Neurocognitive development in first episode psychosis 5 years follow-up: Associations between illness severity and cognitive course

Helene Eidsmo Barde, Kjetil Sundet, Bjørn Rishovd Runde, Julie Evensen, Ulrik Haahr, Wenche ten Velden Hegelstad, Inge Joah, Jan Olav Johannessen, Hans Langeveld, T.K. Larsen, Ingrid Melle, Stein Opjordsmoen, Jan Ivar Røssberg, Erik Simonsen, Per Vaglum, Thomas McGlashan, Svein Friis

Article history:
Received 8 June 2012
Received in revised form 24 May 2013
Accepted 4 June 2013
Available online 1 July 2013

Keywords:
Neurocognition
First-episode psychosis
Longitudinal course
Relapse
Diagnosis

Abstract
Cognitive deficits are documented in first-episode psychosis (FEP), but the continuing course is not fully understood. The present study examines the longitudinal development of neurocognitive function in a five year follow-up of FEP-patients, focusing on the relation to illness severity, as measured by relapses and diagnostic subgroups. The study is an extension of previous findings from the TIPS-project, reporting stability over the first two years. Sixty-two FEP patients (53% male, age 28 ± 9 years) were neuropsychologically examined at baseline and at 1, 2, and 5 year follow-ups. The test battery was divided into five indices: Verbal Learning, Executive Function, Impulsivity, Motor Speed, and Working Memory. To investigate the effect of illness severity, the sample was divided in groups based on number of relapses, and diagnostic subgroups, respectively. Impulsivity and Working Memory improved significantly in the first two years, followed by no change over the next three years. Motor Speed decreased significantly from 2 to 5 years. Number of relapses was significantly related to Verbal Learning and Working Memory, showing a small decrease and less improvement, respectively, in patients with two or more episodes. No significant association was found with diagnostic group. Neurocognitive stability as well as change was found in a sample of FEP-patients examined repeatedly over 5 years. Of potential greater importance for understanding how psychotic illnesses progress, is the finding of significant associations between neurocognition and number of relapses but not diagnostic group, indicating that neurocognition is more related to recurring psychotic episodes than to the descriptive diagnosis per se.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction
Cognitive dysfunction is described as a core symptom of schizophrenia (Townsend and Norman, 2004; Mesholam-Gately et al., 2009; Palmer et al., 2009). It is present before onset of the first psychotic episode (FEP) (Pukrop et al., 2006; Jahshan et al., 2010; Frommann et al., 2011), but the continuing neurocognitive development is less clear (Kurtz, 2005; Seidman et al., 2006; Bonner-Jackson et al., 2010). Longitudinal studies of FEP are still rare (Milev et al., 2005), and follow-up intervals of one to five years appear to characterize most studies of FEP (Townsend and Norman, 2004; Albus et al., 2006; Rodriguez-Sanchez et al., 2008; Leeson et al., 2009; Bozikas and Andreou, 2011), reporting relatively static neurocognitive dysfunction in the first years after onset. A few studies report ten to thirteen year follow-ups (Stirling et al., 2003; Hoff et al., 2005; Oie et al., 2010), but provide conflicting results. Hoff et al.
(2005) found patterns of stability after the acute phase, whereas Stirling et al. (2003) found significant deterioration in visuospatial tasks, and Oie et al. (2010) reported deterioration on memory and processing speed tasks.

Discrepancies in findings generally are attributed to methodological limitations such as medication effects (Keefe et al., 2006), practice effects (Goldberg et al., 2007; Szoke et al., 2008) and different ways of obtaining cognitive data, e.g. follow-back studies deducing cognitive development from standardized achievement test scores (Bilder et al., 2006). Changes in the clinical course of schizophrenia spectrum disorders (SSDs) may also correlate with neurocognitive change, possibly resulting in different trajectories across clinical and/or diagnostic groups (Dominguez Mde et al., 2009). Although studies report no neurocognitive differences between various SSDs at initial presentation (Addington et al., 2003), subgroups may complicate and restrict interpretations in a long-term perspective.

The early course of psychosis is characterized by recurrent relapses. Studies have shown a pooled prevalence of relapse of positive symptoms in FEP to be 54% (40–63%) at 3 year follow-up (Alvarez-Jimenez et al., 2012), and up to 80% of patients will experience a relapse within 5 years of remission from the initial episode (Wiersma et al., 1998). Investigating clinically or diagnostically defined subgroups that would benefit from treatment adapted to potential neurocognitive difficulties seems relevant. Some studies have described an association between reduction in positive symptoms and cognitive improvement (Hoff et al., 1999), but investigations of relapse as a specific factor in long-term neurocognitive course is to our knowledge not previously reported.

