Validation studies of the Clock Drawing Test in Mild Cognitive Impairment

Diogo Fernando Reis Carneiro[1]

[1] Faculty of Medicine, University of Coimbra, Portugal

Address: Rua Virgílio Correia, nº18, 3ºesquerdo. 3000-413 Coimbra

E-mail: diogoreiscarneiro@gmail.com

Index

Abstract

Introduction: Mild Cognitive Impairment (MCI) is a transitional entity between normal aging and Alzheimer's disease. It is assumed that an early identification and intervention in MCI may delay or slow its progression to dementia and several neuropsychological brief-tests have been investigated in this context. The Clock Drawing Test (CDT) is a widely used instrument in this field; however, its application needs further validation in specific clinical populations, mainly in milder forms of cognitive decline and in the distinction between MCI subtypes.

Objectives: To validate three scoring systems of the CDT for the detection of cognitive impairment in a cohort of MCI-patients previously classified in amnestic single-domain (aMCI) and amnestic multidomain (mdMCI) subtypes; to test inter-rater reliability and to compare different subtypes of MCI, attempting to define performance profiles according to qualitative analyses of errors.

Methods: The study includes two clinical groups: aMCI and mdMCI, each with 90 subjects, recruited at the Neurology Department of the Centro Hospitalar e Universitário de Coimbra. Their performance was compared with a cohort of 90 community-dwelling controls matched according to gender, age and education. All participants were assessed with Mini Mental State Examination, Montreal Cognitive Assessment (MoCA) and CDT. Clock drawings of MCI patients were scored by a neuropsychologist and an inexperienced rater using three scoring systems - Rouleau, Cahn and Babins. Data were analysed with the Statistical Package for the Social Sciences.

Results: There was high inter-rater reliability in CDT scoring systems (p<0,001). Significant correlations were found between the cognitive screening instruments and CDT scoring systems as well as a consistent relationship with performance in visuospatial and executive domains of the MoCA. We also observed qualitative differences between both forms of DCL, with higher error rate of "Conceptual deficit" and "Perseveration" in mdMCI, and "Nonspecific spatial error" regarding the aMCI group. There was only sufficient (60%) discriminatory capacity of total scores of the CDT, comparing control and MCI subjects.

Conclusions: Our study showed that CDT scoring systems have high inter-rater reliability to screen for MCI and can be applied in large scale studies and primary health care. Although in this context CDT revealed only sufficient discriminatory capacity and should be used with other cognitive screening tests in order to increase the diagnostic accuracy.

Key-words: Clock Drawing Test, Mild Cognitive Impairment, Reliability, Validity, Neuropsychological tests.

Resumo

Introdução: O Défice Cognitivo Ligeiro (DCL) é uma entidade transitória entre o envelhecimento normal e a demência. Considera-se que o rastreio e intervenção precoces no DCL retardam a evolução para demência e diversos testes neuropsicológicos breves têm sido propostos tendo em vista este objetivo. O Teste do Desenho do Relógio (TDR) é um instrumento vastamente utilizado neste campo, embora a sua aplicação necessite de validação suplementar em populações clínicas específicas, principalmente em formas ligeiras de declínio cognitivo e na distinção entre subtipos de DCL.

Objectivos: Validar 3 sistemas de cotação do TDR para deteção de défices cognitivos numa amostra de pacientes com DCL previamente classificados nos subtipos amnésico monodomínio (DCLa) e amnésico multidomínios (DCLmd). Testar a fiabilidade do acordo inter-avaliadores. Comparar diferentes subtipos de DCL, para definir perfis de desempenho de acordo com a análise qualitativa de erros.

Métodos: Este estudo incluiu dois grupos clínicos: DCLa e DCLmd, cada um com 90 pacientes recrutados no Serviço de Neurologia do Centro Hospitalar e Universitário de Coimbra. Os desempenhos foram comparados com uma amostra de 90 cidadãos, emparelhados de acordo com género, idade e escolaridade. Todos os participantes foram avaliados com *Mini Mental State Examination*, *Montreal Cognitive Assessment* (MoCA) e TDR. Os desenhos de relógios de pacientes com DCL foram cotados por um neuropsicólogo e um avaliador inexperiente, utilizando 3 sistemas de cotação: Rouleau, Cahn e Babins. Os dados foram analisados com recurso ao Statistical Package for the Social Sciences.

