A Digit Symbol Coding Task as a Screening Instrument for Cognitive Impairment in First-Episode Psychosis

César González-Blanch*, Rocío Pérez-Iglesias, José Manuel Rodríguez-Sánchez, Gema Pardo-García, Obdulia Martínez-García, José Luis Vázquez-Barquero, Benedicto Crespo-Facorro

Psychiatry Research Unit of Cantabria, IFIMAV, CIBERSAM, University Hospital “Marqués de Valdecilla”, Santander, Spain

*Corresponding author at: Unidad de Investigación en Psiquiatría de Cantabria, Hospital Universitario “Marqués de Valdecilla”, Avda, Valdecilla s/n, 39008 Santander, Spain. Tel.: +34-942-202537; fax: +34-942-203447.
E-mail address: cgblanch@terra.es (C. González-Blanch).

Accepted 30 October 2010

Abstract

Cognitive impairment may be detected largely by examining the performance on a single neuropsychological measure. The purpose of the present study was to evaluate the validity and diagnostic accuracy of a coding task in comparison with other related tasks. One hundred thirty-one first-episode psychosis patients were administered five cognitive tasks related to a “speed of processing and executive functioning” dimension (Digit Symbol, Trail Making Test [TMT] parts A and B, Cancellation Test, and Digit Span-backward) and an additional measure of functional outcome. Digit Symbol provided good indices of accuracy and correlations with the global composite score of a comprehensive neuropsychological assessment represented large effect sizes. Correlations with a functional outcome were modest. Similar results were observed with the TMT. The processing speed, as measured by Digit Symbol, may be particularly good in capturing the generalized dysfunction which may be causing the widespread cognitive failures in schizophrenia spectrum disorders.

Keywords: schizophrenia; neuropsychological assessment; cognitive screening; coding task

Introduction

Schizophrenia is a clinical syndrome marked by the presence of symptoms such as auditory hallucinations and delusions, as well as social withdrawal, restricted affect, and poor motivation. Patients with schizophrenia show cognitive deficits that range between 1 and 3 SD below healthy controls, the most prominent of which are related to memory, attention, and executive functions (Heinrichs & Zakzanis, 1998). These deficits are clearly present at the time of treatment initiation (González-Blanch et al., 2006) and are largely independent of clinical symptoms (González-Blanch et al., 2008). As a result, these deficits are becoming increasingly recognized as a critical treatment target (Buchanan et al., 2005; Gold, 2004; Hyman & Fenton, 2003). However, formal cognitive assessments are not often conducted in routine clinical practice. Neuropsychological test batteries take a substantial length of time and the interpretation of the results for individual assessment is complex.

To overcome these limitations, several approaches have been investigated to reduce the time and complexity of comprehensive neuropsychological assessments. Some authors have developed a new set of neuropsychological tests with shorter administration times: The Brief Assessment of Cognition in Schizophrenia (BACS; Keefe Goldberg et al., 2004) and the Screen for Cognitive Impairment in Psychiatry (SCIP; Purdon, 2005) are some examples of validated cognitive screening tools in schizophrenia. Another approach consists of selecting only a small number of standardized tests that examine some key cognitive domains; for example, the Brief Cognitive Assessment (BCA; Velligan et al., 2004) and the Brief Cognitive Assessment Tool for Schizophrenia (B-CATS; Hurford, Marder, Keefe, Reise, & Bilder, in press). Alternatively, other authors have created brief interview-based scales with specific questions referring to cognitive deficits and the degree to which they affect day-to-day functioning, such as the Schizophrenia Cognition Rating Scale (Keefe, Poe, Walker, Kang, & Harvey,
2006) and the Clinical Global Impression of Cognition in Schizophrenia (Ventura, Cienfuegos, Boxer, & Bilder, 2008). This strategy seems particularly useful given the clinicians’ difficulties in differentiating between cognitive and other clinical symptoms (Bromley, 2007), and therefore, the questionable validity of evaluations through the cognitive items of psychopathological scales (Harvey et al., 2001; Hofer et al., 2007). All these types of screens are easy to administer, and the administration times vary from 15 to 30 min approximately, but are still scarcely available in clinical settings.

