Original article

Psychometric properties of a brief neuropsychological protocol for use in geriatric populations

Propriedades psicométricas de um protocolo neuropsicológico breve para uso em populações geriátricas

JONAS JARDIM DE PAULA¹, CARLOS GUHLERME SCHLOTTFELD², LAFAIETE MOREIRA³, MARIA COTTÁ³, MARIA APARECIDA BICALHO³, MARCO AURELIO ROMANO-SILVA³, HUMBERTO CORRÊA³, EDGAR NUNES MORAES³, LEANDRO FERNANDES MALLOY-DINIZ⁴

¹ Neurosciences Program, Institute of Biological Sciences, Universidade Federal de Minas Gerais (UFMG).
² Gericatric and Gerontology Center, Clinics Hospital, UFMG.
³ Department of Internal Medicine, School of Medicine, UFMG. Geriatric and Gerontology Center, Clinics Hospital, UFMG.
⁴ Department of Mental Health, School of Medicine, UFMG. Neurosciences Laboratory, School of Medicine, UFMG.
⁵ Department of Psychology, UFMG. Neurosciences Program, Institute of Biological Sciences, UFMG.

Received: 15/3/2010 – Accepted: 25/6/2010

Abstract

Background: The increase in life expectancy and proportion of elderly in the population is causing an increase in dementia prevalence rates. The correct, early diagnosis of dementia is very important to clinical treatment and to improved prognosis. Therefore, it is necessary to adapt and develop assessment tools for the differential diagnosis between pathological and normal aging processes. Objective: Assess the psychometric properties and the factorial structure of a neuropsychological protocol used in geriatric assessment. Method: Subjects (n = 69) with heterogeneous cognitive complaints were assessed at the Centro de Geriatria e Gerontologia at the Hospital das Clínicas da Universidade Federal de Minas Gerais using a protocol composed of the Mini-Mental State Examination, Clock Drawing, Corsi Blocks, Verbal Fluency, Digit Span and Token Test. Statistical analyses included factorial analyses of test results, Pearson’s correlation between obtained factor, age, years of formal education and clinical dementia rating and area under the ROC curve. Results: The factorial analyses of test scores showed a general representative factor that had moderate and significant association with CDR (r = -0.672; p < 0.001) and years of formal education (r = 0.455; p < 0.001), respectively. This factor had weaker and less significant correlation with age (r = -0.282; p < 0.05). Discussion: These results point to the protocol’s good construct and criteria validity in assessing cognitive decline in the elderly. Future works concerning applicability and populational norms are needed to improve the clinical use of this assessment protocol.

Keywords: Neuropsychological assessment, geriatric health, dementia, cognition.

Resumo

Contexto: O aumento na expectativa de vida e na proporção de idosos na população tem acarretado elevação nas taxas de prevalência de demências. O diagnóstico correto da demência é muito importante para o tratamento clínico e para um melhor prognóstico. Por isso, é necessário adaptar e desenvolver instrumentos para o diagnóstico diferencial entre os processos de envelhecimento normal e patológico. Objetivo: Avaliar as propriedades psicométricas e a estrutura fatorial de um protocolo neuropsicológico usado para avaliação geriátrica. Método: Pacientes (n = 69) com queixas cognitivas heterogêneas foram avaliados no Centro de Geriatria e Gerontologia do Hospital das Clínicas da Universidade Federal de Minas Gerais, a partir de um protocolo composto pelo Mini-Exame do Estado Mental, Desenho do Relógio, Cubos de Corsi, Fluência Verbal, Span de Dígitos e Token Test. Análises estatísticas incluíram análise fatorial dos resultados dos testes, correlação de Pearson entre o fator obtido e a idade, escolaridade, anos de educação formal e o Clinical Dementia Rating e a área sob a curva ROC. Resultados: A análise fatorial dos escores do teste mostrou um fator geral representativo que teve associação moderada e significativa com o CDR (r = -0.672; p < 0.001) e anos de educação formal (r = 0.455; p < 0.001). Esse fator teve fraça mas significativa correlação com a idade (r = -0.282; p < 0.05). Conclusões: Esses resultados apontam para uma boa validade de construto e de criterio do protocolo na avaliação do declínio cognitivo de idosos. Estudos futuros sobre aplicabilidade e normas populacionais são necessários para aprimorar o uso clínico deste protocolo de avaliação.

Palavras-chave: Avaliação neuropsicológica, saúde geriátrica, demência, cognição.