Resultados: Verificou-se elevado acordo inter-avaliadores no TDR (p<0,001). Foram encontradas correlações significativas entre os instrumentos de rastreio cognitivo e os sistemas de cotação do TDR, assim como uma relação consistente com os domínios visuoespacial e executivo do MoCA. Foram encontradas diferenças qualitativas entre as formas de DCL, com elevada taxa de erro no "Défice Conceptual" e "Preserveração" no DCLmd e no "Erro espacial não específico" no DCLa. Os resultados totais do TDR apresentaram apenas uma capacidade discriminatória suficiente (60%) na distinção entre controlos e DCL.

Conclusões: Este estudo mostrou que os sistemas de cotação do TDR têm elevada fiabilidade inter-avaliadores para rastrear DCL e podem ser aplicados em estudos de larga escala e em cuidados de saúde primários. No entanto, neste contexto, o TDR revelou uma capacidade discriminatória apenas suficiente e deve ser utilizado juntamente com outros instrumentos de rastreio cognitivo com o objetivo de aumentar a precisão diagnóstica.

Palavras-chave: Teste do Desenho do Relógio, Défice Cognitivo Ligeiro, Fidelidade, Validade, Testes Neuropsicológicos.

Introduction

Mild Cognitive Impairment (MCI) is a transitional entity between normal aging and Alzheimer's disease (AD), hence the discrimination between normal aging and pathology is frequently a difficult challenge[1-4]. It is considered both an incipient stage of dementia and a situation of risk for its development, though this progression does not always occur. About 10-15% per year and 80% over 6 years of these patients develop some type of dementia[5]. Some predictors of conversion from MCI to AD have been identified, including neuropsychological testing, neuroimaging and biomarkers, alone or in combination. It is clinically defined as a self or informant-reported cognitive complaint and an objective cognitive impairment that surpasses what is expectable in subjects with a certain age and education, while functional activities of daily living remain relatively intact.

Different subtypes of MCI are recognized. While amnestic MCI (single or multiple domain) is considered to be the prodromic stage of AD and there is controversy about the impact of this profile impairment (aMCI *vs.* mdMCI) in prognosis. Nonamnestic forms of MCI (single or multiple domain) express other types of cognitive impairment, namely language or executive functions and visuospatial abilities, representing prodromic stages of other forms of dementia, respectively Frontotemporal Dementia Vascular Dementia or Dementia with Lewy Bodies[2,3].

It is assumed that an early identification and intervention in MCI may delay or slow its progression to dementia. This identification in primary care centres depends on the development of brief cognitive screening instruments, easily applicable and properly validated. The Clock Drawing Test (CDT), included in the Montreal Cognitive Assessment (MoCA) and several neuropsychological test batteries, is a globally recognized instrument in this field, but its application requires more studies in specific clinical populations. It is already normalized

for the Portuguese population according to age and education[6], but further validation in clinical groups with MCI is still needed.

In spite of its wide usage in distinguishing normal aging from subjects with MCI, the academic literature lacks consensus on the sensitivity and specificity of the CDT as a screening tool for MCI, if used alone, contrarily to what happens in moderate and severe dementia[7]. Differences in scoring systems are also a subject of discussion. This makes the CDT of doubtful recommendation for MCI screening[8].

The aim of this study was to validate the CDT for the detection of cognitive impairment in a group of patients previously diagnosed with MCI, whose performance was compared with a cohort of cognitively normal individuals, according to age and education. We also tested inter-rater reliability and compared different subtypes of MCI, attempting to define performance profiles according to qualitative analyses of errors.