In a previous research, we found that the speed of processing, as measured by a Digit Symbol coding task, was severely impaired in first-episode schizophrenia spectrum disorders (González-Blanch et al., 2007) and that this might mediate a broader diversity of cognitive deficits (Rodríguez-Sánchez, Crespo-Facorro, González-Blanch, Pérez-Iglesias, & Vázquez-Barquero, 2007). The remarkable importance of speed of processing, as measured by Digit Symbol coding tasks, has been borne out by recent meta-analyses in schizophrenia (Dickinson, Ramsey, & Gold, 2007) and first-episode schizophrenia (Mesholam-Gately, Giuliano, Goff, Faraone, & Seidman, 2009). Likewise, in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study, the test that explained the greatest part of the global cognitive score at baseline assessment was the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III) Digit Symbol subtest, accounting for 61% of the variance of the global score (Keefe et al., 2006).

Digit Symbol coding tasks are multifactorial. Consistent with this, Digit Symbol tasks have not been linked to specific brain structure or function, and poor performance in tasks of this kind has have been associated with diverse syndromes that are either organic or functional (Groth-Marnat, 2003). However, it has been argued that much of what is measured by traditional neuropsychological assessments is general cognitive ability, not independent, domain-specific performance. Thus, much of the severity of cognitive impairments may arise from generalized factors, which might be the fundamental manifestation of schizophrenia (Dickinson, 2008). In pursuit of very brief and simple screening tools, the fact that generalized cognitive impairment could be the core deficit, rather than specific cognitive processes, has an obvious implication: A great part of cognitive impairment in schizophrenia may be detected by examining the performance on a single neuropsychological measure.

The purpose of the present study was to examine the validity and diagnostic accuracy in the assessment of global cognitive functioning in patients with first-episode psychosis of a Digit Symbol coding task in comparison with other tasks that loaded on the same cognitive dimension (i.e., speed of processing and executive functioning) in a previous factor analytic study (González-Blanch et al., 2007).

Materials and Methods

Setting and Subjects

The study participants were part of a cohort of consecutive admissions to the Cantabria Intervention Programme of First-Episode Psychosis (PAFIP, Spanish abbreviation) from February 2001 to February 2005. The PAFIP is located at the University Hospital “Marqués de Valdecilla” (Santander, Spain). Inclusion criteria for the PAFIP were as follows: age between 15 and 60; Diagnostic and Statistical Manual of Mental Disorder, Fourth edition (DSM-IV) criteria for diagnosis of schizophrenia, schizophreniform disorder, schizoaffective disorder, brief psychotic disorder, or psychosis not otherwise specified; never treated with antipsychotic medication; and living in the catchment area. Exclusion criteria were a history of neurological disease, head injury, mental retardation, or current drug dependence (DSM-IV criteria).

Of the 174 consecutive admissions who met criteria for enrolment, 131 (75.3%) completed baseline neuropsychological assessment. All diagnoses were confirmed 6 months after study entry using the Structured Clinical Interview for DSM-IV (SCID-I; First et al., 1995) by a trained clinician. The majority of participants were male (65%). The mean age of the sample was 27 years and the mean number of years of education was 10 (Table 1). All patients were randomly assigned to haloperidol, olanzapine, or risperidone (Crespo-Facorro et al., 2006). Eighty-eight participants (67%) were prescribed a second-generation antipsychotic medication. This study was approved by the Ethics Committee of the University Hospital “Marqués de Valdecilla,” and written informed consent was obtained from all participants or their legal representative.

