# Premorbid IQ in Schizophrenia: A Meta-Analytic Review

**Kristen A. Woodberry, M.S.W., A.M.**

**Anthony J. Giuliano, Ph.D.**

**Larry J. Seidman, Ph.D.**

**Objective:** Over the past three decades, there have been significant changes in the diagnostic criteria for schizophrenia as well as changes in measurement of IQ. The last quantitative review of the literature on premorbid IQ in schizophrenia was published more than two decades ago. Since that time, there have been many published studies of data sets pertaining to this issue. The purpose of the present review was to provide an updated meta-analysis of premorbid IQ in individuals who later develop schizophrenia.

**Method:** The authors performed a systematic literature search, which yielded 18 studies that met criteria for the meta-analysis. Inclusion criteria were 1) premorbid psychometric measures of IQ in subjects who were later diagnosed with schizophrenia, schizoaffective disorder, or schizophreniform disorder, 2) similar comparison data, and 3) sufficient data for calculation of an effect size. The analogue to the analysis of variance method was used to model between-study variance due to key study-design features.

**Results:** Overall, schizophrenia samples demonstrated a reliable, medium-sized impairment in premorbid IQ. The heterogeneity of effect sizes was minimal and almost exclusively the result of one study. Methodological differences, such as diagnostic criteria, type of IQ measure, sample ascertainment, and age at premorbid testing, contributed minimally to the effect size variance. A cross-sectional analysis of all studies by age and a descriptive review of studies that used repeated measures of IQ in a single sample did not support the presence of a relative decline in IQ during the premorbid period in individuals with schizophrenia. However, all studies with pre- and post-onset testing within the same sample suggested that a significant decline in the IQ of individuals with schizophrenia, relative to comparison subjects, was associated with the onset of frank psychosis.

**Conclusions:** Years before the onset of psychotic symptoms, individuals with schizophrenia, as a group, demonstrate mean IQ scores approximately one-half of a standard deviation below that of healthy comparison subjects.

Schizophrenia has been consistently associated with a range of early neurodevelopmental abnormalities (1–3). One measure that may reflect early neurodevelopmental abnormality and has received considerable attention is general intellectual functioning—or IQ. Estimates of premorbid IQ are attainable through several study designs, including follow-back studies of school-, conscript-, or clinic-based testing, longitudinal birth or conscript cohort studies, and studies of samples at genetic risk for schizophrenia. The last quantitative review of the literature, which was published in 1984, suggested that premorbid IQ in individuals with schizophrenia was, on average, 0.43 standard deviations below that of comparison samples (4).

However, the diagnostic criteria for schizophrenia and related disorders became more restrictive with the publication of the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) (5) and its subsequent editions. This diagnostic narrowing is also reflected in related international systems (6) and diagnostic assessment tools such as the Research Diagnostic Criteria (RDC) (7). Studies published during the past 20 years have provided data from samples diagnosed using these more narrow criteria. Several have also used more comprehensive psychometric measures of IQ. In addition, a substantially larger group of studies with actual premorbid data is currently available.

The purpose of the present quantitative review was to re-evaluate the mean premorbid effect size of IQ in schizophrenia samples relative to comparison samples using more recent studies. The following five specific questions were considered in our assessment: 1) What is the mean effect size for premorbid IQ impairment in schizophrenia? 2) Is the mean effect size reliable? 3) Are discrepancies in verbal and nonverbal IQ present during the premorbid period? 4) Is there evidence of gender differences in premorbid IQ? 5) Is there evidence of increasing impairment over time or with age?

This article is the subject of a **CME** course (p. 659).

TABLE 1. Study Descriptors and Mean Effect Sizes for 18 Studies Included in Review and Categorized by Methodological Level[a]

