Correlation between the CAMCOG, the MMSE, and three clock drawing tests in a specialized outpatient psychogeriatric service

Jeremia Heinik a,b,*, Isaac Solomesh a,b, Pinhas Berkman b,c

a Margoletz Psychogeriatric Center, Ichilov Hospital, 6 Weizman Street, Tel-Aviv 64239, Israel
b Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel
c Shmuel Harofe Hospital, Geriatric Medical Center, Beer Yakov, Israel

Received 11 March 2003; received in revised form 14 August 2003; accepted 19 August 2003

Abstract

Our objective was to assess the correlation between (1) the Cambridge Cognitive Examination (CAMCOG) (including the Mini-Mental State Examination [MMSE]) score and three clock drawing tests (CDT) and (2) the three CDTs independently, in a specialized outpatient psychogeriatric service. One hundred and fourteen subjects completed a comprehensive evaluation and were allocated to one of the following groups: dementia of the Alzheimer’s type (DAT) in 52; vascular dementia (VD) in 36; non-dementia (ND; Mood or Anxiety Disorders) in 26. When the entire sample of patients is considered, all three CDTs used were highly and significantly correlated to the MMSE score, the CAMCOG score, and to each other. In this patient population, these cognitive tests may be interchangeable for providing an initial objective measure of cognitive function. However, when the same correlations were studied in the separate diagnostic groups, in the dementia group (DAT and VD) even though the high correlations between the various CDTs themselves did not change, the correlations between the MMSE score, the CAMCOG score and the CDTs decreased, more evidently in the VD group. This trend became even more conspicuous in the ND group, where some of the above mentioned correlations became non-significant. We hypothesize that in a real clinical situation the clinician initially assumes the role of cognitive “evaluator” (in terms of the total sample) followed by the role of cognitive “monitor” (in relation to specific diagnostic groups). In the first instance, CDTs, the MMSE, and the CAMCOG might be considered interchangeable as an initial objective

* Corresponding author. Tel.: +972-3-6973606; fax: +972-3-6974658. E-mail address: heinik@post.tau.ac.il (J. Heinik).
measure of cognitive dysfunction, while in the second role, different CDTs might be diversely used, presumably supplemented by other cognitive tests and clinical methods.

Keywords: Clock drawing; Mini-Mental State Examination; Cambridge Cognitive Examination; Psychogeriatric service; Dementia of the Alzheimer’s type; Vascular dementia

1. Introduction

While it is commonly believed that the clock drawing test (CDT) appears to be, at the very least, an important adjunct to the cognitive screening process (Shulman, 2000), it is still debated at what level (general practitioner versus specialized service) the screening should be performed (Brodaty and Moore, 1997), which groups should be screened (all subjects at risk, those with subjective complaints and/or those with cognitive impairment as reported by informants) (Jorm and Jacomb, 1989; Brodaty et al., 1998), who should administer and interpret scorings (naïve raters versus professionals) and what specific scoring system is the “best” (complex versus simple) (Shulman, 2000; Tuokko et al., 2000; Storey et al., 2001; Schramm et al., 2002). The psychometric properties of all CDTs are remarkably consistent, with sensitivity and specificity levels at a mean of 85% (Scanlan et al., 2002), even though this affirmation needs to be further investigated. In fact, several investigators suggest that the accuracy of some CDTs is modest at best (Storey et al., 2001), that studies generally report a wide range of sensitivities for CDT (Tuokko et al., 2000; Storey et al., 2001; Scanlan et al., 2002; Schramm et al., 2002), that CDT’s role in the detection of dementia is more relevant in the primary context than in a specialist setting (Kirby et al., 2001), and that the many versions of CDT are not necessarily comparable or equal in utility (Tuokko et al., 2000; Royall and Espino, 2002).