For comparison purposes, 28 healthy volunteers (13 male, 15 female) were recruited from the local area through advertisements and underwent the same cognitive battery as patients. Healthy control participants met the following exclusion criteria: Past or current mental disorders screened by the Comprehensive Assessment of Symptoms and History (Andreasen, Flaum, & Arndt, 1992); regular medical prescriptions; neurological or general medical illnesses; or presence of psychosis in first-degree relatives. Patients and controls were similar in age and gender distribution (Table 1). They did, though, differ in years of education and premorbid intelligence quotient (IQ). Given the difference between controls and patients in premorbid IQ, the possibility of having a control sample not representative of the general population was indeed a concern. However, the estimation of current IQ (with Verbal Comprehension Index, a composite measure of verbal intellectual functioning highly correlated with full IQ scale) was similar to the average expected in a general population (mean, 101.2; SD, 14.5). When available, data from controls were
compared with Spanish standardized norms for age in individual tests (Peña-Casanova, Gramunt, & Gich, 2004; Wechsler, 1999); the controls’ performance was always within the range of average scores of population norms. Therefore, controls recruited in this study showed a cognitive performance that can be considered representative of the age-matched general population.

**Procedures**

On the basis of a previous factor analytic study, 30 cognitive subtest scores were grouped into eight cognitive dimensions (González-Blanch et al., 2007). To maximize collaboration and avoid acute psychosis state effects of acute psychosis, the cognitive battery was applied following the clinical stabilization of acute psychotic symptoms, with a mean of 10.58 (SD = 3.98) weeks after treatment initiation. For the purpose of test–retest reliability analysis, we used the data of a second application of the same tests 3 months later. Data were standardized to z-scores (with an average of 0 and SD = 1) using baseline healthy control group data. The mean of subtests z-scores was used to compute factor scores. For the purpose of the present study, we used the five variables out of four neuropsychological tests included in the factor called “Speed of processing and executive functioning” and a composite score of the whole battery.

**Neuropsychological Measures**

**Digit Symbol coding test (Wechsler, 1999).** This task consists of rows containing small blank squares, each paired with a randomly assigned number from one to nine. Above these rows is a printed key that pairs each number with a different
symbol. Using the reference key, the examinee has 120 s to pair specific numbers with given geometric figures. Scaled scores were obtained.

*Trail making test (TMT; Spreen & Strauss, 1998)*. The TMT consists of two parts (A and B). Part A requires the participant to connect series of numbered circles arrayed randomly on a sheet of paper using a pencil. In part B, the array consists of both numbers and letters, and the participant must connect them in alternating order. Part B demands simultaneous processing capacity for two sets of mental operations (number and letter sequencing) as well as a rule-following instruction to alternate between the sets. For parts A and B, scoring is expressed in terms of the time to completion.

*Cancellation test (Lezak, Howieson, & Loring, 2004)*. The participant is asked to scan a large field of either letters or symbols for a target letter (A). The participant is asked to cancel all the targets as quickly as possible for a minute. The score is the number of correct cancellations.

*Digit Span-backward subtest (Wechsler, 1999)*. This test is considered to be related to working memory. In this task, an increasing number of orally presented digits, starting with two, at a rate of one per second have to be repeated in an exactly reversed order. Every other trial the number of digits is increased by one. The test is ended when errors in two consecutive trials are made. Raw score is used as the key-dependent measure.

*Global deficit score (GDS)*. In order to summarize the neuropsychological data of the whole battery into a single score, we used the GDS method. The GDS is computed by converting standard scores (T-scores; with a mean of 50 and an SD of 10) on individual neuropsychological test variables into deficit scores ranging from 0 (no impairment) to 5 (severe impairment). A sum of all deficit scores divided by the total number of tests administered is calculated for each subject in order to obtain the final GDS (for further details on the calculation of this variable, see Carey et al., 2004). A higher value reflects a greater degree of overall impairment. In previous studies, a GDS of ≥0.50 has accurately predicted expert clinical ratings of overall impairment (Carey et al., 2004). The GDS method appears to be relatively resistant to modifications in test batteries or missing data. This method has shown convergent validity with other cognitive composite scores used in schizophrenia studies, such as the Clinically Significant Cognitive Impairment method (Reichenberg et al., 2009). The GDS was based on the eight cognitive dimensions (González-Blanch et al., 2007).