Introduction

According to the Summary of Social Indicators¹, Brazil has approximately 19 million people aged 60 years or older. These individuals are concentrated in the Southeast region of the country, including São Paulo (4.4 million), Minas Gerais (2.1 million) and Rio de Janeiro (2.1 million), which together represent 45% of the country’s elderly population. Lima-Costa and Veras² reported that the Brazilian elderly population will likely exceed the 32 million mark in 2020. The growing proportion of elderly in the population is a direct result of increasing life expectancy at birth. In Brazil, for example, life expectancy at birth was 55.7 years in 1960 and today it is 72.3³. As noted by Lopes and Bottino⁴, one of the consequences of increased life expectancy and the proportion of elderly in the population is increased dementia prevalence rates since aging is one of the most important risk factors for the development of dementia. After 65 years of age, the prevalence of disease doubles every five years⁵. The findings of Herrera et al.⁶ demonstrate the relationship between ag-
ing and the emergence of dementia. They found that the prevalence of dementia in the Brazilian elderly population ranged from 1.3% (from 60 to 69 years) up to 36.9% (85+ years).

The most frequently diagnosed demential neurodegenerative syndrome is Alzheimer’s disease, accounting for 50%-70% of cases, while the frontotemporal is responsible for approximately 5% of cases and non-degenerative diseases caused by vascular problems make up about 20% of cases. These diseases significantly affect the quality of life of patients, besides being strongly related to mortality in elderly patients. Normal or pathological cognitive decline associated with aging is closely related to independence in daily activities. For example, the impairment of executive functions is one of the main factors in predicting good adaptation of the elderly in everyday activities. A compromised memory system, besides causing serious impediments in day-to-day tasks, is associated with depression and significant worsening of quality of life in this population.

In a Brazilian study (Pereira et al.), Vasconcelos et al. argue that there are few standardized neuropsychological instruments adapted for cognitive assessment in the elderly and, in addition, there are few validity and normative studies on the instruments commonly used in this population. However, in recent years several studies were carried out in Brazil aimed at obtaining norms and assessing the validity of cognitive batteries including the Dementia Rating Scale, ADAS-Cog, CAMCOG and neuropsychological tests like the Rey Auditory-Verbal Learning Test and Phonemic Verbal Fluency Task. Additional studies concerning the use of neuropsychological tests that assess different cognitive domains in informal batteries would be useful for both clinical and research activities in Brazil, especially when a differential diagnosis is required. Furthermore, brief batteries that are easy to use and in the public domain would improve neuropsychologists’ access to diagnostic and longitudinal treatment tools in elderly patients with cognitive complaints. In view of these deficits, the present study seeks to evaluate the psychometric properties of a neuropsychological protocol for assessing elderly patients. To that end, aspects related to internal consistency, diagnostic power, construct and criterion validity were evaluated.

Six neuropsychological instruments were chosen for the protocol: the Mini-Mental State Examination (MMSE), Clock Drawing, Corsi Blocks, Verbal Fluency, Digit Span and Token Test. All protocol items are easy to administer, quick, low cost and widely recognized in the neuropsychological literature. In addition to assessing the protocol’s psychometric properties, the following assumptions were investigated based on the scientific literature concerning neuropsychological assessment:

1. Patients’ performance on the neuropsychological trial protocol would present strong-to-moderate correlation with their performance on the clinical dementia rating (CDR).
2. Formal schooling and age would be a good predictor of cognitive performance and 3) the instruments would be able to generate a representative general factor for the protocol as a whole.

Methodology

Sample

The draft protocol was applied to a group of 69 patients with heterogeneous cognitive complaints who underwent neuropsychological assessment during an open event about memory and dementia held at the Geriatric and Gerontologic Center at the Hospital das Clinicas of the Universidade Federal de Minas Gerais (UFMG). To participate in the study subjects had to have expressed, at least subjectively, a cognitive complaint on any functional field. No subjects had received a diagnosis prior to the evaluation process. All patients were also classified according to the CDR.

Instruments

a. MMSE: a cognitive screening instrument that enables the overall assessment of the patient’s mental state using 11 simple tasks. It evaluates temporal orientation, spatial memory, attention, language and praxia. The current study employed a Brazilian version based on which adapted by Bertolucci et al. The test scores range from 0 to 30 and the most used cutoff score is 22.

b. Clock Drawing: a test that is part of several batteries of psychological assessments and found in many versions. In the present protocol, a subject draws an empty circle, representing the shape of the clock, and then he/she must draw the numbers, pointers and time. The test assesses visuospatial skills, praxia and higher cognitive skills, including executive functions. The version developed by Shulman et al., with scores ranging from 0 (worse) to 5 (better) was used here.