Eventhough correlations alone, in this field, are concerned only with relationship without reference to other psychometric properties (e.g., sensibility, specificity etc.), correlation studies between the various clock drawing techniques and scoring systems themselves and between CDTs and other cognitive tools used for cognitive evaluation, might indicate to us whether, to what extent, and which of these tools are interchangeable as an initial measure of cognitive function. In clinical practice, the clinician at the specialized psychogeriatric ambulatory setting has many tests of cognitive impairment at his disposal (Burns et al., 1999) of which CDT is but one. Some are short and easy to score, e.g., the MMSE (Folstein et al., 1975), others are more lengthy and comprehensive; e.g., the ADAS (Rosen et al., 1984) and the CAMCOG (Roth et al., 1986). As is the case with general practitioners, specialist clinic diagnosticians are pressed for time. In addition, they also might consider the administration of one test at one stage of the assessment process, and another test at another stage based on the continuously changing clinical information gathered. Since even the most complex-to-score CDT is about a 2 min task (Burns et al., 1999), this remains a valid option in the cognitive instruments armamentarium.

Our outpatient psychogeriatric service (as many others) is prevalently concerned with dementia patients (the vast majority Alzheimer’s and/or vascular type) and sufferers of
depressive and anxiety disorders treatable on an ambulatory basis (Heunik et al., 1998). To this patient population, we routinely administer the CAMCOG (including the MMSE). CAMCOG’s clock is scored using Shulman’s method and/or Freedman’s method. A literature survey provided inconclusive data regarding correlations studies. Of the 14 CDT original scoring systems described thus far (Shulman, 2000), seven report moderate to high significant correlation between different CDTs and the MMSE score (Shulman et al., 1986, 1993; Mendez et al., 1992; Manos and Wu, 1994; Shua-Haim et al., 1997; Lam et al., 1998; Royall et al., 1998). Several replication and other studies show similar results (Ben-Yehuda et al., 1995; Bourke et al., 1995; Brodaty and Moore, 1997; Juby, 1999).

Surprisingly enough, although the CAMCOG also includes the CDT, only a few studies reported correlation (highly significant) between CDTs and the CAMCOG score (Bourke et al., 1995).

Therefore, the purpose of this study was to examine presumed correlations between two cognitive evaluation tests (the MMSE and the more complex CAMCOG) and three CDTs with increasing levels of administration and scoring complexities (CAMCOG clock, Shulman’s clock, Freedman’s clock), at a specialized psychogeriatric setting, as well as the correlations between the three CDTs.

2. Methods

One hundred and fourteen community-dwelling outpatients were included in the study. Patient evaluation and the methods used have been described elsewhere (Heunik et al., 2003) and will be briefly summarized here. Each patient underwent a comprehensive multi-disciplinary psychiatric and medical assessment. The first encounter with the patient was conducted by a geriatric psychiatrist who was also responsible for the administration of the CAMCOG (Roth et al., 1986) (including the MMSE (Folstein et al., 1975)). Hebrew versions of both tests were validated (Heunik et al., 1999; Werner et al., 1999). Laboratory investigations, including imaging studies, were conducted in each case to exclude potentially treatable causes for cognitive impairment and physical causes for the emotional disorders. Subjects met DSM-IV (American Psychiatric Association, 1994) criteria for DAT or VD. The non-dementia (ND) group included outpatients who went through the same assessment process and met the criteria for either DSM-IV Mood Disorders or Anxiety Disorders.

The CAMCOG tests clock drawing ability in its Praxis subscale. One point is given respectively for correctly drawing a circle, placement of the numbers in the correct position, and setting the hands at the correct time (11:10). The maximum score is therefore three. All the CAMCOG derived clock drawings were photocopied on clear sheets of paper and blindly scored by one of the authors (J.H.) on different occasions also using Shulman et al’s (1993) scoring system ranging from 1 (perfect) to 6 (not a reasonable representation of a clock), and Freedman et al’s (1994) scoring system for a free-drawn clock ranging from 0 (severely impaired) to 15 (not impaired). The latter outlines the following clock drawing components— contour: two points, numbers: six points, hands: six points, center: one point. Using CAMCOG derived clock drawings we had to modify both Shulman’s method (which uses a predrawn circle) and Freedman’s method (where setting the hands at 6:45 is requested). Altogether three different clock drawing scores were obtained for each subject.
Descriptively, the CAMCOG derived clock may be considered a gross, elementary scoring method. Shulman’s method is more complex, based on a hierarchical classification of errors (Shulman et al., 1986), while Freedman’s method (Freedman et al., 1994) where specific indicators for the scoring of the various clock drawing components are given, relies upon descriptors of positive responses. The statistical package SPSS-8 for Windows was used for data analysis. Pearson moment correlation and partial correlation controlling for age and education were used to determine the correlation between variables. Differences between the three groups were assessed with one-way analysis of variance. The Bonferroni correction was used to adjust P-value for multiple comparisons.

