Cognitive Decline in Schizophrenia from Childhood to Midlife: A 33-Year Longitudinal Birth Cohort Study

William S. Kremen, Ph.D.1, Sophia Vinogradov, M.D.2, John H. Poole, Ph.D.3, Catherine A. Schaefer, Ph.D.4, Raymond F. Deicken, M.D.2, Pam Factor-Litvak, Ph.D.5, and Alan S. Brown, M.D., M.P.H.6

1Department of Psychiatry, Center for Behavioral Genomics, University of California, San Diego, and VA San Diego Healthcare System, La Jolla, CA
2Department of Psychiatry, University of California, San Francisco; San Francisco Department of Veterans Affairs Medical Center, San Francisco, CA
3Veterans Affairs Healthcare System, Palo Alto, CA
4Kaiser Permanente Division of Research, Oakland, CA
5Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY
6Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York State Psychiatric Institute, and Mailman School of Public Health, New York, NY

Abstract

Background—To examine cognitive deficits before and after onset of schizophrenia in a longitudinal study that: 1) covers a long time interval; 2) minimizes test unreliability by including the identical measure at both childhood and post-onset cognitive assessments; and 3) minimizes bias by utilizing a population-based sample in which participants were selected neither for signs of illness in childhood nor for being at risk for schizophrenia.

Methods—Participants in the present study, Developmental Insult and Brain Anomaly in Schizophrenia (DIBS), were ascertained from an earlier epidemiologic study conducted in Oakland, CA. The original version of the Peabody Picture Vocabulary Test (PPVT), a test of receptive vocabulary, was administered at age 5 or 9 and repeated as part of the DIBS study at an average age of 40. There were 10 DIBS cases with DSM-IV schizophrenia or schizoaffective disorder and 15 demographically similar DIBS controls with both child and adult PPVT scores.

© 2009 Elsevier B.V. All rights reserved.

Correspondence to: William S. Kremen, Ph.D., Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive (MC 0738), La Jolla, CA 92039-0738.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Portions of this research were presented at the annual meeting of the International Congress on Schizophrenia Research, Colorado Springs, CO, March 28-April 1, 2007.

Contributors: Alan Brown, William Kremen, Sophia Vinogradov, John Poole, Catherine A. Schaefer, and Raymond Deicken all contributed to designing the study and writing the protocol. William Kremen and Alan Brown managed the literature searches and analyses. William Kremen and Anna Kochetkova undertook the statistical analysis. William Kremen wrote the first draft of the manuscript. Pam Factor-Litvak contributed to the interpretation of the data. All authors contributed to and have approved the final manuscript.

Conflict of Interest: All authors declare that they have no conflicts of interest.
**Results**—Cases scored significantly lower than controls in childhood (d=.95) and adulthood (d=1.67). Residualized scores indicating the number of SDs above or below one’s predicted adult score revealed a mean case-control difference of −1.51 SDs, consistent with significant relative decline over time among the cases (p<.0013).

**Conclusions**—In this prospective study, individuals who developed adult schizophrenia manifested impaired receptive vocabulary during childhood and further relative deterioration (or lack of expected improvement) between childhood and midlife. Limitations should also be acknowledged, including the small sample size and the fact that we cannot be certain when the continued deterioration took place.

**Keywords**
premorbid ability; deterioration; verbal ability; IQ

---

**1. Introduction**

It is widely accepted that cognitive deficits are core features of schizophrenia and that they are present before the onset of frank psychosis in individuals who later go on to develop the illness. There is also evidence of a decline in cognitive function after illness onset, although it appears that this usually occurs early in the course of illness. Beyond these early declines, the data largely suggest that any further changes are not greater than what would be expected from normal age-related declines (Heaton et al., 2001). However, there may be 2 exceptions to this general finding: treatment-resistant chronic patients may show late-life cognitive declines (Harvey et al., 1999), and cross-sectional data suggest relatively selective declines in executive function that exceed normal age-related changes (Fucetola et al., 2000).