2. Verbal comprehension abilities: WAIS-III subtests (Wechsler, 1999): Vocabulary, Similarities, Information, and Comprehension. Additionally, the Vocabulary subtest was used to estimate premorbid IQ.
3. Speed of processing and executive functioning: TMT parts A and B (Spreen & Strauss, 1998); Cancellation test (Lezak et al., 2004); Digit Symbol coding; and Digit Span-backward (Wechsler, 1999).
4. Visual memory: Rey Complex Figure Test immediate and Delayed Recall (Rey, 1987).
7. Sustained attention/vigilance: Continuous Performance Test Degraded-Stimulus (CPT-DS) hits and reaction time (Cegalis & Bowlin, 1991); and Brief Test of Attention (Schretlen, Bobholz, & Brandt, 1996).

**Functional Measures**

To rate functional status, we used the global disability item from the Spanish version of Disability Assessment Schedule (DAS; Mañá, Ivorra, & Girón, 1998). The DAS evaluates the ability of the subjects to carry out particular social roles normally expected of them in their environment. Ratings were based on the clinician’s judgement of the information obtained from the patient, relatives, case notes, and observation of the patient during the month prior to 1-year follow-up. The global disability item has a score range of 0 (no disability) to 5 (gross disability). The patient and their relatives were assessed by the treating psychiatrist and the social worker independently, with a consensus being reached between both.

**Clinical Measures**

For characterization purposes, Table 1 shows the values for well-known clinical scales.

The presence of positive and negative psychotic symptoms was assessed by the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1983) and the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1981).
The Brief Psychiatric Rating Scale was used to assess the severity of general psychopathological symptoms (Overall & Gorham, 1962).

Duration of untreated psychosis (DUP) was defined as the time from the first continuous psychotic symptom corresponding to a score of 4 or more on one of the SAPS items to initiation of an adequate antipsychotic drug treatment. DUP was determined after interviewing the patient and a close relative. Duration of untreated illness was defined as the time from the first nonspecific symptom related to psychosis (for such a symptom to be considered, there should be no return to a previous stable level of functioning) to the initiation of adequate antipsychotic drug treatment.

Data Analyses

We examined the relationship between test variables and full cognitive assessment (i.e., GDS) by the Pearson correlation coefficients. To avoid spuriously inflated correlations, we computed the corrected item-total correlation (CITC) for the five test scores relative to the composite score (the GDS) using the reliability procedure. The CITC is the correlation between a test variable and the rest of the battery, without that item considered part of the test battery, thereby controlling for part-whole correlation.

For the five test variables, test–retest reliability was assessed by using intraclass correlation coefficients (ICC) and the Pearson correlation coefficients in the patient and control groups separately. Test–retest reliability data were available for 111 patients and 20 controls that were tested 3 months apart. Practice effects were measured by comparing data collected at test session 1 to those collected at test session 2 with within-group t-tests. These were determined in the patient and control groups separately.

Receiver operating characteristic (ROC) curves were calculated for the five tests to evaluate their screening performance. The ROC curve is a graphical representation of the sensitivity and specificity values of all possible cut-off values of a continuous diagnostic variable. The presence of a cognitive impairment assigned according to the 0.5 cut-off score in the GDS served as an external criterion. The area under the curve (AUC) directly represents the accuracy of the instrument in screening for cognitive impairment. Differences between the AUC values were tested using the Hanley and McNeil method (Hanley & McNeil, 1983). Determination of optimal cut-off scores was done by maximizing the Youden index (Y = sensitivity + specificity − 100; Youden, 1950).

We calculated sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV) for the optimal cut-off score for each test. We then used Mallow’s Cp (a measure of goodness of fit) to rank these single variable models and to determine the best subset of test variables.

Finally, we examined Spearman’s correlations between global functional outcome and each test variable. All statistical analyses were two-tailed and performed with Statistical Package for Social Sciences (SPSS) version 17.0 and MedCalc version 11.1 for Windows. The latter was mainly used for calculating the differences in AUC values.

Results

Correlation of Individual Test Variables to Global Score

We examined the relationship between the individual tests and GDS for full cognitive assessment. The Pearson correlation coefficients range from 0.39 to 0.71 indicating medium to large effect sizes. The CITC provided large effect sizes for Digit Symbol, TMT-A, and TMT-B (Table 2).