| Study | Diagnostic Measure/Method | Comparison Group | Group Ns (Schizophrenia/Comparison) | Age at Testing | IQ Measures | Adjusted Effect Size[b] |
|---|---|---|---|---|---|---|
| **Level 1** | | | | | | |
| Amminger et al. (15)[c] | RDC/OPCRIT using semi-structured phone interviews and medical records | Child psychiatric patients born within the same 11 years but without adult disorder | 8/13 | Mean=13.1 years (SD=3.2) | For subjects <16: WISC (11 subtests) | −1.78 |
| | | | 5/18 | | For subjects ≥16: WAIS (11 subtests) | −0.53 |
| Brewer et al. (16) | DSM-IV/SCID | Healthy subjects similar in age (range=15–33) and gender | 18/37 | Mean=20.7 years (SD=4.3) | WAIS-R (short form) | −0.42 |
| Cannon et al. (17) | DSM-IV/ medical record review | Cohort members with no adult treatment history or sibling with schizophrenia | 72/7,941 | 4 | Stanford-Binet (form L–M) | −0.66 |
| | | | 72/7,941 | 7 | WISC (7 subtests) | −0.53 |
| Ott et al. (18) | RDC/SADS-L | Subjects from similar school districts or matched on socioeconomic status | 18/189 | Mean=9.4 years (SD=1.8), range=7–12 | WISC/WISC-R | −0.78 |
| | | | 13/146 | Mean=15.2 years (SD=2.0), range=11–19 | WISC, WISC-R, WAIS, WAIS-R | −0.66 |
| Seidman et al. (19) | DSM-IV/ best consensus SCID diagnosis and available records | Cohort members without major lifetime psychopathology | 26/59 | 7 | WISC (7 subtests) | −0.64 |
| | | | 26/59 | 35 | WAIS-R (vocabulary and block design) | −0.89 |
| Sørensen et al. (20)[d] | DSM-III-R/SADS-L and Present State Examination (current and lifetime) | Age-, gender-, social class-, and education-matched subjects at low genetic risk with no diagnosis | 35/59 | Mean=15.1 years, range=8–20 | WISC (12 subtests) | −0.21 |
| Whyte et al. (21) | ICD-10/Present State Examination/CATEGO | Age- and sex-comparable individuals from social networks of a genetic high-risk sample | 13/30 | Mean age=19.3 years (SD=2.4), (schizophrenia group) Mean age=21.7 years (SD=2.4) (comparison group) | WAIS-R | −0.82 |
| **Level 2** | | | | | | |
| Cannon et al. (22) | DSM-IV/Diagnostic Interview Schedule | Cohort members without schizophrenia, schizophreniform disorder, mania, or anxiety/depression | 36/642 | 3 | Peabody Picture Vocabulary[e] | −0.55 |
| | | | 36/642 | 5 | Stanford-Binet (form L–M) | −0.46 |
| | | | 36/642 | 7 | WISC-R | −0.48 |
| | | | 36/642 | 9 | WISC-R | −0.55 |
| | | | 36/642 | 11 | WISC-R | −0.46 |
| Jones et al. (23) | DSM-III-R/evaluation of hospital admissions and cohort surveys | Cohort members without schizophrenia | 30/4,176 | 8 | Group-administered nonverbal/verbal IQ tests | −0.40 |
| | | | 30/4,176 | 11 | Same | −0.30 −0.52 |
| | | | 30/4,176 | 15 | Same | |
| Kremen et al. (24) | ICD-9/hospital staff | One twin from remaining pairs of Vietnam Era Twin Registry sample matched on parental education | 21/860 | Mean=19.1 years (SD=1.0/1.2), (schizophrenia group/comparison group) | Armed Forces Qualification Test | −0.64 |
| Lencz et al. (25)[d] | DSM-IV/K-SADS-E and SIDP-IV | Sex- and age-matched healthy subjects | 7/39 | Schizophrenia mean=16.9 years (SD=1.4); Comparison mean=15.8 years (SD=2.7) | WISC-III or WAIS-R vocabulary and block design | −1.03 |
| Reichenberg et al. (26)[d] | ICD–10/ board-certified psychiatrist | Non-hospitalized remainder of birth cohort | 1,856/549,466 | 17 | Draft Board Tests | −0.57 |

*(continued)*

**Table 1. Study Descriptors and Mean Effect Sizes for 18 Studies That Met Inclusion Criteria Categorized by Methodological Level[a]** *(continued from previous page)*

| Study | Diagnostic Measure/ Method | Comparison Group | Group Ns (Schizophrenia/ Comparison) | Age at Testing | IQ Measures | Adjusted Effect Size[b] |
|---|---|---|---|---|---|---|
| **Level 3** | | | | | | |
| Albee et al. (27) | Diagnosis determined in a North Ohio hospital or clinic | Cleveland school children in a representative school year | 122/2,613 | Grade 2 | Kuhlmann-Anderson | –0.68 |
| | | | 154/4,166 | Grade 6 | Cleveland Classification Test | –0.64 |
| | | | 103/4,960 | Grade 8 | Terman-McNemar | –0.52 |
| Bower et al. (28) | Diagnosis determined by hospital staff | Next student in high school yearbook not known by state agencies to have mental illness | 44/44 | High school | Last group administered IQ test completed in high school | –0.45 |
| Lubin et al. (29) | Diagnosis determined by senior hospital staff | 1) Enlisted patients with no history of brain injury or mental illness (N=47); 2) enlisted employees of field hospital and troop command center (N=115) | 159/162 | Mean=26 years (SD=7.3), range=18–51 | Army Classification Battery | –0.10 |
| Offord (30) | Diagnosis determined during hospitalization (before 1971) | School system peers matched on race, sex, age, and social class | 116/116 | Within first 9 years of school | Average of school-based IQ assessments | –0.54 |
| Watt and Lubensky (31) | Diagnosed schizophrenia in majority of hospital admissions (1958–1964) | Classmates matched on sex, race, social class of origin, and migratory status | 50/50 | Kindergarten–grade 12 | Average of school-based IQ assessments | –0.40 |
| | | | 36/36 | Kindergarten–grade 6 | Kuhlmann-Anderson (grades 3 and 6); Otis Self-Administering Test (occasionally grade 6) | –0.34 |
| | | | 48/48 | Grades 7–12 | Otis Self-Administering Test (grade 8); ACE Psychological Test (grade 10 or 11) | –0.36 |
| Zammit et al. (32)[d] | Nordic ICD-8 or ICD-9/ hospital admission | Conscript cohort (1969–1970) without diagnosis of schizophrenia, bipolar disorder, severe depression, or other non-affective psychoses | 362/49,159 | Mean=18.3 years (SD=0.6), range=18–20 | Swedish conscript tests | –0.53 |

[a] Level 1: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "long" IQ estimates; Level 2: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "short" IQ estimates; Level 3: study that used older diagnostic systems (pre-DSM-III or RDC) and "short" IQ estimates.