The primary pathophysiological process in psychotic disorders has been argued to be mechanisms leading up to psychosis, not the psychosis per se (Becker et al., 2010). The exact nature of this mechanism remains unclear, but both lower values of a brain-derived neurotropic growth factor (Buckley et al., 2007) and volumetric changes in certain cortical regions are reported during the transition to psychosis (Pantelis et al., 2003). Although subtle, these changes reflect microscopic alterations that could cause the dysfunction of the brain–behavior relation (Caviness et al., 1999). Cognitive functioning may decline minimally with each psychotic episode (Becker et al., 2010), and studies with short follow-up periods probably do not allow for all progressive changes to be observed (Mane et al., 2009).

In a previous report from the TIPS-group, Rund et al. (2007) reported stability of neurocognitive deficits in a sample of FEP over the first two years after initiation of treatment. They also found an association between more relapses in the first year and greater severity of deficits in Verbal Learning and Working Memory (Rund et al., 2007). In the present study we expand the findings of Rund et al. (2007) to the 5 year follow-up, and investigate whether the association between neurocognitive domains and illness severity observed at the 2 year follow-up is evident, or possibly stronger, after five years of illness duration. More specifically, we ask if the five neurocognitive domains previously found in the TIPS-sample (Friis et al., 2002) will remain stable over a 5 year course from baseline assessment, and if there is a differential relationship between any of the indices and subgroups defined by either diagnosis or number of relapses over the first five years. The present study is among the very few studies that incorporate illness severity into analysis of neurocognition over time.

2. Material and methods

2.1. The TIPS project

The present report originates from the Early Treatment and Intervention in Psychosis Study (TIPS), a prospective longitudinal study of the relationship between duration of untreated psychosis (DUP) and outcome in FEP. The study was carried out in four Scandinavian health care sectors; Oslo, Stavanger, Haugesund and Roskilde.

A total of 301 patients were included in the TIPS study. The patients were 15–65 years of age, met the DSM-IV criteria for non-organic psychosis, and were actively psychotic without previously receiving adequate antipsychotic treatment. All patients were included in a defined treatment program (Melle et al., 2004). Symptom ratings were obtained at the start of treatment, at 3 months, and at 1-year follow up. Patients were tested neuropsychologically for the first time after remission of the psychotic symptoms (defined as a score lower than 4 on the PANSS (Kay et al., 1987) psychosis items (defined next paragraph) or following the 3-month follow-up). They were reassessed with the same test battery 1, 2 and 5 years after baseline testing.

2.2. Measures

2.2.1. Clinical instruments

The structured clinical interview for the DSM-IV (SCID) (First et al., 1995) was used for diagnostic purposes. Trained clinical research personnel carried out diagnostic evaluations. Symptom levels were assessed with the Positive and Negative Syndrome Scale, PANSS (Kay et al., 1987) and global functioning with the Global Assessment of Functioning Scale — split version (GAF).

DUP was measured as the time from the first onset of psychotic symptoms (defined as the first week with a PANSS score of 4 or more on one of the Positive scale items 1, 3, 5, 6 or General scale item 9) to the start of first adequate treatment of psychosis (defined as start of adequate antipsychotic medication or admission to hospital for treatment of acute psychosis).

Relapse was defined as the reappearance of positive psychotic symptoms (as defined above) for at least 7 days. Data on relapses were obtained from interview with the patient, and later confirmed by hospital records and/or discharge reports from the hospital.

Premorbid functioning was measured using the Premorbid Adjustment Scale (PAS) (Cannon-Spoor et al., 1982). A previous analysis identified two premorbid dimensions: social consisting of PAS item social isolation and peer relationships and academic which comprise school performance and school adaptation (Larsen et al., 2004).

Alcohol and drug use was measured with the Clinician Alcohol/Drug Use Scale (Drake et al., 1990), which assesses abuse over the last 6 months.

Satisfactory inter-rater reliability was found with overall agreement for DSM-IV diagnostic categories at baseline, Kappa: 0.76. PANSS: ICC (1, 1): 0.88 for positive symptoms, 0.76 for negative symptoms, and 0.53 for general symptoms.

2.2.2. Neurocognitive measures

The median time between start of treatment and first neurocognitive testing was 88 days (range: 46–234).

The subtests Similarities, Block Design, and Digit Span from WAIS-R (Wechsler, 1981) were used to calculate an IQ-estimate at baseline.

The neurocognitive test battery was found to validly assess five separate domains, as identified in a factor analysis of baseline data (Friis et al., 2002). The domain scores were calculated as the mean z-score of the tests included based on means and standard deviations at baseline. The five neurocognitive indices with the corresponding subtests and raw scores at each time point are presented in Table 1.