Test–Retest Reliability and Practice Effects of Individual Tests

Test–retest reliability was assessed using a single-measure, random model ICC from baseline to the 3-month follow-up assessment for individual tests. Practice effects were measured by comparing data collected at baseline to those collected 3 months apart with within-group t-tests. Table 3 shows data for 111 patients and 20 controls which were assessed at both time-points. Digit Symbol provides acceptably high values for both samples (>0.70). However, all but the Cancellation Test showed marked practice effects in the patients’ sample. As can be seen in Table 3, z-scores are consistently higher in a second administration. Controls performed significantly better in the second administration only on TMT-A, although this may be due to the lack of power to detect differences.
Table 2. Correlation of individual test variables to global cognitive score

<table>
<thead>
<tr>
<th>Test</th>
<th>Pearson r</th>
<th>CITC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit Symbol</td>
<td>-.657*</td>
<td>0.561</td>
</tr>
<tr>
<td>TMT-A</td>
<td>-.591*</td>
<td>0.492</td>
</tr>
<tr>
<td>TMT-B</td>
<td>-.714*</td>
<td>0.542</td>
</tr>
<tr>
<td>Digit Span-backward</td>
<td>-.394*</td>
<td>0.327</td>
</tr>
<tr>
<td>Cancellation</td>
<td>-.548*</td>
<td>0.432</td>
</tr>
</tbody>
</table>

Notes: CITC = corrected item-total correlation; TMT = Trail Making Test.
*p < .001.

Table 3. Mean performance and reliability of test–retest z-scores of cognitive tests in patients and controls

<table>
<thead>
<tr>
<th>Tests</th>
<th>FEP patients (n = 111)</th>
<th>Controls (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time 1</td>
<td>Time 2</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>-1.38</td>
<td>0.92</td>
</tr>
<tr>
<td>TMT-A</td>
<td>-.83</td>
<td>1.68</td>
</tr>
<tr>
<td>TMT-B</td>
<td>-1.97</td>
<td>2.50</td>
</tr>
<tr>
<td>Digit Span-backward</td>
<td>-.83</td>
<td>0.75</td>
</tr>
<tr>
<td>Cancellation</td>
<td>-1.00</td>
<td>1.94</td>
</tr>
</tbody>
</table>

Notes: TMT = trail making test; ICC = intraclass correlation coefficient;
*a data were standardized to z-scores (with an average of 0 and SD = 1) using baseline healthy control group data.

ROC Curve Analysis of Individual Cognitive Tests

Figure 1 shows the ROC curve for discriminating between patients with and without marked cognitive deficits (i.e., GDS > 0.5) as a function of their individual test scores.

The AUC serves as an overall measure of discrimination; perfect discrimination would yield an area of 1, whereas a test that failed to discriminate would have an area of 0.50. According to Hosmer and Lemeshow (2000), an AUC between 0.70 and 0.80 indicates acceptable discrimination, and at least 0.80 indicates excellent discrimination. All but one test, the Digit Span-backward, yielded an AUC over 0.70 (Table 4). Pairwise comparison of ROC curves revealed no statistical differences between AUCs, except for Digit Span-backward with Digit Symbol (ΔAUC = 0.139; z = 2.030; p = .042), TMT-A (ΔAUC = 0.158; z = 2.181; p = .029), and TMT-B (ΔAUC = 0.206; z = 3.188; p = .001). The more balanced cut-off scores of the tests, based on the Youden index, provided reasonable specificity values (>0.65), but TMT-A and Digit Span-backward had poor sensitivity values. All cognitive tests had high PPV (>85%), which is the probability that a person whose screening test result is positive really does have the condition of interest (i.e., marked cognitive impairment). This result, related to the specificity of the test, is particularly high for TMT-A and TMT-B. However, NPV, which is the probability that a person whose test result is negative does not have marked cognitive impairment, reaches, in the best of the cases, 57% (Table 4).