[b] More than one effect size per study is included when data were provided separately by test or age of testing. Effect sizes were estimated using Cohen's d: $(M_1+M_2)/\sqrt{[(n_1-1)(SD_1)^2+(n_2-1)(SD_2)^2]/[n_1+n_2-2]}$, with the following qualifications: the Brewer et al. (16) effect size is from the original article. For Cannon et al. (17), since actual Ns were not available for the analysis of verbal versus nonverbal IQ, the df for these t tests, in conjunction with the Ns reported for the larger analysis, were used in the equation Cohen's $d=t(n_1+n_2)/[\sqrt{df}\sqrt{n_1*n_2}]$ to provide a more accurate effect size estimate. For Cannon et al. (22), effect size was estimated from graphs providing z scores (based on the entire sample) for the schizophrenia and comparison groups. For Jones et al. (23), Cohen's $d=\sqrt{[F(n1+n2)/(n1*n2)]}$. For Albee et al. (27), the effect size was estimated from the Cohen's equation for d using t and unequal Ns (see equation provided for Cannon et al. [17]); effect size differs from that reported in Aylward et al. (4) who calculated d using 2t/$\sqrt{df}$. For Bower et al. (28), the effect size was estimated from the critical ratio using the same equation provided for Cannon et al. [17] and Albee et al. [27]). Although Offord (30) reported data from matched pairs, the effect size was calculated using sample means and standard deviations per Dunlap et al. (33). For Watt and Lubensky (31), effect size estimated from matched pairs t test for the larger sample: $d=t/\sqrt{df}$; results differ from those reported in Aylward et al. (4) who appear to have calculated Cohen's d as if the t value was from independent samples in addition to using data from only the male and female subsamples with data from both time points.

[c] Although the comparison group in this study might not be considered "healthy" because all had been seen in a child psychiatric clinic, most other studies defined comparison groups by adult diagnostic status without full information on possible childhood psychopathology. This study was included because none of the comparison subjects was diagnosed with schizophrenia as an adult. The –1.78 effect size was excluded as an extreme outlier (z>3).

[d] Complete data not available in reported study but provided upon request.

[e] This study would have been excluded if it reported only on the age 3 time point because we excluded other studies in which the IQ estimate was based on a single test. We included this effect size in our analyses because it was one of five time points included in the study and the other four time points used extended IQ measures.

## Method

### Literature Search and Study Selection

Our literature search included both an online PubMed database search and a careful evaluation of the references from reviews and original studies pertaining to premorbid IQ in schizophrenia that were published before March 1, 2007. The keywords used in the computer search included all combinations of the following words: "IQ," "intelligence," "cognition," "neuropsychological," "neurocognitive," "schizophrenia or psychosis," and "premorbid or predictors." All potentially relevant studies were examined manually to assess inclusion and exclusion criteria. Inclusion criteria were the following: 1) published in English, 2) use of standardized psychometric IQ (or equivalent) tests, 3) testing conducted prior to the onset/diagnosis of schizophrenia, and 4) test results provided separately for a group consisting solely of subjects diagnosed with schizophrenia, schizoaffective disorder, or schizophreniform disorder and for a group of healthy comparison subjects. Exclusion criteria were as follows: 1) IQs estimated from achievement tests or single tests, such as word reading, or solely from verbal IQ or performance IQ assessment; 2) absence of premorbid IQ data from a relatively healthy comparison group or comparison data reported solely from a group at high risk for psychosis, with known cognitive delays in childhood or matched for childhood IQ; 3) insufficient data to closely estimate effect sizes (with each study required to have one of the following combinations: means, standard deviations, and number of subjects for each group; either a t or F score and the degrees of freedom or number of subjects for this statistic; or group differences expressed in standard deviation or z score units); and 4) data reported on the same or overlapping sample as a more complete or relevant study.

### Methodological Categorization

Although there are several important differences in the methodologies of studies on premorbid IQ in schizophrenia, we primarily focused on two issues we thought were likely to have significant impact on estimates of premorbid IQ deficits in schizophrenia patients. The first relates to diagnostic assessment criteria and methods. A significant change was made in the diagnostic criteria for schizophrenia from DSM-II to DSM-III (i.e., narrowing criteria for schizophrenia and altering the boundary with affective psychoses) and with the advent of the RDC. The only prior meta-analysis of premorbid IQ drew primarily from studies in which samples were diagnosed according to pre-DSM-III criteria or before RDC. Thus, we categorized the studies in our review based on whether or not diagnoses were made by criteria according to ICD-10, RDC, DSM-III, DSM-III-R, or DSM-IV.

The second important methodological issue considered was the type and number of tests used to estimate IQ. Individually-administered test batteries of both verbal and nonverbal subtests are now considered the gold standard for assessing IQ, with the Wechsler scales (Wechsler Adult Intelligence Scale [WAIS] [8], Wechsler Intelligence Scale for Children [WISC] [9]) being the most well established. However, given the extensive time needed to administer these test batteries, studies with large samples or many neuropsychological measures typically estimate IQ using two to four subtests. Although these estimates are known to be highly comparable with IQ scores based on full test batteries in standardization samples, it has been argued that short forms of IQ tests may attenuate reliability and validity, particularly in clinical or racially diverse samples (10–12). Similarly, group-administered IQ tests, typical of school settings, reportedly yield more variable estimates relative to a full, individually-administered IQ test battery (12). For this reason, we characterized each study based on the number of tests used to estimate IQ and whether tests were individually or group administered. Studies using at least one-half of an individu-

ally-administered IQ test battery were considered to have "long" IQ estimates, and studies using two to four subtests of an IQ test battery or group-administered IQ tests were considered to have "short" IQ estimates.

Using these two methodological considerations, we organized studies a priori into the following three levels: 1) level 1, studies that used recent diagnostic criteria (RDC, ICD-10, DSM-III, or DSM-IV) and "long" IQ estimates; 2) level 2, studies that used recent diagnostic criteria and "short" IQ estimates; and 3) level 3, studies that used older diagnostic systems (pre-DSM-III or pre-RDC) and "short" IQ estimates. We found no studies using older diagnostic systems and "long" IQ estimates.

Other (secondary) methodological issues that were considered in our analyses included age at premorbid IQ testing and sample ascertainment. In order to analyze effect sizes by age at testing, we categorized studies according to the following general age ranges during which testing was conducted: 1) exclusively during childhood (age <13); 2) exclusively during adolescence or early adulthood (age ≥13); and 3) across a range of ages, including both childhood and adolescence, or without specifying age. If IQ estimates declined with the onset of acute psychosis or during the "prodrome" (the period of active increase of subthreshold symptoms leading into psychosis), we expected effect sizes to be larger in samples of older individuals who were presumably closer to psychosis onset.

