968 findings raises serious doubts about the validity of the findings" (1, p. 1066). This objection seems to assume that if the second half of data collected in a study does not closely resemble the first half, a "failure . . . to replicate" has occurred. This assumption is invalid for three reasons. First, it fails to square with the way follow-up studies are actually conducted. Second, it uses the term "replicate" in a soleticisic way. Third, it makes too much of small numbers of observations, ignoring the role of sampling error in producing differences of outcome in a series of observations.

On the first point, in a follow-up study the last subjects observed are usually those who were hard to find or for whom observation was put off until more important subjects had been secured. In this study, the offspring of diagnostically certain parents were usually ascertained first, those of diagnostically uncertain parents second. It is obvious that these two groups would differ diagnostically. Because it is a different question whether the offspring of diagnostically uncertain parents should have been observed at all, I postpone discussion of this matter until later.

Second, use of the word "replicate" could lead readers to misunderstand the original study. The observations added in the 1971 report were not a replication but a completion of the study reported in preliminary fashion in 1968. If a study is planned as a whole, but one has reason to report on it while data are still being collected, this is in and of itself no reason to call a later, more complete report a "replication." This is not merely a semantic issue, because Lidz and his colleagues' comment makes it sound as if two studies had been conducted, one failing to confirm the other. Of course, divergence in diagnoses between two parts of a series is well worth examining. The difference may illuminate causes of the illness under study. However, divergence (even if statistically significant) between parts of a series is not a mark against a study unless it leads one to suspect that some experimental procedure changed during the study, making the early and later results incomparable. There is no reason to think that this occurred in the Rosenthal study.

Lidz and associates did point out a relevant apparent difference between the first and second parts of the adoptee series, namely, that the latterly observed offspring of parents judged schizophrenic by consensus were apparently less ill than the formerly observed ones. However, they failed to report that this difference was not statistically significant (x^2 with Yates' correction = .78, df = 1, p < .37). This brings out an important problem in evaluating adoption studies, one these critics ignored: With such small numbers, even fairly large apparent differences can be nonsignificant. Without additional considerations, nonsignificant trends are difficult to interpret because sampling error competes with substantive hypotheses as an explanation.

Another criticism Lidz and associates advanced is that too little chronic schizophrenia was found to support a genetic hypothesis (1, p. 1066). They returned to this point in their discussion, writing that "one important outcome of the [reanalysis of the] study is the failure to replicate Heston's findings [5; their reference 18], which had been widely accepted as strong evidence that schizophrenia is essentially a genetic disorder" (1, p. 1068). However, the position that "schizophrenia is essentially a genetic disorder" is not at issue. Neither Rosenthal, his colleagues, nor Heston came to such a conclusion. More to the point, the results are not in nearly as much conflict with Heston's as Lidz and associates would have had the reader believe. Shields and associates (6, p. 180) pointed out that if one considers that subset of the
Rosenthal adoptees who were the offspring of parents who were possibly chronically schizophrenic or were judged by consensus as having the disorder, three of 44 offspring (7%) were chronically schizophrenic, compared with five of 47 offspring (11%) in Heston’s sample.

If this (nonsignificant) difference needs explanation, it may be found in several facts. First, the rate of nuclear schizophrenia in Danish mothers who are allowed to put a child up for adoption is evidently lower than in American populations. This is because a Danish mother may obtain a legal abortion on the grounds of “a history of schizophrenia in herself or in the father” (7, p. 360). Moreover, those who are psychotic at childbirth may not be allowed to place the child for adoption. A related point is that Heston’s schizophrenic mothers were all hospitalized at the birth of the index adoptee; just 10% of the Rosenthal mothers were hospitalized before the index adoptee’s birth (3, p. 308). Because the peak years of onset for schizophrenia in women come later than the peak childbearing years, most schizophrenic mothers fall ill after bearing a child (8, p. 15). Those hospitalized for schizophrenia before childbirth are more severely ill as a group than those not so hospitalized. For at least one type of relative, severity in the proband correlates with increased morbid risk in the relatives: Severely schizophrenic monozygotic twins are usually found to have more ill cotwins than are twins with milder cases of the illness (9, pp. 225–230 and the references cited therein). Although a finding that less severely ill probands have fewer ill relatives than more severely ill probands is important, it does not argue against a role for genes in schizophrenia.

A final objection Lidz and associates made was that the offspring of seven parents judged manic-depressive by consensus were included as index adoptees. They wrote, “It is erroneous to include them as part of an index series selected to find out if adopted-away children of schizophrenic parents develop schizophrenia or schizophrenia spectrum disorders” (1, p. 1066).