Best Subset Selection

We used Mallow’s Cp to rank single variable models and to select the best subset of all possible combinations of test variables. Cp is expected to be approximately equal to p + 1, where p is the number of predictors in the fitted model, with smaller values preferred. If the model fitted is not adequate, the larger values of Cp are expected (King, 2003). On the basis of Mallow’s Cp statistic, Digit Symbol yields the best-fit model for a subset of 1 variable (Cp = 6.53; AUC = 0.80; Nagelkerke $R^2 = .31$), followed by TMT-B (Cp = 12.22; AUC = 0.87; Nagelkerke $R^2 = .48$), TMT-A (Cp = 13.65; AUC = 0.82; Nagelkerke $R^2 = .36$), Cancellation (Cp = 19.29; AUC = 0.78; Nagelkerke $R^2 = .29$), and Digit Span-backward (Cp = 26.93; AUC = 0.66; Nagelkerke $R^2 = .08$). Furthermore, Digit Symbol was also part of the best subset of two variables with TMT-A (Cp = 1.41; AUC = 0.87; Nagelkerke $R^2 = .47$) and was also part of the best combination of three variables with TMT-A and TMT-B (Cp = 2.01; AUC = 0.90; Nagelkerke $R^2 = .54$). The best of all possible subsets was the combination of Digit Symbol and TMT-A for the patients’ sample; and Digit Symbol, TMT-A, and TMT-B for the whole sample (i.e., patients and controls; Cp = 1.60; AUC = 0.91; Nagelkerke $R^2 = .61$).
Correlations Between Individual Test and a Global Measure of Social Disability

The Spearman correlations between cognitive tests and DAS global score indicated a small negative relationship with Digit Symbol \(r = -0.25; p = .007\), TMT-A \(r = -0.23; p = .010\), and TMT-B \(r = -0.28; p = .002\), but a lack of significant relationship with Cancellation and Digit Span-backward tests.

Discussion

All the tests included in the present study are pen-and-paper tasks with desirable characteristics for screening tools, such as being easy to administer and score so that they will be more widely accepted among busy clinicians, and are also well tolerated by patients. Furthermore, these neuropsychological tests are already used worldwide in research and clinical settings. The data in this study indicate that single neuropsychological tests such as Digit Symbol, TMT (parts A and B), and Cancellation are potentially useful screening instruments in a sample of first-episode psychosis patients. The results for Digit Span-backward were rather disappointing; suggesting that this particular test should only be used in combination with other measures in order to obtain a valid estimation of general cognitive ability. On the other hand, based on ROC analyses, there were no significant differences in the accuracy for detecting marked cognitive impairment between the tests used, except Digit Span-backward. Depending on the criteria used the order of the rank will vary, but in all cases Digit Symbol, TMT-A, and TMT-B have

---

**Table 4.** Sensitivity, Specificity, PPV, NPV, and AUC for cognitive tests optimal cutoff z-scores

<table>
<thead>
<tr>
<th>Tests</th>
<th>Optimal cut-off z-score</th>
<th>Se</th>
<th>Sp</th>
<th>PPV</th>
<th>NPP</th>
<th>AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit Symbol</td>
<td>Less than −1.12</td>
<td>0.78</td>
<td>0.67</td>
<td>0.87</td>
<td>0.51</td>
<td>0.798 (0.713–0.882)**</td>
</tr>
<tr>
<td>TMT-A</td>
<td>Less than −0.74</td>
<td>0.60</td>
<td>0.94</td>
<td>0.97</td>
<td>0.46</td>
<td>0.817 (0.743–0.890)**</td>
</tr>
<tr>
<td>TMT-B</td>
<td>Less than −0.80</td>
<td>0.77</td>
<td>0.85</td>
<td>0.94</td>
<td>0.57</td>
<td>0.865 (0.800–0.929)**</td>
</tr>
<tr>
<td>Digit Span-backward</td>
<td>Less than −0.84</td>
<td>0.56</td>
<td>0.74</td>
<td>0.86</td>
<td>0.37</td>
<td>0.659 (0.549–0.768)*</td>
</tr>
<tr>
<td>Cancellation</td>
<td>Less than −0.18</td>
<td>0.71</td>
<td>0.71</td>
<td>0.87</td>
<td>0.46</td>
<td>0.774 (0.687–0.862)**</td>
</tr>
</tbody>
</table>

Notes: TMT = Trail Making Test; ROC = receiver operating characteristic; Se = sensitivity; Sp = specificity; PPV = positive predictive value; NPV = negative predictive value; AUC = area under the curve; CI = confidence interval.