Similarly, the method of sample ascertainment might lead to discrepant estimates of premorbid impairment in schizophrenia. We categorized studies according to whether patient samples were identified through 1) diagnostic screening or hospital record linkage of large cohort or population samples; 2) follow-up of genetic risk samples; 3) follow-up or hospital record linkage of conscript samples; 4) selective hospital samples with follow-back assessment of school, conscript, or clinic records; and 5) follow-up of prodromal samples. Comparison samples were either 1) remaining members of cohort or population samples or 2) healthy individuals matched with a patient sample on various demographic variables.

### Statistical Analysis

Effect sizes for each study were calculated using Hedges' adjusted, standardized mean differences (13). We calculated both weighted and unweighted means. Weighted means were calculated using the inverse variance weight (14). A single mean effect size was calculated for studies with multiple effect sizes. For studies with independent samples (e.g., effect size reported by gender), the study effect size was the weighted mean. An unweighted mean effect size was calculated for studies that used repeated measures or other dependent-sample measures (e.g., effect size for verbal and nonverbal subtests). One effect size ([15]; for the comparison of WISC scores only) was an extreme outlier (z >3) and therefore excluded from all analyses.

We used the analogue to the analysis of variance method (13) to model between-study variance due to key study design factors, such as a priori methodological level, patient and comparison sample ascertainment methods, and age at testing. Heterogeneity was determined by a significant Q statistic based on chi square tables (14). Finally, we calculated a mean IQ for all studies that reported IQ scores by group for tests with known means and standard deviations.

## Results

In our literature search, we identified 69 studies that reported original standardized psychometric IQ data measured in subjects prior to the onset of schizophrenia, schizoaffective disorder, or schizophreniform disorder. Of

these, eight were excluded because premorbid IQ was estimated from achievement tests, from single tests (such as word reading), or solely from verbal IQ or performance IQ scores. Eighteen studies were excluded because data on healthy comparison subjects were not reported. Of the remaining 43 studies, 17 reported insufficient data for calculating a standardized mean effect size specific to the schizophrenia subgroup. Requests for additional data from authors of studies with appropriate designs but incomplete data for estimating premorbid IQ yielded data from four studies. Eight of the remaining 26 studies provided data on samples for which another published study reported more complete or up-to-date data. Table 1 details the basic study descriptors and mean effect sizes for the 18 studies included in our analyses (15–32). (For the list of excluded studies and the study descriptors [location and design], see the data supplement accompanying the online version of this article.)

### Mean Effect Size for Premorbid IQ

The mean weighted (and unweighted) effect size for studies included in our analyses was: Cohen's d=–0.54. This suggests a medium-sized deficit in global cognition prior to the onset of schizophrenia. On average, the premorbid IQ scores in schizophrenia samples had an estimated 33% nonoverlap with scores from comparison samples. Although five studies reported unreliable differences between schizophrenia and comparison samples, overall differences were statistically reliable (i.e., confidence intervals did not contain zero). In converting available IQ scores to a standardized score with a mean of 100 and standard deviation of 15, the mean premorbid IQ estimate for schizophrenia samples was 94.7 (0.35 standard deviations below the mean and at the lower end of the average range).

### Analysis of Heterogeneity

Effect sizes were moderately heterogeneous (Q=29.60, df=17, p<0.05 [Figure 1]), but no significant heterogeneity between methodological levels was found (Q=4.56, df=2, p>0.05). According to the general age at which premorbid testing was conducted, mean effect sizes were also not significantly heterogeneous (Q=0.11, df=1, p>0.05 [comparing groups with specific childhood or adolescent/adult test scores]; Q=0.16, df=2, p>0.05 [including the group for which premorbid testing was not age-specific]). Method of schizophrenia sample ascertainment accounted for almost no mean effect size heterogeneity (Q=0.27, df=4, p>0.05). The only methodological variable that accounted for significant effect size heterogeneity was the method of ascertainment of comparison subjects (Q=10.11, df=1, p<0.01). The mean effect size for studies that used healthy matched comparison subjects (Cohen's d=–0.36) was significantly smaller than the mean effect size for studies that used comparison subjects from large-population or cohort samples (Cohen's d=–0.56). However, this effect was

### FIGURE 1. Mean Effect Size by Study and Methodological Level[a]
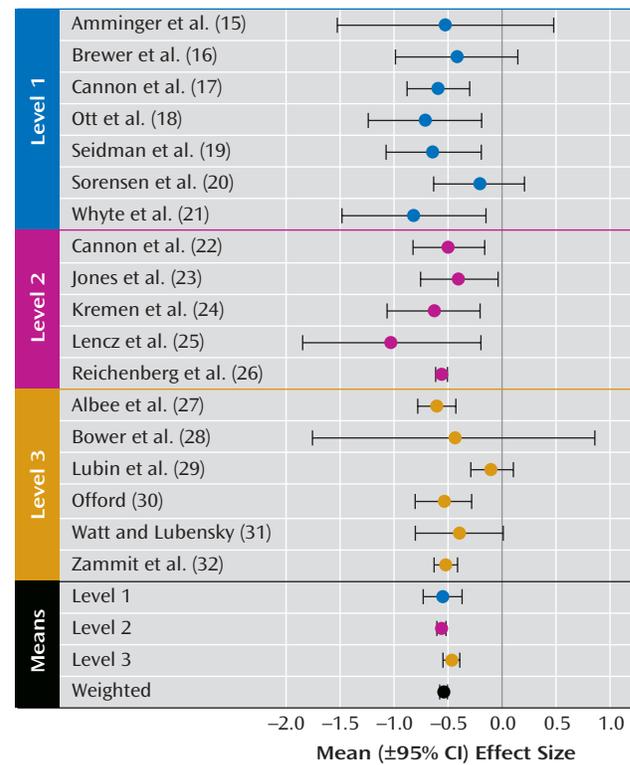


[a] Level 1: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "long" IQ estimates; Level 2: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "short" IQ estimates; Level 3: study that used older diagnostic systems (pre-DSM-III or RDC) and "short" IQ estimates.