Despite this growing body of pre- and post-onset findings, it is still rare to have cognitive measures in the same individuals both before and after the onset of schizophrenia. It is even more unusual to have the identical measures obtained at both time points. Particularly for very long-term follow-ups in which practice effects are not an issue, the state-of-the-art in longitudinal studies is to use the same—rather than alternate-form—measures in order to minimize the impact of test unreliability and method variance. Moreover, because of the issue of developmental changes in test performance, it is important to have a well-characterized healthy comparison group. For example, many cognitive functions show small improvements through late midlife (Schaie, 2005).

Participant in 2 early samples studied prior to the well-defined DSM-III diagnostic criteria were tested initially in their mid-20s. Schizophrenia patients exhibited small IQ declines, whereas healthy comparison participants exhibited increases averaging about the equivalent of 10 IQ points (Heaton and Drexler, 1987). Seidman et al. (2006) compared schizophrenia patients and controls who were tested at age 7 and again at an average post-onset age of 35, for a 28-year follow-up. They compared childhood scores on the Wechsler Intelligence Scale for Children (Wechsler, 1949) and adult scores on the Wechsler Adult Intelligence Scale – Revised (Wechsler, 1981) vocabulary (verbal ability), block design (perceptual organization), and digit span (short-term memory) subtests. Effect sizes at age 7 were 0.42, 0.18, and 0.48. The effect sizes at age 35 were larger: 0.87, 0.89, 0.72. These results indicated poorer cognitive functioning in schizophrenia patients in both childhood and adulthood. Although there was not a test for a group by time interaction, the effect size increases strongly suggest differential change over time. Controls declined very slightly but the patients declined even more.
In the Israeli Draft Board Study, participants were tested at age 16–17 and retested approximately seven years later at an average age of 23 (Caspi et al., 2003). Testing included a modified Raven’s Progressive Matrices (abstract reasoning), Otis-R (mental speed and comprehension), and modifications of the Wechsler similarities (abstract reasoning) and arithmetic (attention-concentration) subtests (Gal, 1986). Effect sizes at initial testing were .61, .88, 1.01, and 1.06, respectively. After onset of illness, the effects were 1.16, 1.72, 1.01, and 1.43. Individuals who developed schizophrenia did not manifest significant declines over this 6–7 year interval, but when compared with the change exhibited by controls, there were significant group by time interactions for the Raven’s Matrices and the Otis test.

Taken together, these results consistently indicate existing premorbid cognitive deficits in schizophrenia, evident both in early childhood and in adolescence, with further relative decline that is evident on post-onset testing either early in the course of illness or at midlife. The largest differential change from childhood to midlife was in perceptual organization, whereas increases in effect size was similar for basic verbal ability and short-term memory (Seidman et al., 2006). With respect to adolescent-to-young adult change, the less verbal of the 2 abstraction measures (Ravens’ Matrices) showed significant differential change (i.e., greater decline among patients than among controls) (Caspi et al., 2003). Attention-concentration did not show significant differential change, although the measure of mental speed and comprehension did. Overall, there is not a clear pattern of greater decline in one cognitive function over another.

In the present study, we compared individuals who developed schizophrenia with matched healthy comparison participants, all of whom were in the Developmental Insult and Brain Anomalies in Schizophrenia (DIBS) study, which was based on a birth cohort that was representative of the Oakland, CA area. Because scores on a cognitive test—the Peabody Picture Vocabulary Test (Dunn, 1959)—were available from childhood, we were able to conduct a longitudinal investigation over approximately 33 years. We examined whether individuals who developed schizophrenia were impaired relative to healthy comparison participants during childhood and during middle age, and whether there was differential change over time on this measure between these 2 groups.