3. Results

Descriptive statistics of the demographic and cognitive characteristics of the patients in the diagnostic groups studied is shown in Table 1.

There were no significant between group differences in age, gender and education. As expected, the dementia patients significantly differed from ND patients in the MMSE score, the CAMCOG score and the three clock drawing scores. VD patients performed significantly worse than DAT patients on Freedman’s clock ($P < 0.003$) but not on CAMCOG clock and Shulman’s clock.

The MMSE score and the CAMCOG score correlated highly and significantly in the total group ($r = 0.933$, $P < 0.0001$), DAT group ($r = 0.895$, $P < 0.0001$), VD group ($r = 0.897$, $P < 0.0001$), and more moderately in the ND group ($r = 0.559$, $P < 0.003$).

Table 2 shows high and significant correlation between the MMSE score, the CAMCOG score and all three clock drawing methods found in the total sample. In the DAT and VD groups, correlations generally decrease, compared with the entire group, although 4 out of 6 remain high ($r > 0.6$) in the DAT group and 2 out of 6 in the VD group. The other correlation in these dementia groups became moderate but still very significant.

In the ND group, CAMCOG clock did not correlate with the MMSE score or with the CAMCOG score, and Shulman’s clock did not correlate with the MMSE score. In this group,

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Descriptive statistics, average (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total group ($N = 114$)</td>
<td>DAT ($N = 52$)</td>
</tr>
<tr>
<td>Age</td>
<td>77.49 (6.30)</td>
</tr>
<tr>
<td>Sex (%F)</td>
<td>49.1</td>
</tr>
<tr>
<td>Education (years)</td>
<td>10.78 (3.06)</td>
</tr>
<tr>
<td>MMSE score</td>
<td>20.47 (5.75)</td>
</tr>
<tr>
<td>CAMCOG score</td>
<td>65.57 (19.59)</td>
</tr>
<tr>
<td>CAMCOG clock</td>
<td>1.71 (0.97)</td>
</tr>
<tr>
<td>Shulman’s clock</td>
<td>3.21 (1.46)</td>
</tr>
<tr>
<td>Freedman’s clock</td>
<td>8.93 (4.48)</td>
</tr>
</tbody>
</table>

$^a$ Differences between DAT and ND groups ($P < 0.0001$).

$^b$ Differences between VD and ND groups ($P < 0.0001$).

$^c$ Differences between DAT and VD groups ($P < 0.003$).
Table 2
Correlation between MMSE and CAMCOG scores and the CDTs in the various diagnostic groups

<table>
<thead>
<tr>
<th>Diagnostic sample group</th>
<th>Cognitive measure</th>
<th>CAMCOG’s clock</th>
<th>Shulman’s clock</th>
<th>Freedman’s clock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Group (N = 114)</td>
<td>MMSE</td>
<td>0.710***</td>
<td>−0.718***</td>
<td>0.730****</td>
</tr>
<tr>
<td></td>
<td>CAMCOG</td>
<td>0.770***</td>
<td>−0.779***</td>
<td>0.795****</td>
</tr>
<tr>
<td>DAT (N = 52)</td>
<td>MMSE</td>
<td>0.657****</td>
<td>−0.563*****</td>
<td>0.574****</td>
</tr>
<tr>
<td></td>
<td>CAMCOG</td>
<td>0.714****</td>
<td>−0.632*****</td>
<td>0.651*****</td>
</tr>
<tr>
<td>VD (N = 36)</td>
<td>MMSE</td>
<td>0.438***</td>
<td>−0.582******</td>
<td>0.531***</td>
</tr>
<tr>
<td></td>
<td>CAMCOG</td>
<td>0.542***</td>
<td>−0.678******</td>
<td>0.662****</td>
</tr>
<tr>
<td>ND (N = 26)</td>
<td>MMSE</td>
<td>0.064</td>
<td>−0.208</td>
<td>0.459*</td>
</tr>
<tr>
<td></td>
<td>CAMCOG</td>
<td>0.364</td>
<td>−0.496*</td>
<td>0.568**</td>
</tr>
</tbody>
</table>