Methods

Participants

The clinical group was composed of ambulatory patients from the Neurology Department of the *Centro Hospitalar e Universitário de Coimbra* (CHUC). MCI diagnosis was previously established by the attending Neurologist according to Petersen criteria: memory complaints (preferably confirmed by an informant), preserved global cognition, activities of daily living essentially preserved, objective memory deficit according to age, and no dementia[1,9]. Clinical diagnosis was further supported by comprehensive neuropsychological assessment and extensive biochemical and image studies in order to exclude other causes for cognitive impairment. In the selection of MCI patients the following exclusion criteria were also considered: a) psychiatric comorbidities diagnosed less than six months prior to the neuropsychological study assessment; b) motor and/or sensory deficits that could represent confounding variables in the assessment of higher nervous functions; c) unstable clinical situation (e.g., recent significant deterioration).

Controls were community-dwellings integrating the CDT normative study sample[6]. In brief, they were recruited according to the following criteria:

a) Informed consent;
b) Portuguese as mother language and formal education received in Portuguese schools;
c) Normal score according to age and education on two cognitive screening instruments validated for the Portuguese population, the Mini-Mental State Examination (MMSE) and the MoCA;
d) Preserved independence and functionality;
e) No severe depressive symptomatology (Geriatric Depression Scale-30 item score ≤20);

f) No history of psychiatric, neurologic or other diseases with a negative impact on cognition;
g) No medication with a negative impact on cognition;
h) No significant motor, visual or auditory deficits with a possible negative impact on cognition;
i) No present or past history of alcoholism or drug abuse.

Instruments

For the assessment of global cognitive function we used the MMSE[10-12] and the MoCA[13,14].

The MMSE is the most widely used cognitive screening instrument for dementia. It assesses several cognitive domains with a set of questions and tasks: orientation, memory, attention and calculus, language (oral, written and reading), and constructive abilities. Each of the subtests is scored according to norms established by the authors, with a maximum of 30 points[10]. The Portuguese validation and normalization considered the significant effects of education, and cut-off scores for cognitive impairment were proposed accordingly: illiterates – ≤ 15; 1 to 11 years – ≤ 22; more than 11 years – ≤ 27[11].

The MoCA is a cognitive screening instrument created by Nasreddine[13] for the detection of MCI, as the MMSE systematically revealed a low sensitivity in patients with milder forms of cognitive impairment. The instrument has a total score of 30 points and assesses six cognitive domains through different tasks: short-term memory (delayed recall task, 5 points); visuospatial abilities (clock drawing, 3 points; cube copy, 1 point); executive functions (trail-making test B adapted, 1 point; phonemic verbal fluency, 1 point; verbal abstraction, 2 points); attention, concentration and working memory (A's detection, 1 point; serial subtractions, 3 points; digit span, 2 points); language (naming, 3 points; sentence repetition, 2 points); and

orientation (day, month, year, week day, place, and city, 6 points)[13,14]. The MoCA is normalized for the Portuguese population and there are norms according to age and education for subjects with 25 years of age and older[15].

Finally, the CDT was scored according to the three scoring systems currently validated for the Portuguese population: the Rouleau[16] 10-point quantitative system; the Cahn[17] system which combines the Rouleau quantitative score with the analysis of the 8 more common types of error that occur during clock drawing[18]; and the Babins[19] 18-point scoring system. The instructions given to the participants were the following: "*I want you to draw a round clock, place all the numbers, and set the time for ten past eleven*". The word "hands" and "minutes" should have been avoided in all cases as they could constitute hints for the execution process.

Procedures

All study participants in the MCI group were assessed by certified and experienced neuropsychologists, and the final neuropsychological diagnosis was validated by the Head of the Laboratory of Higher Nervous Functions of the CHUC. All clock drawings from both the clinical and control groups were scored by one Neuropsychologist, thus minimizing the risk of differences in CDT scores attributable to interpersonal factors. Additionally, a subgroup of protocols was scored by an independent inexperienced rater previously trained by the responsible Neuropsychologist in order to assess inter-rater reliability.

Complete methods for the selection and assessment of control subjects can be found in Santana[6].

Statistical analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS, version 19.0) and a *p* value of 0,05 was considered statistically significant. Descriptive statistics were used for sample's characterization and two-sample *t*-test as well as independence qui-square (χ^2) test allowed the group comparisons. The inter-rater reliability was calculated using the Pearson correlation coefficient between the scoring of two independent raters. The convergent validity was determined using Pearson correlations coefficients between the CDT, MoCA and MMSE scores. The group differences were examined using two-sample *t*-test, analysis of variance (ANOVA) and analysis of covariance (ANCOVA).