c. Digit Span: one of the main tasks used in assessing short-term memory in which an examiner asks a patient to repeat a series of numbers (e.g.: 2-5; 3-6-7). The first sequence begins with two digits. After each correct answer, the examiner adds a digit in the next sequence. The test is composed of two parts, direct and reverse order but only direct order was used in the current study. Digit Span evaluates storage capacity in short-term memory and its executive component, especially regarding remembering in reverse order. The version adapted from Wechsler Adult Intelligence Scale is used here, but with the scoring criteria as in the Corsi Blocks.

d. Corsi Blocks: a spatial instrument analogous to the Digit Span test that evaluates the span of short-term memory using visual space. Initial studies found a pattern of double dissociation between performance on digit span and Corsi Blocks in patients with cerebral lesion in the right and left hemispheres. The test consists of a square base with nine blocks. Subjects are instructed to replicate the examiner’s manipulation of the pieces. The test score is computed by multiplying the number of correct sequences by the maximum span achieved by the patient.

e. Semantic Verbal Fluency: an exercise where participants name as many fruits as possible in 60 seconds without repeating any. This task measures the fluency component of the executive functions and is related to activity in both the dorsolateral prefrontal and temporal cortex of the left hemisphere.

f. Token Test-Short Version: this instrument assesses language comprehension. The task contains 36 commands divided in six parts where increases in the quantity of information of each successive command raise the difficulty of the item. The Token Test is significantly sensitive to left hemisphere dysfunctions, the effects of childhood brain development and demential processes.

g. CDR: a scale that sorts the various degrees of dementia and identifies questionable cases (e.g. those that are not considered normal). The cases may be age-associated cognitive decline or mild cognitive impairment, which epidemiological studies have shown to have a higher conversion rate to dementia.

Procedures

The medical and neuropsychological assessments were performed at the Centro de Geriatria e Gerontologia do HC-UFGM. Patients were initially examined by geriatricians and geriatric residents and then referred for neuropsychological evaluation. At the time of evaluation, the neuropsychologists were unaware of the outcome of the geriatric...
assessments. The neuropsychological assessment procedures were also done in UFMG's Gerontology Service facilities. The examination rooms had natural ventilation, were quiet and were appropriately equipped to perform the tests and implement the protocols for the neuropsychological assessment. All procedures were applied in the same manner and order by trained researchers. The study proposal was submitted to and authorized by the UFMG Ethics Committee (COEP-334/06).

Data analysis

Descriptive statistics were performed to characterize the sample concerning sociodemographic variables and each subject's scores. Considering the assessment protocol, each subtest result was standardized to acquire the total score for the battery. To ensure a normal distribution, the scores were scaled producing a distribution with a mean of 10 and standard deviation (SD) of three. After summing the six scores, the resulting score was computed into a linear transformation generating a single value, named total score, with a mean of 50 and SD of 10. Also, to further investigate the sensitivity and specificity of the diagnostic power of the battery, all seven scores (six subtests, total score) were analyzed through a receiver operating characteristics (ROC) procedure considering CDR classification (cognitive impairment absence when equal or minor to 0 or presence when greater than 0).

The raw test scores were submitted to exploratory factor analysis using the principal axis factoring method. Using a matrix of correlations between the variables (raw scores), this method provides a common set of latent dimensions called factors. However, two data adequacy tests must be done before factor analysis can be performed, namely the Kaiser-Meyer-Olkin (KMO) and Bartlett tests. The KMO adequacy tests must be done before factor analysis can be performed, indicating whether the partial correlations between variables are weak while the Bartlett test examines whether the correlation matrix generated by the crossing of variables is significant enough for factor extraction.

Following this procedure, a single dimension was transformed through a regression process into a new variable named general cognition (GC) that was used to estimate the association between multiple variables. The Pearson correlation was calculated for the GC values, schooling, CDR and age.

Results

A total of 69 subjects were assessed. The group was 62.9% female and had a mean age of 71.7 years (SD = 6.54 years) and average schooling of 4.7 years (SD = 4.0 years). Of the total, 45.7% had no dementia. 24.3% had mild cognitive decline and 30% had pathological decline (CDR ≥ 1). Table 1 illustrates the mean and standard deviation results of the tests that make up the protocol.

Considering the other psychometric properties of the battery, it was found that the Total Score, MMSE, Verbal Fluency, Clock Drawing, Corsi Blocks and Token Test had an area under the ROC curve (AUC) between 0.7 and 0.9 (Table 2) meaning a moderate discrimination power for cognitive impairment. The Digit Span achieved a ROC AUC between 0.5 and 0.7 and this result can be considered weak. Taken together the subtests produced a Cronbach’s alpha of 0.857 pointing to moderate-to-high internal consistency of the assessment protocol.