2.3. Subjects

The patient sample in the present study is based on the sample from the Rund et al. (2007) study, plus six more patients available after their analysis, i.e. 213 patients at baseline. Sixty-two of these were available for neurocognitive assessment at all four follow-ups, and are referred to as the follow-up sample (n = 62). The group of patients who missed at least one assessment will be referred to as the remaining sample (n = 151).
The follow-up sample consisted of approximately the same number of men and women, they were in their late twenties and had an estimated average IQ. Symptom ratings (PANSS) were severe at baseline, and the follow-up sample had one more year of education, a slightly higher IQ-estimate and a slightly lower GAF-Function and GAF-Symptoms scores than the remaining sample at baseline. At 3 months the groups obtained the same GAF-scores, but PANSS positive and negative scores were significantly higher in the remaining group. The remaining sample obtained a significantly weaker Verbal Learning index at baseline (data no shown).

The diagnostic distribution did not differ between groups at any of the follow-ups.

2.4. Illness severity as measured by number of relapses and diagnostic groups

The follow-up sample was divided into three groups; “No relapses after first episode” (n = 22), “One relapse after first episode” (n = 17), and “Two or more relapses after first episode” (i.e. 2–6 relapses (n = 23)) including eight patients who were continuously psychotic throughout the five years. Analyses did not show significant differences between the group with two or more relapses (i.e. 2–6 relapses) and those continuously psychotic in neurocognitive development, hence the groups were collapsed.

The three groups did not differ in demographic characteristics (gender, age, education), symptoms (PANSS positive/negative/general), function (GAF-S/F) or substance abuse (alcohol/drugs). No difference was found for diagnostic distribution, nor in time before first neurocognitive assessment. However, patients with no relapse had higher IQ and less negative change in premorbid social function than those with two or more relapses (see Table 2).

A separate grouping was made by dividing the follow-up sample in three groups based on diagnosis at 5 years: “Schizophrenia and schizophreniform disorder” (n = 35), “Affective psychosis with mood incongruent symptoms, and schizoaffective disorder” (n = 20), and “Brief psychotic episode, delusional disorder, and psychotic disorder NOS” (n = 7). The diagnostic groups did not differ in demographics, symptoms, function, or substance abuse at baseline. (Table 2)

2.5. Medication

At the 5 years follow-up 44 patients were using antipsychotic medication whereas 18 were medication free. Significantly more patients with two or more relapses used medication compared to those with no relapse. The defined daily dosage (DDD) was also significantly higher in the multiple-relapse group. The patients with one relapse were not significantly different from the other two groups.

2.6. Attrition and missing data

The five index scores consisted of a total of 13 subtests. At baseline, six of the subtests had missing data for one subject, one subtest had missing data for 2 subjects, and in three subtests there was missing data for 5 subjects. At 1-, 2- and 5-year follow-ups there was less missing data than was the case at baseline. In cases of missing data the group mean was inserted. This applied to less than 4% of the sample at each time point. The five index scores were calculated after missing scores were replaced.

2.7. Statistical methods

Analyses were conducted using the statistical package SPSS (PAWS) for Windows (version 18).

Statistical analyses were evaluated with t-tests for continuous variables and with chi-square tests for categorical variables. We used a within-group repeated measure multiple analysis of variance, MANOVA, to examine the neurocognitive development over time (four assessments) with the five neurocognitive indices as dependent variables.

Five separate one-way repeated measures ANOVAs were conducted to follow up results from the MANOVA; one for each of the neurocognitive indices, to analyze change over time.

The hypothesis of an association between neurocognitive functioning and number of relapses was examined by a second repeated measure MANOVA, with the sample divided in three relapse-groups as the between-subjects factor, and indices and time as the within-subjects factors. Five separate repeated measures ANOVAs were performed to follow up the results from the MANOVA. A multiple analysis of covariance (MANCOVA) using the variable “Duration of periods of medical