The diagnostic accuracy of the CDT scoring systems for the prediction of the clinical diagnosis of MCI (as well as subtypes of MCI) was assessed through the receiver operating characteristics (ROC) curve analysis. In this analysis, the area under the curve (AUC) can vary between 0.5 and 1 and a larger AUC indicates better diagnostic accuracy. The optimal cut-off points for each scoring system that yielded the highest Youden index were selected, with the higher value indicating maximization of sensitivity and specificity. For the analysis of the predictive value of this test we calculated, for each cut-off point, the sensitivity (the probability for subjects with cognitive impairment to have a positive test), specificity (the probability for subjects without cognitive impairment to have a negative test), positive predictive value (PPV, the probability of disease in subjects who have a positive test), negative predictive value (NPV, probability of the classification "lack of disease" in subjects who have a negative test) and classification accuracy (probability of correct classification of subjects with or without cognitive impairment).

Results

The final study sample was composed of 270 subjects, distributed equally amongst three groups: 90 amnestic (single domain) MCI subjects (aMCI), 90 amnestic multidomain subjects (mdMCI), and 90 control subjects (CNT). In the selection process, all participants were matched according to gender, age and education. The total sample had 141 females (52,2%), a mean age of 69,39 (±7,32) years and 7,93 (±4,54) years of education. Additionally, the participants in the MCI groups had a mean age at onset of cognitive complaints of 66,42 (±7,54) years. As expected, there were no gender distribution differences between the three groups (χ^2=0,000, p=1,000), as well as no differences in age ($F_{(2,267)}$=0,232, p=0,793), education ($F_{(2,267)}$=0,088, p=0,915), or age of onset ($t_{(137)}$=1,146, p=0,254). A complete description of the demographic variables from the study samples can be found in Table 1.

Table 1. Demographic variables

	aMCI	mdMCI	CNT	Group differences
Gender (%females)	47 (52,2%)	47 (52,2%)	47 (52,2%)	χ2= 0,000, p=1,000
Age	69,37 (7,38)	69,77 (7,47)	69,02 (7,19)	$F_{(2,267)}$= 0,232, p=0,793
Education	7,87 (4,49)	8,09 (4,71)	7,82 (4,47)	$F_{(2,267)}$= 0,088, p=0,915
Age of onset	65,68 (7,40)	67,14 (7,65)	N/A	$t_{(137)}$=1,146, p=0,254

Note: Results are presented as *mean* (*standard deviation*) except for gender.

Inter-rater reliability was assessed in a subgroup of 70 randomly selected MCI subjects that were scored by two independent raters. We found high significant correlations (Pearson's R) between total scores of both raters for the three scoring systems: Rouleau – r=0,895, $p\leq$0,001; Cahn – r=0,871, $p\leq$0,001; and Babins – r=0,897, $p\leq$0,001.

There were moderate to high significant correlations between the cognitive screening instruments and the CDT scoring systems (Table 2). As for demographic variables of interest, we found significant low negative correlations between all scoring systems and age, as well as low to moderate significant correlations with education (Table 2).

Table 2. Correlation between the MMSE, MoCA and CDT scoring systems

	Education	MMSE	MoCA	Rouleau	Cahn	Babins
Age	0,133*	-0,049	-0,107	-0,167**	-0,164**	-0,153*
MMSE	0,261***	-----	0,625***	0,320***	0,335***	0,350***
MoCA	0,416***	0,625***	-----	0,424***	0,437***	0,502***
Rouleau	0,279***	0,320***	0,424***	-----	0,979***	0,941***
Cahn	0,269***	0,335***	0,437***	0,979***	-----	0,929***
Babins	0,354***	0,350***	0,502***	0,941***	0,929***	----

*** Significant at the 0,001 level

** Significant at the 0,01 level

* Significant at the 0,05 level

We performed an ANOVA in order to analyze the differences between the groups in the MMSE, MoCA and CDT scoring systems. We found significant differences in the MMSE ($F_{(2,267)}$=13,503, p≤0,001), MoCA ($F_{(2,213)}$=26,798, p≤0,001), Rouleau ($F_{(2,267)}$=5,976, p≤0,01), Cahn ($F_{(2,266)}$=7,729, p≤0,001), and Babins scoring systems ($F_{(2,267)}$=4,278, p≤0,05). Post hoc analyses are described in Table 3.