primarily accounted for by the effect size of one study, which had a somewhat atypical comparison sample ascertainment method (Cohen's d=–0.10 [29]). When the effect size of this particular study was removed from the analysis, the effect sizes of the two different comparison sample ascertainment methods (method 1: identification from the remaining members of cohort or population samples; method 2: identification from healthy individuals matched with a patient sample on various demographic variables) were homogeneous (Q=1.25, df=1, p>0.05). In fact, when the effect size of this study was removed from the analysis of the overall effect size heterogeneity, the remaining effect sizes were homogeneous (Q=8.73, df=16, p>0.05).

### Verbal Versus Nonverbal IQ

Seven studies provided data for calculating a mean effect size by verbal and nonverbal domains (Figure 2). No significant difference in weighted mean effect size was found by verbal versus nonverbal domains. Interestingly, examination of comparison data suggested that, in at least three of these seven studies (15, 18, 20), the discrepancies that were found reflected discrepancies in comparison group data rather than discrepancies in schizophrenia group

**FIGURE 2. Effect Size for Verbal and Nonverbal IQ Domains**



**FIGURE 3. Mean Effect Size by Age and Methodological Level[a]**



[a] Level 1: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "long" IQ estimates; Level 2: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "short" IQ estimates; Level 3: study that used older diagnostic systems (pre-DSM-III or RDC) and "short" IQ estimates.
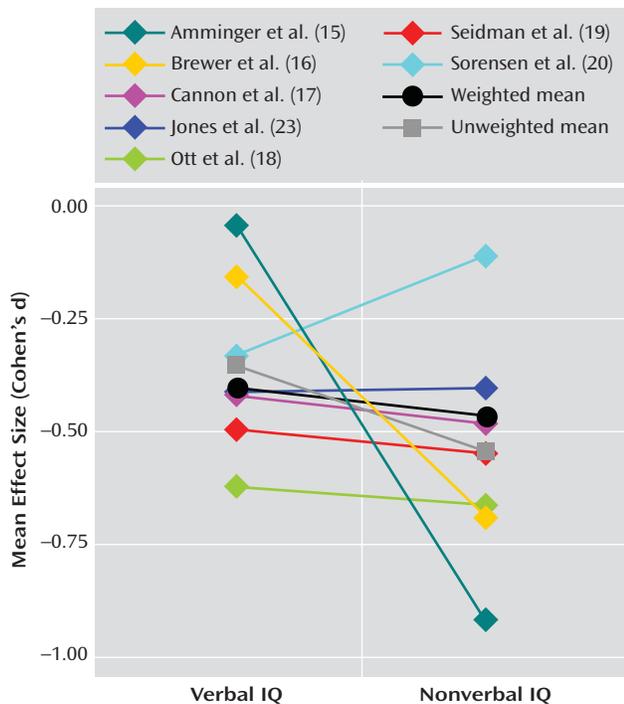
data. Three of the remaining four studies did not provide adequate comparison data to make this determination.

### Premorbid IQ by Gender

Only three studies provided data on premorbid IQ by gender (26, 30, 31). Although male subjects had a significantly lower mean premorbid IQ relative to female subjects in one study (30), the overall weighted mean effect size for premorbid IQ was not significantly different for male (Cohen's d=–0.53) and female (Cohen's d=–0.52) subjects.

### IQ Impairment Over Time

Given recent evidence suggesting that IQ may decline from the premorbid to post-onset stages of schizophrenia (e.g., 19, 34), we examined both cross-sectional and longitudinal data for evidence of changes in IQ impairment with age or over time prior to the onset of illness. For the cross-sectional analysis, three studies contributed effect sizes from testing conducted exclusively during childhood (17, 19, 22). Six studies contributed effect sizes from testing conducted during childhood and adolescence or at unspecified ages (18, 20, 23, 27, 30, 31), and nine studies contributed effect sizes from testing conducted exclusively during adolescence and early adulthood (15, 16, 21, 24–26, 28, 29, 32). As shown in Figure 3, we found no cross-sectional evidence of a decline in IQ with age during the premorbid period.

Sufficient data to conduct a longitudinal analysis of IQ over time were not available for studies with repeated measures of IQ. (Not all studies reported repeated-mea-

sures analyses or the correlations across repeated measures necessary for estimating change over time effects.) Findings from these studies are detailed in Table 2. Although the data reported suggest a possible increase in premorbid effect size over time in some samples (23, 31), none of the studies that reported analyses of IQ over time exclusively during the premorbid period (17, 18, 22, 23, 31, 35) reported a significant increase in IQ impairment in schizophrenia samples relative to comparison samples. In fact, the only study that reported a significant group-by-time interaction found a larger decrease in the vocabulary scores of comparison subjects relative to those subjects who later developed schizophrenia (18).
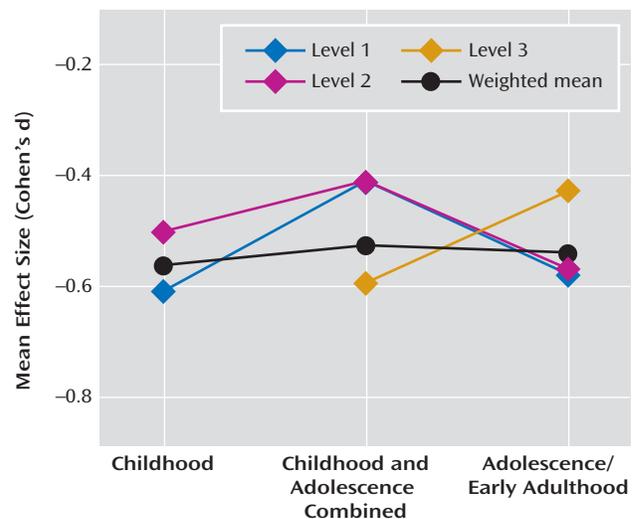
Three studies that used longitudinal designs reported both premorbid and post-onset estimates of IQ for the same sample. All three of these studies (one with data reported in two separate articles), found a significant decline in IQ from premorbid to post-onset testing in schizophrenia samples relative to comparison samples (19, 29, 34, 36).