2. Materials and methods

2.1 Recruitment and clinical assessment

The DIBS study was a follow-up investigation that included schizophrenia cases and healthy comparison participants ascertained from the earlier Prenatal Determinants of Schizophrenia (PDS) study. The PDS has been described in detail elsewhere (Susser et al., 2000). Briefly, it originated from the collaborative efforts of the Child Health and Development Study in Berkeley, CA and the Kaiser Permanente Medical Care Plan (Kaiser-Permanente) in Oakland, CA. The CHDS recruited nearly every pregnant woman who received obstetric care from Kaiser-Permanente in Alameda County, California. Information obtained from the CHDS on pre-, peri-, and postnatal exposures—including measures of child neurocognitive function, and parental background—was part of the PDS database. The cohort for the PDS study consisted of the subsample of 12,094 live births who were Kaiser-Permanente members from January 1, 1981 through December 31, 1997. There were no differences on a wide range of maternal and paternal characteristics between participants who remained in Kaiser-Permanente and those who were lost to follow-up (Bresnahan et al., 2000; Susser et al., 2000). Of those who did leave, the vast majority did so before the age of 10 (Susser et al., 2000); hence, early manifestations of schizophrenia probably did not influence the likelihood of loss to follow-up.
Screening based on ICD-9 diagnoses or pharmacy records of antipsychotic medication use resulted in 170 living potential cases. Of these, 146 (86%) were contacted and 107 were administered the Diagnostic Interview for Genetic Studies (Nurnberger et al., 1994), followed by a consensus diagnostic procedure. Diagnoses based on chart reviews by experienced clinicians were made for the 76 potential cases who were not interviewed. These procedures yielded 71 cases of DSM-IV schizophrenia or other schizophrenia spectrum disorders that included schizophrenia, schizoaffective disorder, delusional disorder, schizotypal personality disorder, and psychosis not otherwise specified; 83% of the cases had either schizophrenia or schizoaffective disorder.

Up to 8 controls for each case were matched on 5 characteristics: Kaiser-Permanente membership at the time of case ascertainment (time of first treatment for schizophrenia spectrum disorder); date of birth (± 28 days); sex; number of maternal blood samples drawn during the index pregnancy; and number of weeks after the last menstrual period of the first maternal blood draw during the index pregnancy (± 4 weeks). Matching for Kaiser-Permanente membership ensured that the controls for each case were representative of the population at risk at the time of case ascertainment. Further details on control selection in the PDS study can be found in Susser et al. (2000).

All PDS cases and one matched control per case were targeted for recruitment into the DIBS study. We enrolled 26 cases (13 with schizophrenia, 7 with schizoaffective disorder, and 6 with other schizophrenia spectrum disorders) and 25 controls. Written informed consent was obtained after description of the study. The protocol was approved by the Institutional Review Boards of the New York State Psychiatric Institute, the Kaiser Foundation Research Institute, and the University of California San Francisco VA Medical Center. Participants were paid for their involvement.

2.2 Demographic and cognitive characteristics

The focus of this report is on the Peabody Picture Vocabulary Test (PPVT) which was administered to participants during childhood as part of the CHDS birth cohort investigation and in adulthood as part of the DIBS study. There were 10 cases and 15 controls who had both childhood and adult PPVTs. Table 1 shows the sample demographic characteristics. Cases and controls were well matched on age, parental education, and sex distribution. As would be expected, the controls had significantly more education than the cases (about 2 years).

2.3 PPVT

The PPVT is a widely used receptive vocabulary measure. Participants are shown a series of sheets, each of which has 4 pictures. The tester says a word aloud and the participant must point to the picture that depicts the word. The test does not require reading, writing, or even verbal responses. The PPVT manual indicates test-retest reliabilities greater than 0.90. An unusual feature of the PPVT is that the same tests can be administered to individuals ages 2.5 and older. This feature made it possible to obtain meaningful scores from our adult sample on the identical version of the test that they took in childhood. Participants were administered the original (1959) version of the PPVT in childhood, and we administered the original version in the DIBS study in order to reduce method variance and test-retest unreliability. Moreover, we did not have to be concerned with practice effects given the length of the test-retest interval. Scores on the test are scaled according to age so that the scores are comparable across different ages at testing. Adult scores are based on a mean of 100 and SD of 15. Child scores were based on T-scores (mean=50, SD=10); we converted these to equivalent scores based on a mean of 100 and SD of 15 so that all scores being analyzed were comparable.
Prior PPVT scores were available at ages 5, 9–11, and 15–17. If a participant was administered the test more than once, only the score at first testing was included. There were very few available age 15–17 scores, and all of them were repeat tests. We therefore included scores from either age 5 or 9–11, depending on availability of the data. There were 6 cases and 8 controls with age 5 scores, and 4 cases and 7 controls with age 9 scores.