Table 3. Group differences in MMSE, MoCA, and CDT scoring systems

	aMCI	mdMCI	CNT	Group comparisons	*p*
MMSE	27,81 (2,21)	27,18 (1,92)	28,63 (1,44)	CNT > aMCI CNT > mdMCI	0,010 0,000
MoCA	20,63 (3,76)	19,35 (3,50)	23,58 (3,43)	CNT > aMCI, mdMCI	0,000
Rouleau	8,08 (1,84)	7,29 (1,99)	8,24 (2,10)	CNT > aMCI aMCI > mdMCI	0,004 0,022
Cahn	6,97 (2,64)	5,78 (2,85)	7,34 (2,87)	CNT > mdMCI aMCI > mdMCI	0,001 0,013
Babins	13,82 (3,81)	12,87 (3,53)	14,46 (3,66)	CNT > mdMCI	0,011

Note: Results are presented as *mean* (*standard deviation*).

Considering the significant effect of age and education in all tests' scores, we performed an ANCOVA using as covariates both demographic variables. We confirmed a significant effect of diagnosis in all scoring systems, as well as MMSE and MoCA scores (Table 4).

Table 4. Analysis of covariance: effect of Diagnosis in tests' scores (covariates: age and education)

	F	Sig.	R^2	Partial Eta square	Observed power
MMSE	14,953	0,000	0,169	0,101	0,998
MoCA	37,705	0,000	0,411	0,263	1,000
Rouleau	6,776	0,001	0,163	0,049	0,917
Cahn	8,718	0,000	0,168	0,062	0,969
Babins	5,200	0,006	0,198	0,038	0,827

Afterwards we explored the differences between the groups regarding subjective errors included in the Cahn scoring system. We found significant differences in "Conceptual deficit",

with the mdMCI and CNT groups having a significantly higher error rate than the aMCI group (χ^2=13,034, p≤0,001); the mdMCI group also had a significantly higher rate of "Perseveration" errors than the other two groups (χ^2=6,067, p≤0,05); finally, regarding "Non specific spatial error", the aMCI group had a significantly higher error rate that the CNT group (χ^2=8,258, p≤0,05). There were no significant differences regarding any of the other subjective errors.

In order to analyze the relationship between the CDT and different cognitive areas, we performed correlation analyses between the three scoring systems and the cognitive domains assessed by the MoCA. We found different results according to each group, but there was a common significant relationship between the CDT, visuospatial abilities, and executive functions. The results for each group are detailed in Tables 5, 6 and 7.

Table 5. Correlation between the CDT scoring systems and MoCA cognitive domains – aMCI

	Memory	**Visuospatial abilities**	**Executive functions**	**Attention/ C/WM**	**Language**	**Orientation**
Rouleau	0,188	0,755**	0,245*	0,181	0,027	0,250*
Cahn	0,173	0,730**	0,224	0,185	0,006	0,258*
Babins	0,239*	0,804**	0,362**	0,246*	0,083	0,207

Note: Attention/C/WM (Attention, Concentration and Working Memory)

** Significant at the 0,001 level

* Significant at the 0,05 level

Table 6. Correlation between the CDT scoring systems and MoCA cognitive domains – mdMCI

	Memory	Visuospatial abilities	Executive functions	Attention/ C/WM	Language	Orientation
Rouleau	-0,179	0,723**	0,284*	0,113	0,392**	-0,224
Cahn	-0,139	0,689**	0,299*	0,103	0,347*	-0,200
Babins	-0,175	0,699**	0,305*	0,176	0,440**	-0,220

Note: Attention/C/WM (Attention, Concentration and Working Memory)