## Discussion

To our knowledge, this is the first quantitative review of the literature on premorbid IQ in schizophrenia since the well-cited meta-analysis by Aylward et al. (4) was published in 1984. As such, it is the first to incorporate studies that used the most recent diagnostic criteria for schizophrenia and samples that reached the age of schizophrenia onset since the Aylward et al. (4) review. With the increase in both the number of available studies and diagnostic specificity, we were able to apply more rigorous in-

**TABLE 2. Studies Reporting Multiple Measures of IQ Over Time in a Given Sample**

| Level[a] and Study | Period of Change | Findings in Schizophrenia Relative to Comparison Group |
|---|---|---|
| **Level 1** | | |
| Cannon et al. (17) | Ages 4 to 7 | Non-significant change or interaction in schizophrenia sample. |
| Ott et al. (18) | Ages 7 and 9 (4- to 7-year period in subjects within this age range) | Significant interaction only for vocabulary subtest (comparison decline > schizophrenia decline). |
| Seidman et al. (19) | Ages 7 to 35 | Significant interaction for full-scale IQ estimated from vocabulary and block design subtests (schizophrenia decline > comparison decline). |
| **Level 2** | | |
| Cannon et al. (22) | Ages 3 to 11 | Values do not support overall decline in IQ; statistics unavailable. |
| Caspi et al. (34) | Age 16 or 17 to first episode (mean age=23) | Significant deterioration on two out of four draft board subtests in schizophrenia sample relative to comparison sample (subsample of sample analyzed in Reichenberg et al. [26]). |
| Jones et al. (23) | Ages 8, 11, 15 | Greater verbal and nonverbal effect sizes at age 15 relative to age 8; statistical significance unavailable. |
| **Level 3** | | |
| Lane and Albee (35) | Grades 2, 6, 8 | No relative decline reported for grades 2 to 6 (same sample reported in Albee et al. [27]). |
| Lubin et al. (29) | Induction to hospitalization | Significant decline in schizophrenia sample scores relative to increase in comparison sample scores (36). |
| Watt and Lubensky (31) | Kindergarten to grade 6; grades 7 to 12 | Non-significant relative decline in schizophrenia sample. |

[a] Level 1: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "long" IQ estimates; Level 2: study that used recent diagnostic criteria (RDC, ICD-10, or DSM-III or later) and "short" IQ estimates; Level 3: study that used older diagnostic systems (pre-DSM-III or RDC) and "short" IQ estimates.

clusion criteria. Analyzing only studies that used standardized measures of psychometric IQ administered during the premorbid period in samples of individuals who later developed schizophrenia, schizoaffective disorder, or schizophreniform disorder, we found a moderate and reliable effect size (Cohen's d=–0.54) for premorbid IQ impairment (relative to IQ estimates in comparison samples) that was comparable with the mean effect size reported by Aylward et al. ([Cohen's d=–0.43] As noted in Table 1, we used different methods than Aylward et al. Applying these methods to the data summarized in Table 2 of the review by Aylward et al. [4, p. 437], we obtained a mean Cohen's d of –0.49.). The effect size we found is almost exactly one-half of the mean effect size found in chronically ill samples (Cohen's d=–1.10 [37]). Furthermore, it is consistent across studies using more recent as well as older diagnostic criteria and across studies using longer and individually-administered and shorter or group-administered IQ test batteries (with one exception [29]).

Interestingly, in examining factors that might account for the unusually small premorbid IQ deficit found in the schizophrenia group in the Lubin et al. (29) study, we found that the mean IQ score of the comparison group (standard score equivalent=95.3) was slightly below the standard score mean for IQ (100). While this was not the lowest IQ score of a comparison group among the studies we examined (mean IQ=90.6 for comparison subjects at age 7 [17]), it may represent a more poorly matched comparison sample relative to other studies. More precisely, the study sample examined by Lubin et al. is the only one in which comparison subjects were recruited from specific subsamples of a larger cohort. An earlier published report on this same study sample indicated that nearly one-third of the comparison subjects were recruited from

the same hospital patient population as schizophrenia subjects. The other two-thirds were recruited from specific work sites, a field hospital, and a troop command (36). It is possible that this recruitment method yielded comparison subjects with a lower mean IQ score relative to the larger conscript cohort from which the schizophrenia subjects were identified.

Appropriateness of comparison group sampling is important to this type of analysis. Based on standardized psychometric IQ test scores (mean=100 [SD=15]), one would expect an effect size of Cohen's d=–0.54 to equate to a mean premorbid IQ score of 91.8 in schizophrenia samples. The mean premorbid IQ score of 94.7 in the studies we reviewed reflects a mean IQ in comparison groups that was slightly above standardization means of 100. In other words, the mean deficit of Cohen's d=–0.54 may be slightly inflated because of a potential confound in comparison sampling. However, effect sizes in large samples with comparison subjects, either well matched on several demographic variables or highly representative of the larger population (17, 23, 26, 30–32), were homogeneous in demonstrating a mean premorbid IQ deficit of approximately one-half the standard deviation. This suggests that comparison sampling was not a significant overall confound, even if comparison subjects who were poorly matched or had lower mean IQ scores contributed to atypical results at an individual study level (16, 25, 29).