Indeed, if schizophrenic and manic-depressive psychoses are relatively distinct genetically, as is commonly thought, then including the offspring of manic-depressive parents will not be informative in a study such as Rosenthal and associates wanted to conduct. It is interesting that these affectively ill parents contributed two cases to the spectrum count, although one adoptee was diagnosed only as having either schizophrenia, manic-depressive illness. Including this case in the spectrum gives a rate of 28.6%, not less than the 26.9% rate for offspring of parents judged schizophrenic by consensus. However, the numbers are too tiny to bear speculations about the breadth of the spectrum. It would have been a considerably clearer report if these offspring had not been taken as index cases. This point was originally made by Gottesman and Shields (10) in a review of the study. However, Lidz and associates recomputed the index-control comparison with the offspring of manic-depressive parents left out. I will try to show here why their reanalysis is misleading.

Lidz and colleagues objected to the inclusion of six offspring of parents whom at least one judge called manic-depressive, as well as to the inclusion of 11 offspring of parents whom at least one diagnostician judged as in the spectrum but for whom there was no consensus. Lidz and associates thought that these parents, like the manic-depressive parents, should never have been included in a study of adopted-away offspring of schizophrenic parents. However, whether one includes parents diagnosed with or without consensus is a difficult judgment and depends on several considerations. First, is the parent’s diagnosis closer to schizophrenia than to some competing diagnosis? Second, is it likely that including the parents diagnosed without consensus will inflate the count of spectrum offspring for reasons other than their genetic contribution to these offspring? Third, what are the scientific costs associated with mistakenly excluding the parents with nonconsensus diagnoses?

The first point cannot be fully evaluated without recourse to the detailed case notes on each index parent. One knows only that parents judged by consensus as not having a spectrum disorder who were also not manic-depressive were not included as index parents. It would be much easier for the critical reader to evaluate these diagnoses if the case material were published.

The second point is rather a reason for including parents diagnosed without consensus. What except genes could such parents contribute to the spectrum count in their offspring? Study of half-siblings of schizophrenic adoptees suggests that uterine factors probably do not contribute most of the diathesis for later schizophrenia. The maternal half-siblings of schizophrenic patients, who share a common uterus, are no more often schizophrenic than are paternal half-siblings, who share only the paternal genetic contribution (11). Although some evidence indicates that perinatal birth complications may amplify the risk of schizophrenia in genetically low-risk adoptees (12), and although such obstetrical complications could be more common in index parents with nonconsensus diagnoses than in control parents, the effect seems too weak to explain all of the spectrum pathology in offspring of index parents with nonconsensus diagnoses. The data from offspring of a parent mildly ill with a spectrum disorder and a schizophrenic parent suggest that the less severe spectrum illnesses may be genetically related to schizophrenia. These matings were associated with elevated risk for spectrum disorders in the Rosenthal adoptees (13), but Kety and associates (11) could not confirm this association.

A difficulty with including parents with nonconsensus diagnoses in the index series is the “specificity problem.” If parents with nonconsensus diagnoses produce more offspring with schizophrenia spectrum disorders, is this because they themselves had such disorders? Or is it because they contributed a genetic
factor predisposing to psychopathology in general, some of which fell in the schizophrenia spectrum? This question could be answered through further research. It is just possible that what is inherited is a more general predisposition to severe psychopathology than most investigators now think. Although the bulk of the evidence supports the separation of schizophrenia from such other disorders as affective illnesses, the issue has not been considered settled by all investigators (14). However, this important issue is strictly irrelevant to the question "Do genes play a part in the etiology of schizophrenia spectrum disorders?" It is irrelevant because the role of genes in schizophrenia can be decided without knowing whether predispositions to other disorders may be inherited. Moreover, it can be answered even if schizophrenia and some other clinically distinct disorder share some genetic causes. Disclosing possible genetic links between various disorders would logically seem, if anything, to follow a demonstration that each is partly caused by an inherited predisposition. Thus the "specificity problem" affects the interpretation of adoption study data regarding the specificity of a genetically transmitted diathesis for schizophrenia but not the validity of inferences from such data about the presence or absence of genetic factors in that disorder.