** Significant at the 0,001 level

* Significant at the 0,05 level

Table 7. Correlation between the CDT scoring systems and MoCA cognitive domains – CNT

	Memory	Visuospatial abilities	Executive functions	Attention/ C/WM	Language	Orientation
Rouleau	-0,057	0,699**	0,451**	0,337**	0,215*	0,111
Cahn	-0,057	0,731**	0,478**	0,361**	0,209*	0,106
Babins	-0,039	0,759**	0,550**	0,354**	0,284**	0,114

Note: Attention/C/WM (Attention, Concentration and Working Memory)

** Significant at the 0,001 level

* Significant at the 0,05 level

The diagnostic/discriminative capacity of the CDT was determined by using ROC curve analysis; sensitivity and specificity values, as well as the optimal cut-off points were determined by using Youden index. When distinguishing CNT from all MCI subjects, we found an AUC of 0,614, 0,623 and 0,623 for the Rouleau, Cahn and Babins scoring systems total scores, respectively (Figure 1). The optimal cut-off score for the Rouleau system was 9 points, with 86% sensitivity and 34% specificity; for the Cahn system a cut-off score of 8 points represented 55% sensitivity and 66% specificity; as for the Babins system, a cut-off score of 15 points

showed 53% sensitivity and 70% specificity (Annex 1). The PPV, NPV and diagnostic accuracy values for each scoring system are presented in Table 8.

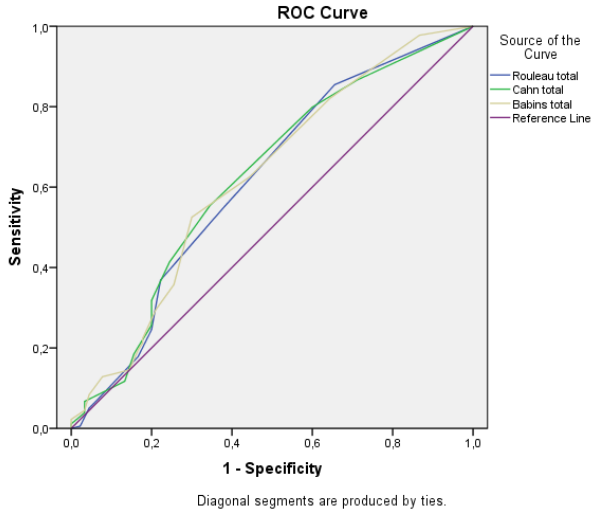

Figure 1. ROC curve analysis: CNT vs MCI.

Table 8. Sensitivity, specificity, PPV, NPV and diagnostic accuracy of the CDT: CNT vs. MCI.

		Positive	Negative	PPV/NPV	Accuracy
Rouleau	Presence	86	66	57	60%
	Absence	14	34	71	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Cahn	Presence	55	33	63	61%
	Absence	45	66	59	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Babins	Presence	53	30	64	62%
	Absence	47	70	60	

Next we examined the discriminative capacity of the CDT regarding both subtypes of MCI separately. When comparing CNT and aMCI subjects, we found an AUC of 0,556, 0,565, and 0,575 for the Rouleau, Cahn and Babins systems, respectively (Figure 2). The optimal cut-off score for the Rouleau system was 9 points, with 78% sensitivity and 34% specificity; for the Cahn system a cut-off score of 8 points represented 73% sensitivity and 40% specificity; as for the Babins system, a cut-off score of 14 points showed 43% sensitivity and 70% specificity

(Annex 2). The PPV, NPV and diagnostic accuracy values for each scoring system are presented in Table 9.