Regarding the factorial analyses of the protocol, a KMO test value of 0.863 and Bartlett test (p < 0.001) indicated that the data were suitable for extracting factors. This analysis revealed a single factor, GC, which had an eigenvalue of 3.512 that alone accounted for 58.5% of variance. All the items comprising this factor were weighted above 0.6: MMSE 0.820; Verbal Fluency (fruits) 0.683; Clock Drawing Test 0.681; Corsi Blocks 0.667; Digit Span 0.601; Token Test 0.777.

The correlations revealed moderate associations and were significant for the GC values, schooling, CDR and age. The correlations revealed moderate associations and were significant for the GC values, schooling, CDR and age.

Table 1. Subjects performance on protocol according CDR

<table>
<thead>
<tr>
<th></th>
<th>No Impairment CDR = 0</th>
<th>Possible cognitive impairment CDR = 0.5</th>
<th>Early cognitive impairment CDR = 1</th>
<th>Established cognitive impairment CDR = 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (standard deviation)</td>
<td>Mean (standard deviation)</td>
<td>Mean (standard deviation)</td>
<td>Mean (standard deviation)</td>
</tr>
<tr>
<td>Age</td>
<td>70 (5)</td>
<td>71 (6)</td>
<td>73 (6)</td>
<td>77 (4)</td>
</tr>
<tr>
<td>Gender</td>
<td>23</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Male</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Years of formal education</td>
<td>5 (4)</td>
<td>4 (4)</td>
<td>4 (4)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Mini Mental Status Examination</td>
<td>26.06 (3.62)</td>
<td>21.88 (4.11)</td>
<td>20.54 (2.57)</td>
<td>13.57 (1.13)</td>
</tr>
<tr>
<td>Semantic Verbal Fluency (Fruits)</td>
<td>12.66 (4.51)</td>
<td>8.24 (2.77)</td>
<td>6.92 (3.17)</td>
<td>6.00 (2.00)</td>
</tr>
<tr>
<td>Clock Drawing Test</td>
<td>3.66 (1.64)</td>
<td>2.94 (1.48)</td>
<td>1.85 (1.34)</td>
<td>1.43 (0.98)</td>
</tr>
<tr>
<td>Digit Span</td>
<td>7 (2)</td>
<td>6 (2)</td>
<td>6 (2)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Corsi Blocks</td>
<td>28 (14)</td>
<td>20 (11)</td>
<td>20 (14)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>Token Test</td>
<td>30 (4)</td>
<td>27 (3)</td>
<td>27 (4)</td>
<td>21 (3)</td>
</tr>
<tr>
<td>Total Score</td>
<td>56 (8)</td>
<td>48 (6)</td>
<td>45 (6)</td>
<td>36 (8)</td>
</tr>
</tbody>
</table>

Table 2. ROC Analyses of tests and total score of the protocol

<table>
<thead>
<tr>
<th></th>
<th>Cut-off point</th>
<th>Sensibility</th>
<th>Sensibility</th>
<th>AUC*</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>21</td>
<td>.816</td>
<td>.750</td>
<td>0.865**</td>
<td>0.782 - 0.952</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>9</td>
<td>.796</td>
<td>.650</td>
<td>0.818**</td>
<td>0.720 - 0.918</td>
</tr>
<tr>
<td>Clock Drawing Test</td>
<td>3</td>
<td>.694</td>
<td>.700</td>
<td>0.794**</td>
<td>0.690 - 0.898</td>
</tr>
<tr>
<td>Corsi Blocks</td>
<td>18</td>
<td>.714</td>
<td>.700</td>
<td>0.715**</td>
<td>0.570 - 0.860</td>
</tr>
<tr>
<td>Digits Span</td>
<td>6</td>
<td>.653</td>
<td>.600</td>
<td>0.663***</td>
<td>0.508 - 0.819</td>
</tr>
<tr>
<td>Token Test</td>
<td>27</td>
<td>.776</td>
<td>.700</td>
<td>0.755**</td>
<td>0.620 - 0.890</td>
</tr>
<tr>
<td>Total Score</td>
<td>48</td>
<td>.857</td>
<td>.800</td>
<td>0.846**</td>
<td>0.749 - 0.943</td>
</tr>
</tbody>
</table>

* According to Swets: AUC < 0.5 means lack of statistical significance; AUC from 0.5 to 0.7 means low accuracy; AUC from 0.7 to 0.9 means moderate accuracy; AUC ≥ 0.9 means high accuracy.
** Significant at p < 0.01.
*** Significant at p < 0.005.
Discussion