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Baseline</th>
<th>1 year</th>
<th>2 year</th>
<th>5 year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Verbal Learning-index</td>
<td>0.20 (0.72)</td>
<td>0.28 (0.77)</td>
<td>0.37 (0.76)</td>
<td>0.20 (0.82)</td>
</tr>
<tr>
<td>CVLT immediate recall (total over five trials)</td>
<td>55.6 (10.8)</td>
<td>56.6 (12.2)</td>
<td>58.2 (13.2)</td>
<td>57.1 (11.5)</td>
</tr>
<tr>
<td>CVLT delayed free recall</td>
<td>12.8 (2.7)</td>
<td>12.9 (2.6)</td>
<td>13.0 (2.9)</td>
<td>12.7 (2.9)</td>
</tr>
<tr>
<td>CVLT total errors</td>
<td>0.38 (0.63)</td>
<td>0.30 (0.48)</td>
<td>0.25 (0.43)</td>
<td>0.44 (0.63)</td>
</tr>
<tr>
<td>Motor Speed-index</td>
<td>0.04 (0.91)</td>
<td>0.11 (0.87)</td>
<td>0.22 (0.82)</td>
<td>0.16 (0.72)</td>
</tr>
<tr>
<td>FTT (dom.)</td>
<td>50.3 (8.3)</td>
<td>50.6 (8.2)</td>
<td>52.1 (6.8)</td>
<td>48.7 (6.4)</td>
</tr>
<tr>
<td>FTT (non-dom.)</td>
<td>46.5 (8.3)</td>
<td>47.5 (8.1)</td>
<td>47.8 (8.2)</td>
<td>44.6 (6.8)</td>
</tr>
<tr>
<td>Executive Functioning-index</td>
<td>0.15 (0.71)</td>
<td>0.23 (0.78)</td>
<td>0.19 (0.92)</td>
<td>0.25 (0.63)</td>
</tr>
<tr>
<td>WCST categories completed</td>
<td>5.5 (1.2)</td>
<td>5.5 (1.4)</td>
<td>5.5 (1.4)</td>
<td>5.5 (1.4)</td>
</tr>
<tr>
<td>WCST perseverative responses</td>
<td>13.8 (12.0)</td>
<td>11.4 (9.6)</td>
<td>11.9 (11.9)</td>
<td>9.9 (8.2)</td>
</tr>
<tr>
<td>WCST attempts to first category</td>
<td>18.2 (12.9)</td>
<td>18.4 (16.7)</td>
<td>20.0 (22.5)</td>
<td>20.1 (18.6)</td>
</tr>
<tr>
<td>Working Memory-index</td>
<td>0.11 (0.70)</td>
<td>0.27 (0.68)</td>
<td>0.35 (0.70)</td>
<td>0.47 (0.66)</td>
</tr>
<tr>
<td>COWA (sum of F-, A-, and F-words)</td>
<td>37.7 (9.7)</td>
<td>32.7 (10.7)</td>
<td>35.4 (11.5)</td>
<td>37.0 (10.4)</td>
</tr>
<tr>
<td>DSST (Digit Span with distractor)</td>
<td>78.2 (21.2)</td>
<td>82.6 (18.2)</td>
<td>81.4 (19.0)</td>
<td>85.5 (14.6)</td>
</tr>
<tr>
<td>DSST (Digit Span without distractor)</td>
<td>77.5 (19.5)</td>
<td>82.3 (16.6)</td>
<td>84.3 (16.4)</td>
<td>85.3 (13.9)</td>
</tr>
<tr>
<td>CPT-IP hits (correct responses to target trials)</td>
<td>0.60 (0.17)</td>
<td>0.66 (0.18)</td>
<td>0.66 (0.19)</td>
<td>0.67 (0.18)</td>
</tr>
<tr>
<td>Impulsivity-index</td>
<td>−0.04 (0.47)</td>
<td>0.07 (0.49)</td>
<td>0.18 (0.49)</td>
<td>0.18 (0.53)</td>
</tr>
</tbody>
</table>

Note: CVLT = California Verbal Learning Test (Delis et al., 1987); FTT = Finger tapping test (Lezak, 1995); WCST = Wisconsin Card Sorting Test (Heaton et al., 1993); COWA = Controlled Oral Word Association task (Spreen and Strauss, 1998); DSST = Digit Span Distractibility Test (Oltmanns and Neale, 1975; Rund, 1982); CPT-IP = Continuous Performance Test, Identical Pairs version (Cornblatt et al., 1989).
Follow-up analyses with separate repeated measures ANOVA revealed a significant association with relapse on two neurocognitive indices: VL (F(2, 59) = 5.1, p = 0.009, $\eta^2 = 0.15$) and WM (F(2, 59) = 6.5, p = 0.003, $\eta^2 = 0.18$). Results are shown in Figs. 2 and 3. The more severely ill group of patients scored poorer at all time points compared to the other two, less severely ill groups, although for VL significantly so only on the 1 and 2 year follow-ups.