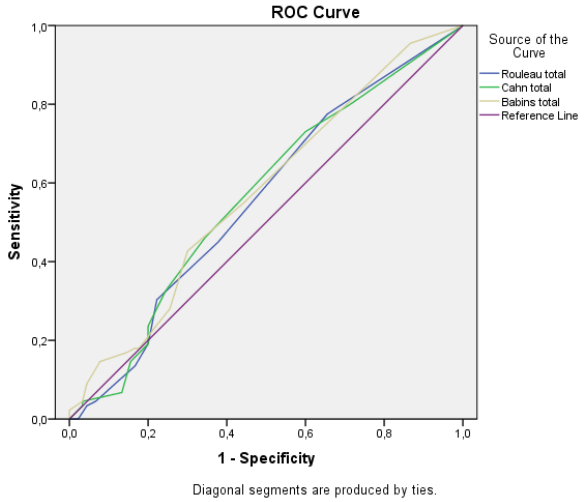

Figure 2. ROC curve analysis: CNT vs aMCI.

Table 9. Sensitivity, specificity, PPV, NPV and diagnostic accuracy of the CDT: CNT vs. aMCI.

		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Rouleau	Presence	78	66	54	56%
	Absence	22	34	61	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Cahn	Presence	73	60	55	57%
	Absence	27	40	60	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Babins	Presence	43	30	59	57%
	Absence	57	70	55	

When comparing CNT and mdMCI subjects, we found an AUC of 0,672, 0,680, and 0,671 for the Rouleau, Cahn and Babins systems, respectively (Figure 3). The optimal cut-off score for the Rouleau system was 9 points, with 93% sensitivity and 34,4% specificity; for the Cahn system a cut-off score of 8 points represented 64% sensitivity and 66% specificity; as for the Babins system, a cut-off score of 15 points showed 62% sensitivity and 70% specificity

(Annex 3). The PPV, NPV and diagnostic accuracy values for each scoring system are presented in Table 10.

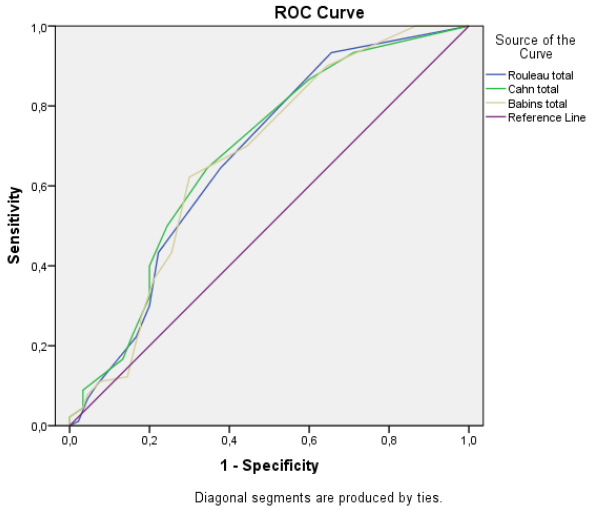

Figure 3. ROC curve analysis: CNT vs. mdMCI.

Table 10. Sensitivity, specificity, PPV, NPV and diagnostic accuracy of the CDT: CNT vs. mdMCI.

		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Rouleau	Presence	93	66	58	64%
	Absence	7	34	83	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Cahn	Presence	64	34	65	65%
	Absence	36	66	65	
		Positive	**Negative**	**PPV/NPV**	**Accuracy**
Babins	Presence	62	30	67	66%
	Absence	38	70	65	

Discussion and Conclusions

MCI is still now a controversial entity, considered by some authors as a situation of high risk for the development of dementia and AD, and assumed by others as prodromic AD. This ambivalence is reflected in the most recent diagnostic criteria: the NIA-AAWG (National Institute on Aging- Alzheimer´s disease Work Group[4]), considers that the MCI-group with a biomarker-profile of AD is still in a pre-dementia high risk stage and propose the designation of "MCI due to AD" for these patients. According to the recent European proposal - Advancing research diagnostic criteria for Alzheimer's disease: IWG-2 criteria[20], MCI-patients presenting evidence of tau and amyloid pathology or harbouring an AD-related mutation on chromosomes 1, 14 or 21, are already classified as Typical AD. There is also controversy about the impact of the profile impairment (aMCI *vs.* mdMCI) in prognosis. Some studies indicate that the most relevant neuropsychological predictor of conversion to dementia is the severity of the memory impairment[21], while others consider a multidomain involvement as the stronger indicator of conversion[22].