Considering the main hypothesis, the instruments presented in the neuropsychological battery were able to generate a representative general factor for the protocol as a whole. The tested protocol was able to generate a general factor that explained 58% of variance and had weighted factorials higher than 0.6. These results indicate that this emerging factor is involved consistently in all the items that compose it and that the results of this item are homogeneously distributed. In short, the results point to good construct validity\(^8\). A moderate-to-high correlation between the CDR and GC indicates that both tests measure constructs close, but are not exactly the same. This result is consistent with the hypothesis that the protocol presents a high correlation with CDR since the protocol evaluates cognition and the CDR evaluates the consequences of cognitive decline. A negative and significant correlation was found between GC and age while a positive and significant correlation was identified between GC and schooling. It is expected that a cognitive assessment test of elderly subjects would present a relationship with both age and schooling, and therefore the protocol is consistent with the human developmental trend that is well established in the literature\(^2\), providing further evidence of construct validity\(^8\). Taken together, the results of the relationship between GC, CDR, age and schooling suggest strong evidence of external validity of the battery.

Considering the AUC, most of the battery tests produced good discrimination power except for the Digit Span subtest. But it must be stressed that such a result cannot be considered independently since the battery includes subtests, like this one, that are not specifically designed to assess early stages of dementia. Nevertheless, the total score obtained as a sum of the subtests produced a high AUC, indicating good diagnostic power. The high internal consistency of the battery and this moderate-to-high accuracy reinforces its good psychometric properties.

The method and techniques used to evaluate the protocol and the subsequent convergence of the various cognitive measures into a strong general factor suggests a relationship with consolidated constructs in psychological studies of individual differences, a General Intelligence Factor. In a meta-analysis, Carroll\(^16\) combined cognitive testing data from more than 130,000 subjects and confirmed the existence of a general factor that permeated all other investigated skills. In view of the inherent property of an instrument that evaluates aspects of cognition, the GC factor can be seen to some extent, as an expression of the general factor. The factor found in our analysis can also be interpreted based on the concept of cognitive reserve, which holds that individual difference in cognitive processes and their underlying neural networks allow some people to deal with cerebral damage resulting from stressors or neurodegenerative diseases better than others\(^23\). Variations in cognitive reserve are also related in psychiatry disorders\(^4\) like bipolar disorder\(^4\), schizophrenia\(^4\) and depression\(^4\). The present study suggests that one of the factors involved in greater cognitive reserve is formal schooling. Subjects with greater cognitive activity throughout life have a tendency to be less cognitively impaired in old age\(^8\). Such data, combined with genetic studies\(^4\), contribute to elucidating the underlying mechanisms of dementia on a epigenetic level.

The present results are consistent with other studies conducted in Brazil. Diniz et al.\(^2\) evaluated elderly community residents and found statistically significant correlations between MMSE and age (r = 0.232, p < 0.003) and education (r = 0.471, p < 0.0001) and a significant correlation linked to sex, with men having better outcomes than women. Yassuda et al.\(^3\) studied the effect of formal education in a neuropsychological assessment battery comparing groups with different educational levels and reported that formal education is correlated with most of the instruments (MMSE, CAMCOG, Trail Making Test Part B, Vocabulary and Block Design, WAIS-R and Short Cognitive Performance Test) although no significant correlation with the Verbal Fluency Test was found. In another study\(^21\) the author evaluated the relationship between memory complaints, education, depression and anxiety in elderly people and found differences in performance on the MMSE and the Clock Drawing test in groups with different schooling levels.

The development of a new tool to characterize cognitive status that is easy-to-use and publicly available is of particular interest for Brazilian neuropsychologists for use in clinical and research settings. Limitations of the current study include not controlling for socioeconomic level, the lack of a control group of elderly subjects with no cognitive complaints and the small sample size. However, it should be noted that the present analysis is of an exploratory nature. New studies are needed to establish consistent findings and for validation and neutralization purposes.

The results presented here suggest that the cognitive assessment protocol under study, besides being convenient and easy to implement, is valid to examine the various cognitive domains in the elderly and presents good psychometric properties. These findings will enhance future clinical studies on the protocol’s applicability in the Brazilian population. Given its importance, the GC construct should be investigated further and additional measures included such as: socioeconomic status, income, educational level and occupational level\(^16\).

References

17. Schultz RR. Desempenho de uma população brasileira na subescala cognitiva da Escala de Avaliação da Doença de Alzheimer (ADAS-Cog) [dissertação]. Escola Paulista de Medicina da Universidade Federal de São Paulo; São Paulo; 1999.