* P < 0.05
** P < 0.01
*** P < 0.001
**** P < 0.0001

only Shulman’s clock and Freedman’s clock correlated moderately with the CAMCOG score, and only Freedman’s method correlated moderately with the MMSE score. In all the diagnostic groups, the CAMCOG consistently correlated higher than the MMSE with the three clock drawing scorings studied. The subtraction of CAMCOG clock scores from the CAMCOG total score did not significantly alter the results neither in the total group nor in the separate diagnostic groups.

Table 3 shows the high and significant correlations (all P < 0.0001) between the three clock drawing tests in the total group, the DAT group and the VD group. In the ND group, a higher correlation is observed between Shulman’s clock and Freedman’s clock, compared to the more moderate correlation (all significant) found between CAMCOG clock and Shulman’s and Freedman’s clock. All correlations described remained significant after controlling for age and education.

Table 3
Correlation between three CDTs in the various diagnostic group

<table>
<thead>
<tr>
<th>Diagnostic sample group</th>
<th>Clock method</th>
<th>Shulman’s clock</th>
<th>Freedman’s clock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Group (N = 114)</td>
<td>CAMCOG’s clock</td>
<td>−0.892****</td>
<td>−0.877****</td>
</tr>
<tr>
<td></td>
<td>Shulman’s clock</td>
<td>1.000</td>
<td>−0.912****</td>
</tr>
<tr>
<td>DAT (N = 52)</td>
<td>CAMCOG’s clock</td>
<td>−0.866****</td>
<td>0.841****</td>
</tr>
<tr>
<td></td>
<td>Shulman’s clock</td>
<td>1.000</td>
<td>−0.885****</td>
</tr>
<tr>
<td>VD (N = 36)</td>
<td>CAMCOG’s clock</td>
<td>−0.848****</td>
<td>0.815****</td>
</tr>
<tr>
<td></td>
<td>Shulman’s clock</td>
<td>1.000</td>
<td>−0.884****</td>
</tr>
<tr>
<td>ND (N = 26)</td>
<td>CAMCOG’s clock</td>
<td>−0.576**</td>
<td>0.476**</td>
</tr>
<tr>
<td></td>
<td>Shulman’s clock</td>
<td>1.000</td>
<td>−0.678****</td>
</tr>
</tbody>
</table>

* P < 0.05
** P < 0.01
*** P < 0.0001
**** P < 0.0001
4. Discussion

Our results suggest that when the entire sample of patients is considered (at this stage the clinician serves as a cognitive “evaluator”), all three CDTs used were highly and significantly correlated to the MMSE score and the CAMCOG score, and very highly to each other. These cognitive tests are therefore interchangeable in this patient population when used to provide an initial measure of cognitive dysfunction, may it be the simple CAMCOG clock or the more complicated and lengthy to score Shulman’s and Freedman’s clocks. The clinician may choose one scoring system and not feel obligated to use several simultaneously. As already mentioned, some studies similarly report high ($r \geq 0.60$) correlations between the MMSE scores and various clock scorings (Shulman et al., 1986, 1993; Ben-Yehuda et al., 1995; Brodaty and Moore, 1997; Lam et al., 1998; Royall et al., 1998; Juby, 1999) at least in some of the CDTs described. The highest correlations between MMSE scores and CDT scores were found with Shulman’s method ($r = -0.77$; negative correlation reflects the scoring system used for the clock test) (Shulman et al., 1993), Mendez’ method ($r = 0.76$) (Mendez et al., 1992) and the CLOX ($r = 0.76$) (Royall et al., 1998). These authors used relatively complex scoring systems of which at least two (CLOX and Mendez) (Royall et al., 1999) tap on executive functions. Freedman’s method, that we used, might be similar to the latter two as the high correlation we found ($r = 0.73$) suggests. However, other authors do not unanimously reach these conclusions. They report only moderate correlations ($r = 0.3–0.59$) between the MMSE scores and some clock scorings (Mendez et al., 1992; Manos and Wu, 1994; Bourke et al., 1995; Brodaty and Moore, 1997; Shua-Haim et al., 1997). As for the CAMCOG, the correlations we found between CAMCOG’s scores of that sample and the three CDTs we used are somewhat higher than those described by Bourke et al. (1995) with Shulman’s method ($r = 0.70$) and Mendez’ method ($r = 0.67$).