2.4 Statistical analysis

We tested for PPVT differences in childhood and in adulthood. In addition, we addressed the issue of change over time with 2 different data analytic approaches. First, we tested for group (case, control) by time (childhood, adulthood) interactions. Next, we performed regression analyses in which observed (actual) adult scores were compared to predicted adult scores. The prediction equation was based on control participants only, and standardized residual (observed minus predicted) scores were then calculated for all participants based on that prediction equation. In this approach, the mean residual score for controls is set at zero. Positive z-scores indicate observed scores that were above one's predicted score, and negative z-scores indicate observed scores that were below one’s predicted score.

3. Results

As can be seen in Table 1, cases had significantly lower PPVT scores than controls during both childhood ($p<0.04$) and adulthood ($p<0.0005$). Although the $p$-value of 0.0005 in adulthood suggests a more highly significant difference than the $p$-value of .04 in childhood, there was not a significant group × time interaction ($F[1,23]=1.91$, $p=0.18$). The lack of statistical significance suggests that our small sample size left us underpowered to detect the interaction. Comparison of the effect sizes, based on Cohen’s (1988) $d$ supports this inference because there was a large increase in the effect size for the case-control differences from childhood (d=0.95) to adulthood (d=1.67).

The regression analysis based on the residual scores lent further strong support to the idea of differential change between childhood and adulthood. Residual scores were highly significantly lower (more negative) in cases than in controls ($-1.51 [1.06]$ vs. $0.00 [0.96]$; $t=-3.67$, $p<0.0013$). The effect size for this comparison was very large (d=1.51). This difference indicates that the adult PPVT scores for the cases were 1.5 SDs below the scores that were predicted on the basis of their childhood scores. This result suggests a relative decline in performance from childhood to adulthood gauged against what would be considered normal change.

4. Discussion

Taken together, these results demonstrate that receptive vocabulary deficits were already present in childhood, and that there was a relative decline in receptive vocabulary between childhood and early midlife in individuals with schizophrenia compared with matched controls. The cases had an average decline of only 1.5 points, but the controls increased by approximately 1.4 points. As such, it is more accurate to say that the cases did not manifest what appears to be a small normally expected increase. An expected increase in PPVT performance is consistent with the results of longitudinal cognitive aging studies that show small increases in several cognitive abilities, including vocabulary, up through at least late midlife (Schaie, 2005). Our data, the results of normative cognitive aging studies, and those of clinical trials of schizophrenia (Goldberg et al., 2007; Gur et al., 2003) also highlight the fact that a healthy comparison group is essential in the study of cognitive change in schizophrenia. Without a healthy comparison group, we would have incorrectly inferred that the lack of change among cases in the present study indicated that there was nothing
abnormal. This point may seem so obvious as to not be worth mentioning, but when cognitive change is examined in pharmacological studies, for example, inclusion of a healthy comparison group is not that common (exceptions are Goldberg et al., 2007; Gur et al., 2003).

The results of the present study are generally consistent with previous studies of other cognitive tests. In an uncontrolled study, Russell et al. (1997) found no change in IQ from childhood to post-morbid adulthood in schizophrenia patients. Caspi et al. (2003) did not find significant changes on cognitive measures administered premorbidly (age 16–17) and readministered post-morbidly in early adulthood, but they did find significant relative deterioration on 2 out of 4 tests in comparison with normal controls. Seidman et al. (2006) found slight declines in the control group from childhood to early midlife, but greater declines in the schizophrenia group. Thus, consistent with our findings, the data for individuals with schizophrenia did not indicate significant change from childhood to adulthood in 2 of these 3 studies.

It is also interesting that individuals in the present study who developed schizophrenia manifested declines over time relative to normal functioning (actually lack of increase) even for a basic verbal ability measure that would be expected to be among those that are most well preserved. This result is consistent with what was found for vocabulary in the study of Seidman et al. (2006).