Results of the repeated measure ANOVA with “diagnostic-groups at 5 years” as the between-subject factor showed a significant main effect of time, only (F(3,57) = 3.1, p = 0.033, $\eta^2 = 0.14$). No significant interactions were found for time and indices (F(12,48) = 1.0, p = 0.423, $\eta^2 = 0.20$), indices and diagnostic groups (F(8,112) = 1.2, p = 0.286, $\eta^2 = 0.08$) or the three-way interaction (F(24,96) = 1.1, p = 0.382, $\eta^2 = 0.21$).

4. Discussion

The five neurocognitive domains demonstrated different developmental trends from the 2 to 5 year follow-ups. Three of the five indices did not change over this time period, which is in line with the concept of a “stable encephalopathy” (Kurtz, 2005).

However, a significant and a near significant decline in the two remaining indices (Motor Speed and Verbal Learning) within the same time period brings nuance to the picture of neurocognitive stability after onset.

For Motor Speed, the significant decline between the 2 and 5 year follow-ups was not significantly correlated with age, and the five year increase in mean age (from 28 to 33 years) falls within the same age range defined by standard norms when analyzing test scores clinically, which also weakens a possible age-related change.

Effects of antipsychotic medication are previously shown to have a small or no influence on Motor Speed (Kopala et al., 2006; Goldberg et al., 2007), as measured by the Finger Tapping Test. Thus, this finding is difficult to explain and may be an accidental dip in performance, illustrating the need for further longitudinal studies.

We found a significant relationship between relapses and the two neurocognitive indices, Verbal Learning and Working Memory. The
Executive Function

Impulsivity

5-year follow-up in Fig. 2. 

In accordance with other previous reports (Alvarez-Jimenez et al., 2012). The 5-year relapse rate in our follow-up sample was 65%, which is in accordance with other previous reports (Alvarez-Jimenez et al., 2012).

Patients with no relapse after the first episode performed significantly better on the Verbal Learning-index, compared to the patients with one episode, or two or more episodes (see Fig. 2). The decline, however not significant, could be observed in spite of possible learning effects on this task.

For Working Memory, the association was less clear since the groups were significantly different at all time points. Still, even if the relapse groups developed in the same linear direction, the baseline difference remained constant throughout the five year follow-up. This indicates that the patients with more relapses did not improve to the same level as the patients having only one episode (no relapse group) (see Fig. 3). Thus, these findings represent new knowledge in the field of neurocognitive development in FEP. Repeated psychotic episodes may be a relevant differentiating factor for future studies.

Nevertheless, an association does not imply causality. A firm conclusion that cumulated relapses affect certain areas of neurocognitive functioning negatively is not warranted. The relationship might as well be the opposite: weaker neurocognitive performance could contribute to an increase in symptomatology, and hence more relapses. Looking back from the five year assessment, data suggests that poor Verbal Learning and Working Memory scores at baseline are associated with higher risk of relapses, as can be seen in Figs. 2 and 3. Also, there may be a yet unidentified third variable causing both relapses and decline in neurocognitive functions.

Diagnostic development constitutes a relevant factor in this regard, as the more severely ill patients by definition represent more patients with schizophrenia or affective psychosis. However, we found that groups based on total relapses were more informative than groups based on diagnosis, when identifying long-term neurocognitive trajectories. Thus, for neurocognitive course the essential factor seems to be relapsing episodes or non-remission, independent of diagnostic category, at least in a five year perspective. This may imply that neurocognitive dysfunction represents a symptom dimension that spans diagnostic categories, as it has previously been found across affective and non-affective psychosis (Lewandowski et al., 2011), and across major depression and schizophrenia (Stordal et al., 2005). Further, in patients with bipolar and SSDs, neurocognitive dysfunction is found to be determined more by history of psychosis than by diagnostic category or subtype (Simonsen et al., 2011).

The results of Rund et al. (2007) are supported by our findings of an association between relapses and deficits in Verbal Learning and Working Memory, still evident five year follow-up.

Our findings could indicate an episode-related cognitive deterioration in SSDs, but associations between disease progression and neurocognitive course is so far investigated to a limited extent. Our results need to be replicated, preferably based on studies beyond five years.

5. Limitations

There is no healthy control group. Neurocognitive test results may be influenced by practice effects, although the three year interval from 2- to 5-year follow-up may minimize this effect for the latest assessment.

6. Strengths

Our sample consists of the same patients followed over as long as five years, and all clinical variables and neurocognitive indices were assessed reliably.
Role of funding source
The project has been approved by the Regional Committee for Medical Research Ethics Health Region II (#5-9189) and the Regional Committee for Medical Research Ethics Health Region East (#1-2007.2177).

Data Inspectorate (License #96/30-17-1 and #2003.2052).

Biological data collection approved by Norwegian Directorate of Health (#200403453) and the Regional Committee for Medical Research Ethics Health Region East (#1-2007.2177).