Discussion may be similar when correlation between various CDTs is concerned. We found high correlation between the three CDTs used. In this respect, our findings are similar to Royall et al. (1999) who compared six methods of clock drawings and the correlations ranged from $-0.73$ (CLOX with Shulman) to $0.95$ (CLOX with Mendez) and to Richardson and Glass (2002) who compared five scoring protocols for the CDT and the correlations ranged from $-0.711$ to $0.913$ for the total sample. However, in Tuokko et al.’s (2000) study the correlations between the different CDTs ranged from $-0.294$ to $-0.715$.

Some of the above discrepancies might be explained on methodological grounds: different diagnostic groups, different settings (community, in-patients, out-patients, residential homes) and different examiners and specialty services.

Once the diagnosis is established, the issue of the correlation between the MMSE score, the CAMCOG score and the CDTs assumes a different perspective. This stage begins at the very moment the diagnosis is made, and the clinician assumes the role of a cognitive “monitor”. The correlations are now to be examined in the specific diagnostic groups. We found that in the specific dementia groups the correlations between the MMSE score, the CAMCOG score and the CDTs decrease. This is more evident in the VD group where four out of six correlations belong to the moderate, and less significant realm, and even more conspicuous in the ND group, where correlations between the CDTs become moderate and less significant (with one exception), and CAMCOG clock loses its correlation with the MMSE and the CAMCOG, while Shulman score loses its correlation with the MMSE.
In practical terms, this means that in the VD group, for example, the clinician may prefer to use Shulman’s clock or Freedman’s clock ($r = 0.678, 0.662$, respectively) instead of CAMCOG clock ($r = 0.438$) for rapid monitoring of the cognitive status and in the ND group Freedman’s method ($r = 0.568$), followed by Shulman’s method ($r = 0.496$). These findings are in accordance with the recent suggestion of Richardson and Glass (2002) that it may be that particular scoring methods are better suited to assess DAT than multi-infarct dementia and vice-versa.

To conclude, this simple correlations study demonstrates that the practical issue of interchangeability between cognitive tests and CDTs depends on whether the sample is considered in total or according to the specific diagnostic groups. Translated into clinical language that means that interchangeability between CDTs and the CAMCOG/MMSE might depend on the stage of assessment. In the different diagnostic groups, some CDTs may be more appropriate than others, perhaps due to some inherent properties (e.g., the presumed sensitivity of Freedman’s clock to executive dysfunction (Heinik et al., 2002)). However, the decreased correlations in the diagnostic groups suggest, that CDTs should be supplemented with other cognitive tests or clinical methods.

Several limitations need to be addressed: (a) This is a retrospective study which tacitly assumes that CAMCOG derived clock drawings and MMSE scores are equivalent to their administration separately. This still needs confirmation; (b) Our study population although grossly representative of the clinic referrals, is biased towards dementia patients (Alzheimer’s and vascular type) referred to a specialized setting. The fact that other etiologies for cognitive impairment and normals were not included, undermines the generalizability of our findings.

Acknowledgements

The authors acknowledge the assistance of Rena Kurs in preparation of the manuscript.

References