On the other hand, the average effect size in a meta-analysis of premorbid IQ in schizophrenia was approximately .5 SD with little variability across studies (Woodberry et al., 2008). A similar effect size has been reported for a receptive vocabulary test (Cannon et al., 2002). In their meta-analysis, Heinrichs and Zackzanis (1998) found that the average postmorbid effect size for WAIS-R IQ was 1.24 SDs and .63 for non-WAIS-R-IQ. As noted by Woodberry et al., the difference in effect size between the meta-analysis of premorbid versus postmorbid IQs may suggest meaningful decline over time, but it could be due to medication effects, ascertainment differences, or measurement artifacts. Medication effects seem an unlikely explanation because antipsychotic medications have not been found to have dramatic negative effects on IQ (Mishara and Goldberg, 2004). Ascertainment differences are also unlikely in the DIBS because it was a longitudinal study and cases were not inpatients; thus, they do not represent the subset of cases who would be most likely to experience the greatest declines. With regard to measurement issues, the difference in WAIS-R and other IQ measures highlights the value of having the same cognitive test at both time points. These considerations suggest that the conclusion to be drawn from our study is consistent with that of previous studies, despite the fact that the effect size was larger.

4.1 Strength and limitations

Limitations of the study include the small sample size, inclusion of only a single cognitive measure, and uncertainty as to whether the declines in performance (or lack of expected improvement) took place before or after onset of illness. The large effect sizes we observed might suggest something atypical about our sample, but it does not appear that the large effect sizes in the present study are due to our having either a very impaired schizophrenia group or a superior control group. For example, the mean education and PPVT scores are not atypical for schizophrenia-control comparisons. The PPVT is generally not considered to be appropriate as an IQ estimate (Lezak et al., 2004) but it is a measure of highly crystallized knowledge, and that type of measure would not be expected to yield larger effect sizes than IQ or other cognitive measures. Although the group means were not atypical, the variability within groups was somewhat low. As such, the large effect sizes may be a function of unstable parameter estimates due to the small sample size. Given the
consistency of the pattern of results with that of previous studies, we are confident that the results reflect a real effect. However, caution should be exercised regarding any strong interpretation of the magnitude of the effect. Thus, there is still a need for a larger study with all of the design features that were emphasized in the present study.

Strengths of this study were the combination of 3 features: 1) a very long (33-year) childhood to early midlife follow-up interval; 2) minimization of potential bias by utilizing a population-based birth cohort that was entirely unselected for risk for schizophrenia; and 3) minimization of test unreliability through the use of the identical test in childhood and adulthood. To our knowledge, no previous study has included all of these 3 features simultaneously.

Acknowledgments

We wish to acknowledge Ezra Susser, Barbara Cohn, Michaeline Bresnahan, Justin Penner, and Anna Kochetkova for their contributions to this work.

Role of the Funding Source: Funding for this study was provided by NIMH Grant 1R01MH-60249 (A.S.B), NIMH Grant 1 K02-MH65422 (A.S.B.), NICHD Grant N01-HD-1-3334 (B.A. Cohn), and NICHD Grant N01-HD-6-3258 (B.A. Cohn). The NIMH and NICHD had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

References


Schizophr Res. Author manuscript; available in PMC 2011 October 3.


TABLE 1
Demographic and Cognitive Characteristics Among Participants with Schizophrenia and Healthy Controls

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Schizophrenia (N=10)</th>
<th>Control (N=15)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Age (years)</td>
<td>39.0</td>
<td>1.8</td>
<td>40.3</td>
<td>1.8</td>
</tr>
<tr>
<td>Education (years)</td>
<td>13.1</td>
<td>1.7</td>
<td>15.3</td>
<td>1.4</td>
</tr>
<tr>
<td>Child PPVT</td>
<td>90.6</td>
<td>13.5</td>
<td>100.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Adult PPVT</td>
<td>89.1</td>
<td>10.9</td>
<td>104.9</td>
<td>8.6</td>
</tr>
<tr>
<td>K-W (\chi^2)</td>
<td>p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal education(c)</td>
<td>2.6</td>
<td>1.1</td>
<td>2.3</td>
<td>0.8</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>8 men, 2 women</td>
<td>11 men, 4 women</td>
<td>0.15</td>
<td>0.70</td>
</tr>
</tbody>
</table>

\(a\) PPVT = Peabody Picture Vocabulary Test

\(b\) K-W \(\chi^2\) = Kruskal-Wallis chi-square

\(c\) Maternal education: 1=less than high school; 2=high school; 3=some college; 4=college graduate. Median maternal education=2 for both groups. Result based on Wilcoxon test.