The third consideration to be weighed in deciding whether to exclude parents with nonconsensus diagnoses, one completely neglected by Lidz and associates, is statistical power. With fewer cases, only larger differences in the rates of spectrum disorders in offspring can be detected. Power calculations show that even with the approximately 75 subjects per group actually obtained, there was only a 50% chance of detecting a true difference in spectrum rates such as 30% versus 15%; yet this is a two-to-one relative risk. For a 70% chance of detecting such a difference, sample sizes of N=138 per group would be needed, but for an 80% or 90% chance the required numbers are 161 or 199 per group! If attention were confined to strictly schizophrenic offspring, the problem would be much worse. Clearly, if adoption studies are to furnish powerful refutations of genetic hypotheses, there is reason to include as many subjects as can be found whose observation may illuminate the transmission of various disorders. Just as clearly, the more the parental group is broadened, the more dilute the genetic hypothesis tested becomes. The hypothesis here becomes one of the heritability of spectrum disorders; it has always been tentatively reported as such. Thus it is not clear that including index parents with nonconsensus diagnoses biased the study's outcome in favor of a genetic hypothesis.

Lidz and associates excluded the offspring of manic-depressive parents and parents with nonconsensus diagnoses from the index series, recalculated the significance test for the difference between index and control adoptees, and reported that it was not significant. In so doing, they not only ignored considerations of statistical power but also failed to discuss a potentially serious statistical problem. The problem follows from the gathering of the data as pairs; ideally they should be treated as pairs. Lidz and associates used the same unpaired analysis that Rosenthal and associates used, but the magnitude of the effect of using an unpaired test was much greater in the reanalysis than in the original reports.

The reason index adoptees were matched with control adoptees was to rule out differences in spectrum rates resulting from differences between index and control adoptees on other variables such as age, sex, age at placement, and social class of the adoptive family. Matching the index adoptees with controls, rather than just collecting an unselected comparison group of adoptees who did not have a psychiatrically registered parent, increased the precision of the design (15). In the original reports, Rosenthal and associates did not report the data as pairs, presumably because attrition left some index and some control adoptees unpaired. The statistical consequence is that comparing the index series with a group of mostly matched control adoptees—but ignoring the matching—increased the estimated sampling error of the estimated rate difference between groups. The precision of the comparison was systematically underestimated, making rate differences appear less significant than they really were. This was pointed out by Rosenthal and associates in another context (16, p. 331). In the original reports, then, the significant rate differences found actually point to a stronger role for genes than indicated by the significance level reported. The use of unpaired data also changed the estimated rate difference itself. However, in the original reports there was only a modest unmatching of the groups, and the effect on the analysis was probably small.

On the other hand, Lidz and associates excluded 24 index adoptees in a study in which more than 20 index adoptees had already been lost by attrition. If we assume that losses of index and control adoptees occur independently, there were about 10 unmatched index adoptees and 34 unmatched controls in the "purified" sample of Lidz and associates. This constituted about 38% of the original data—20% of the index and 50% of the control series. Retaining all the originally reported control probands inflated the sample size with extraneous subjects who not only did not control for anything but also confused the index-control comparison with possibly confounded demographic variation. The variables chosen to match control adoptees are all known demographic risk factors for schizophrenia. The untouched control group contained about 20 adoptees originally matched to the 24 excluded index adoptees. These controls might have been at disproportionately high risk demographically. The resultant comparison might then have been made between an inflated estimate of control spectrum rates and a fair, if inexact, estimate of index spectrum rates. Such a comparison would have been biased against a genetic hypothesis. This conjecture cannot be evaluated from just the published data and is speculative at present.
It is no conjecture, however, that removal of cases greatly lowered the statistical power of the comparison in Lidz and associates' reanalysis. This, indeed, made a nonsignificant difference between index and control groups more likely, but for reasons unrelated to the existence or nonexistence of a genetic diathesis for schizophrenia spectrum disorders. Because of the foregoing considerations, the exclusion of adoptees from the sample, except for the offspring of parents diagnosed as manic-depressive by consensus and their matched controls, was not warranted by argument on the merits. Moreover, because the non-pairwise exclusion of subjects was unsound, reports of reanalyses by independent evaluators who did not have the original paired data are probably less illuminating than confusing.

CONCLUSIONS

I have tried to show that the critique by Lidz and associates is unbalanced and therefore misleading. However, the more important issue is not whether the Rosenthal Danish adoption study method is unassailable but whether the weight of the evidence supports the hypothesis that genetic factors powerfully predispose to schizophrenia. Clearly, the many family studies, twin studies, and the other adoption studies have almost without exception implicated hereditary factors (6). Although the study under criticism is an important piece of evidence, it is after all simply one of the reasons psychopathologists are almost all agreed that heredity plays a significant role in the etiology of schizophrenia.

REFERENCES