The modest impairment of premorbid IQ in schizophrenia samples only underscores the persistent question of whether IQ declines with illness progression. While the discrepancy between pre- and postmorbid effect sizes is consistent with theories of a decline in IQ over the course of schizophrenia onset, it might also be explained by differences in sampling, medication, or clinical state. However, there is no obvious indication that the samples from

which premorbid IQ measures were obtained in our analyses were less chronic overall relative to those in the Heinrichs and Zakzanis meta-analysis of post-onset neurocognition (37). In addition, a significant correlation between neuroleptic dose and IQ deficit was not found in the Heinrichs and Zakzanis review. Moreover, while Heinrichs and Zakzanis did find a smaller mean effect size in studies that used non-WAIS-R estimates of full-scale IQ compared with WAIS-R estimates, most of the non-WAIS-R estimates were based on single tests or word reading tests, which were not included in our analyses. This evidence suggests that the discrepancy between pre- and postmorbid IQ in schizophrenia might be related to clinical state or illness progression instead of sampling, medication effects, or measurement artifact.

Although nine of the 18 studies provided multiple measures of IQ over time in a given sample (including two studies with reports on similar, although not exactly the same, samples), the data were insufficient to calculate a mean change in effect size over time. Only three longitudinal studies reported significantly greater impairment in IQ over time, and these changes in impairment were from pre- to postmorbid testing. Since these changes could have occurred entirely during the postmorbid period, there remain no reliable longitudinal data supporting a premorbid decline in IQ in schizophrenia. Prospective longitudinal data from pre-adolescence through the premorbid, prodromal, and illness stages are needed to determine if and when IQ declines over time for individuals who develop schizophrenia (relative to comparison subjects). However, as demonstrated in the study conducted by Lane and Albee (35), careful matching of comparison groups is critical to analyses of change over time. It is only through the careful matching of comparison groups that we can properly consider the significance of age at assessment, age at symptom and psychosis onset, and age and duration of illness at retesting.

Analyses of premorbid functioning in specific and separable neurocognitive domains over time may also be important in identifying patterns of change not evident in global measures of cognition such as IQ. Deficits in some domains, such as attention and verbal memory, may be apparent during the premorbid or early prodromal phases and represent markers of vulnerability or predictors of illness (38). Cognitive functioning in other domains, such as executive or olfactory functioning, may become increasingly impaired over time or with illness progression (34, 39). However, a focus on single neurocognitive functions may be less promising in indexing vulnerability and predicting outcome than combinations of relatively distinct neurocognitive and behavioral variables (e.g., 40). In any case, interpretations regarding the deterioration of IQ—or lack thereof—should not be generalized to other neurocognitive functions that are not sufficiently measured by IQ tests (e.g., executive functions, memory).

In summary, with the exception of one study that had potentially problematic comparison group sampling, studies that incorporated several different methodologies and IQ measures provided a notably consistent and reliable report of IQ deficit prior to the onset of schizophrenia. This medium-sized premorbid deficit 1) was approximately one-half that found after diagnosis; 2) could be reliably measured in childhood; and 3) did not appear to progress with age or over time (within the limited data available), even during the onset of early prodromal symptoms. In contrast to findings from several individual studies, evidence is lacking at the meta-analytic level to support the theory of a larger premorbid deficit in performance versus verbal IQ or in male versus female subjects. While the overall finding is highly consistent with theories of schizophrenia as a neurodevelopmental disorder, the size of premorbid relative to postmorbid IQ estimates supports the presence of additional progressive deterioration over the transition to acute psychosis. It is not at all clear whether and when this deterioration occurs or whether it is best accounted for by specific or widespread changes. These questions can only be answered with repeated measures across a number of cognitive domains during the premorbid, prodromal, and very early stages of psychosis onset as well as over time.

## References

1. Murray RM, Lewis SW: Is schizophrenia a neurodevelopmental disorder? Br Med J 1987; 295:681–682

2. Weinberger DR: Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry 1987; 44:660–669

3. Seidman LJ: The neuropsychology of schizophrenia: a neurodevelopmental and case study approach. J Neuropsychiatry Clin Neurosci 1990; 2:301–312

4. Aylward E, Walker E, Bettes B: Intelligence in schizophrenia: meta-analysis of the research. Schizophr Bull 1984; 10:430–459

5. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 3rd ed. Washington, DC, American Psychiatric Publishing, 1980

6. World Health Organization: The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva, WHO, 1992

7. Spitzer R, Endicott J, Robins E: Research diagnostic criteria. Arch Gen Psychiatry 1975; 35:773–782

8. Wechsler D: Manual for the Wechsler Adult Intelligence Scale, 3rd ed (WAIS-III). San Antonio, Tex, Psychological Corporation, 1997

9. Wechsler D: Manual for the Wechsler Intelligence Scale for Children, 3rd ed (WISC-III). San Antonio, Tex, Psychological Corporation, 1991

10. Silverstein A: Short forms of individual intelligence tests: psychological assessment. J Consult Clin Psychol 1990; 2:3–11

11. Kaufman J, Kaufman A: Time for the changing of the guard: a farewell to short forms of intelligence tests. J Psychoeducational Assessment 2001; 19:245–267

12. Sattler J: Assessment of Children: Cognitive Applications. San Diego, Jerome M Sattler, Publisher, Inc, 2001

13. Hedges LV: Distribution theory for Glass's estimator of effect size and related estimators. J Educ Stat 1981; 6:107–128