*p < .01.

**p < .001.
the best indices of correlation with GDS, variance explained, and AUC. According to Mallow’s Cp, Digit Symbol was the best single test for screening cognitive impairment, and the addition of TMT-A further increased the accuracy. The combination of Digit Symbol, TMT-A, and TMT-B provided excellent indices of AUC and Mallow’s Cp and accounted for a substantial amount of the variance of the GDS. A caveat regarding the results of the present study is that, in the absence of a real cognitive improvement (see Gonza´lez-Blanch et al., 2006; Rodrı´guez-Sa´nchez et al., 2008), patients’ scores on the three tests were significantly higher in the second administration, 3 months apart. However, the consistency between the two assessments, as measured by ICC, was particularly good for Digit Symbol. These three subtests (Digit Symbol, TMT-A, and TMT-B) were the only ones that showed significant associations with a global measure of social and occupational functioning, but the magnitude of these correlations represented small to medium effect sizes (rho \approx −0.25).

Further limitations of the present study have to be noted. First, although all these tests have alternate forms available, we did not use alternate forms to control for the assessment of potential practice effects and test–retest reliability. Second, it is likely that other speeded tasks that require information processing, visual scanning, flexibility, and motor abilities (or even related to other cognitive domains) may provide similar results. It is worth noting that WAIS Digit Span-backward involves cognitive functions of a different type (such as storage and reversing operations as well as auditory attention) and its correlation with GDS and accuracy in detecting cognitive impairment as a single test was notably poorer than the other tests used. Third, we used a global measure of functional outcome instead of specific measures of social or occupational performance. Although we have questioned the associations between neurocognition and functional status elsewhere (Gonzalez-Blanch et al., 2010), it is possible, however, that more specific measures may provide stronger correlations than those observed in the present study. Fourth, the clinical sample used in this study was mainly composed of subjects with a diagnosis of schizophrenia and schizophreniform disorder (85% of the whole sample). Thus, it is possible that our results cannot be generalized to those first-episode patients with other diagnoses. Finally, since cognitive assessment was conducted after clinical stabilization, the influence of antipsychotics on cognitive functions cannot be analyzed. However, previous analyses suggest the stability of cognitive deficits during the first weeks of treatment (Gonza´lez-Blanch et al., 2006, 2008) and in 1-year follow-up (Rodrı´guez-Sa´nchez et al., 2008). In a drug-naı¨ve first-episode schizophrenia study, Hong and colleagues (2002) suggested that most of the cognitive functions remain stable during the early phase of treatment (i.e., the first 8 weeks).

From the theoretical point of view, there are two complementary interpretations for the fact that several single test scores account for much of the variance of global cognitive functioning and can correctly classify most of the patients. On one hand, it can be assumed that basic information processing represents an essential feature of cognitive dysfunction among individuals with schizophrenia. This is in agreement with previous research (Badcock, Dragovic, Waters, & Jablensky, 2005; Mohamed, Paulsen, O’Leary, Arndt, & Andreasen, 1999; Niendam et al., 2003; Rodrı´guez-Sa´nchez et al., 2007). At the same time, it can be interpreted to mean that the generalized cognitive deficit is a fundamental manifestation of schizophrenia. These multifaceted tasks may well be capturing this essential aspect of the illness. Some factorial studies in schizophrenia report moderate correlations between cognitive domains (Gonzalez-Blanch et al., 2007) and between cognitive domains and a global composite score (Keefe et al., 2006). Likewise, Dickinson, Iannone, Wilk, and Gold (2004) showed that about two-thirds of the diagnosis-related variance in cognitive performance was mediated through a general factor, with more specific domains (such as verbal memory and processing speed) having small direct effects. In this line, Dickinson (2008) suggests that generalized cognitive impairment should not be dismissed as an artifact that obscures more cognitive domain-specific effects. Coding tasks will be particularly sensitive to the biological underpinnings of the illness, which may be general and systemic in nature, rather than focal abnormalities in specific brain structures (Dickinson, 2008). This point of view is largely consistent with pathophysiological models of schizophrenia such as the “cognitive dysmetria” model. This model postulates that diverse symptoms reflect abnormalities in connectivity in the cortical–cerebellar–thalamic–cortical circuit (Andreasen, Paradiso, & O’Leary, 1998). An abnormality in this circuitry may lead to difficulties in coordinating both motor and mental activities. Thus, processing speed deficits may be reflecting a generalized dysfunction that causes the widespread cognitive failures in attention, memory, motor dexterity, and executive functions.

From the clinical perspective, the results from this study suggest that it is possible to detect cognitive deficits by means of a brief test that accounts for a considerable amount of the variance associated with a more comprehensive assessment of cognitive functioning in first-episode cases. TMT and Digit Symbol have also been included in other cognitive screening batteries which are created by selecting a small number of standardized neuropsychological tests. For example, the BCA (Velligan et al., 2004) includes as part of the battery the Trails A and B, and the B-CATS (Hurford et al., in press) includes Trails B and Digit Symbol substitution. Although using a set of two or three tests can improved the psychometric properties of the screening instrument, this study has shown that the Digit Symbol test alone had correlations with the global score and had test–retest reliability similar to that of BCA and B-CATS. These very short administration times may be particularly valuable for busy, time-pressured clinicians interested in identifying those cases with marked cognitive impairment which would benefit from a more comprehensive neuropsychological assessment or in need of special rehabilitation programs.
On the other hand, it could be assumed that if the processing speed deficit could be improved, then the other deficits would be less severe (Duff & Grabowski, 2008). Importantly, growing evidence suggests that such deficits in speed of processing may be open to amelioration through cognitive remediation therapies (Hogarty, Greenwald, & Eack, 2006; Wykes, Reeder, Corner, Williams, & Everitt, 1999). As global measures of cognition may be more relevant for the determination of functional outcomes than specific neurocognitive constructs (Green, Kern, Braff, & Mintz, 2000), the improvement in individual tests that reflect global cognitive deficits may be particularly important for functional recovery in people with schizophrenia and related disorders.

In conclusion, identification of early cognitive impairment in schizophrenia spectrum disorders is important but often missed in clinical settings. Full neuropsychological assessments are time-costly and not feasible in routine practice, whereas brief screening tools such as the above cited are still not widely applicable to date. The use of a single standardized test that taps information processing, visual scanning, flexibility, and motor abilities may represent a suitable and cost-effective tool for the screening of cognitive impairment in clinical settings.

Funding

This work was supported by Instituto de Salud Carlos III (FIS 00/3095 and G03/032) and SENY Fundació Research Grant (2005-0308007), Fundación Marqués de Valdecilla (A/02/07). These institutions 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.

Conflict of interest

J.L.V.-B. and B.C.-F. have received unrestricted research funding from AstraZeneca, Pfizer, Bristol-Myers Squibb, and Johnson & Johnson, which has been deposited into research accounts at the University of Cantabria. J.L.V.-B. has received honoraria for his participation as a speaker at educational events from Johnson & Johnson. B.C.-F. has received honoraria for his participation as a speaker at educational events from Pfizer, Bristol-Myers Squibb, and Johnson & Johnson and consultant fees from Pfizer. R.P.-I. has received honoraria for his participation as a speaker at educational events from Bristol-Myers Squibb-Otsuka.

The remaining authors report no competing interests.

Acknowledgements

We gratefully acknowledge the participation of patients and controls in the present study. The authors wish to thank Dr Mario Álvarez-Jiménez of ORYGEN Research Centre, University of Melbourne, for his methodological advice, and the contributions made by the staff of the Cantabria Intervention Programme of First-Episode Psychosis (PAFIP).

References

Andreasen, N.C. (1981). Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City, IA.