14. Lipsey MW, Wilson DB: Practical Meta-Analysis. Thousand Oaks, Calif, Sage Publications, 2001

15. Amminger GP, Schlögelhofer M, Lehner T, Looser Ott S, Friedrich MH, Aschauer HN: Premorbid performance IQ deficit in schizophrenia. Acta Psychiatr Scand 2000; 102:414–422

16. Brewer WJ, Francey SM, Wood SJ, Jackson HJ, Pantelis C, Phillips LJ, Yung AR, Anderson V, McGorry PD: Memory impairments identified in people at ultra-high risk for psychosis who later develop first-episode psychosis. Am J Psychiatry 2005; 162:71–78

17. Cannon TD, Bearden CE, Hollister JM, Rosso IM, Sanchez LE, Hadley T: Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. Schizophr Bull 2000; 26:379–393

18. Ott SL, Spinelli S, Rock D, Roberts S, Amminger GP, Erlenmeyer-Kimling L: The New York High-Risk Project: social and general intelligence in children at risk for schizophrenia. Schizophr Res 1998; 31:1–11

19. Seidman LJ, Buka SL, Goldstein JM, Tsuang MT: Intellectual decline in schizophrenia: evidence from a prospective birth cohort 28 year follow-up study. J Clin Exp Neuropsychol 2006; 28:225–242

20. Sorensen HJ, Mortensen EL, Parnas J, Mednick SA: Premorbid neurocognitive functioning in schizophrenia spectrum disorder. Schizophr Bull 2006; 32:578–583

21. Whyte M-C, Brett C, Harrison LK, Byrne M, Miller P, Lawrie S, Johnstone EC: Neuropsychological performance over time in people at high risk of developing schizophrenia and controls. Biol Psychiatry 2006; 59:730–739

22. Cannon M, Caspi A, Moffit TE, Harrington H, Taylor A, Murray RM, Poulton R: Evidence for early-childhood, pan-developmental impairment specific to schizophreniform disorder. Arch Gen Psychiatry 2002; 59:449–456

23. Jones PB, Rodgers B, Murray R, Marmot M: Child developmental risk factors for adult schizophrenia in the British 1946 birth cohort. Lancet 1994; 344:1398–1402

24. Kremen WS, Lyons MJ, Boake C, H. X, Jacobson KC, Waterman B, Eisen SA, Goldberg J, Faraone SV, Tsuang MT: A discordant twin study of premorbid cognitive ability in schizophrenia. J Clin Exp Neuropsychology 2006; 28:208–224

25. Lencz T, Smith CW, McLaughlin D, Auther A, Nakayama E, Hovey L, Cornblatt BA: Generalized and specific neurocognitive deficits in prodromal schizophrenia. Biol Psychiatry 2006; 59:863–871

26. Reichenberg A, Weiser M, Rapp MA, Rabinowitz J, Caspi A, Schmeidler J, Knobler HY, Lubin G, Nahon D, Harvey PD, Davidson M: Elaboration on premorbid intellectual performance in schizophrenia: premorbid intellectual decline and risk for schizophrenia. Arch Gen Psychiatry 2005; 62:1297–1304

27. Albee GW, Lane EA, Reuter JM: Childhood intelligence of future schizophrenics and neighborhood peers. J Psychol 1964; 58:141–144

28. Bower EM, Shelhammer TA, Daily JM: School characteristics of male adolescents who later became schizophrenic. Am J Orthopsychiatry 1960; 4:712–729

29. Lubin A, Gieseking CF, Williams HL: Direct measurement of cognitive deficit in schizophrenia. J Consult Psychol 1962; 26:139–143

30. Offord DR: School performance of adult schizophrenics, their siblings and age mates. Br J Psychiatry 1974; 125:12–19

31. Watt NF, Lubensky AW: Childhood roots of schizophrenia. J Consult Clin Psychol 1976; 44:363–375

32. Zammit S, Allebeck P, David AS, Dalman C, Hemmingsson T, Lundberg J, Lewis G: Longitudinal study of premorbid IQ score and risk of developing schizophrenia. Arch Gen Psychiatry 2004; 61:354–360

33. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ: Meta-analysis of experiments with matched groups or repeated measures designs. Psychol Methods 1996; 1:170–177

34. Caspi A, Reichenberg A, Weiser M, Rabinowitz J, Kaplan Z, Knobler H, Davidson-Sagi N, Davidson M: Cognitive performance in schizophrenia patients assessed before and following the first psychotic episode. Schizophr Res 2003; 65:87–94

35. Lane EA, Albee GW: On childhood intellectual decline of adult schizophrenics: a reassessment of an earlier study. J Abnorm Psychol 1968; 73:174–177

36. Williams HL, Lubin A, Gieseking CF: Direct measurement of cognitive deficit in brain-injured patients. J Consult Psychol 1959; 23:300–305

37. Heinrichs RW, Zakzanis KK: Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology 1998; 12:426–445

38. Cornblatt BA, Lencz T, Smith CW, Correll CU, Auther A, Nakayama E: The schizophrenia prodrome revisited: a neurodevelopmental perspective. Schizophr Bull 2003; 29:633–651

39. Brewer WJ, Wood SJ, McGorry PD, Francey SM, Phillips LJ, Yung AR, Anderson V, Copolov DL, Singh B, Velakoulis D, Pantelis C: Impairment of olfactory identification ability in individuals at ultra-high risk for psychosis who later develop schizophrenia. Am J Psychiatry 2003; 160:1790–1794

40. Cornblatt BA, Obuchowski M, Roberts SA, Pollack S, Erlenmeyer-Kimling L: Cognitive and behavioral precursors of schizophrenia. Dev Psychopathol 1999; 11:487–508