We present a validation study of the CDT in specific groups of MCI patients (amnestic *vs.* amnestic/multidomain), whose performance was compared with a control group selected from the Portuguese normative study population[6]. This study had a large sample, almost perfectly matched between the three groups regarding major demographic variables (gender, age and education), which increased the reliability of results. Considering the study groups, they can be deemed as representative because demographic characteristics are equivalent to previously published cohorts, namely in terms of median age and female gender dominance. Educational level, although lower than usually observed in cohorts from USA or northern Europe, is typical of our geriatric population.

One aim of this study was to test inter-rater reliability. We verified parallel results between an experienced and a non-experienced rater, a fact that supports the reliability of quotation systems and their application in clinical practice by health professionals with no major experience in neuropsychology. As a brief and easily applicable test, the CDT can be safely used in large scale, mainly in primary care centres as an additional instrument for the screening of cognitive impairment. Previous studies also showed that most of the existing CDT scoring systems have inter-rater reliability above 90% to screen for moderate and severe dementia[7,23] and similar values for mild Alzheimer's Disease[24,25]. Our results expanded these conclusions to cohorts of patients with milder forms of cognitive decline.

All three scoring systems presented moderate significant correlations with other cognitive screening instruments analysed. This relationship was stronger with the MoCA, which is explained by the recruitment of multiple brain areas or cognitive functions while performing the CDT, mainly executive functions, which are not assessed by the MMSE. Similar to our results, Mainland and collaborators[26] also found significant correlations between the CDT and the MMSE, using different quotation systems (e.g., Shulman and Babins), although with little impact in the MCI-diagnosis.

The MoCA is generally considered more sensitive and specific in screening dementia or MCI than the MMSE[27] and the same results were found in studies concerning Portuguese population[15]. In addition, when comparing MoCA, MMSE and the CDT, the first one proved to be the most appropriate tool to screen MCI and in differentiating it from AD and from normal ageing[28]. Our study corroborated these results, confirming the superiority of the MoCA relatively to the CDT and the MMSE.

Similar to other instruments of neuropsychological assessment, in this study the CDT showed significant correlations with demographic variables of interest, particularly education. However, since this effect was controlled (together with age), the variable diagnosis (CNT *vs*.

aMCI *vs*. mdMCI) maintained a significant effect on the test results. Age and education have been proved to be very important variables with a significant effect on cognitive screening instruments[29,30]. It is known that lower education can even compromise the results on the CDT as is reflected in very low normative scores in individuals with less than 4 years of formal education[31]. For the Portuguese population, cut-off points according to age and education were recently defined which allowed a more accurate use of the CDT as a brief cognitive screen test[6]. Nonetheless, carefulness it is still suggested when interpreting CDT results in elder and people with low education based exclusively in global quantitative scores.

Concerning differences between the analysed groups, we verified an overall trend towards a superiority of the performance of control subjects with respect to MCI individuals. However, when we analysed aMCI and mdMCI separately, the mdMCI group showed a significantly poorer performance than the aMCI group, a difference that was not detected by standard screening instruments (MMSE and MoCA). Previous neuropsychological studies compared the performance of CDT between normal population, MCI and AD[32], but not between different subtypes of MCI. Few studies have attempted to distinguish subtypes of MCI patients, although they used more developed instruments or techniques, such as Event Related Potentials in the Simon Task[33]. The Cahn scoring system, as a qualitative measure, was important to identify the pattern of errors made by each subtype of MCI, with "Conceptual Deficit" and "Perseveration" more typical of mdMCI, and "Nonspecific spatial error" with a higher error rate committed by aMCI patients.

Differences between MCI subtypes were further enhanced by the correlation with several cognitive domains (sets of specific MoCA subtests). As was previously described[34], we confirmed the importance of executive functions and visuospatial abilities in the CDT performance. This correlation with visual and executive functions is shared by the three groups.

Other results also appeared interesting: first, in CNT group there was a high correlation with the attention domain, including concentration and working memory, which did not happen in any MCI subtype; second, correlation with language domain is shared with MCI-md; third, orientation skills and CDT were only significantly correlated in the aMCI group. We do not have a global explanation for these findings, but we can speculate that these are critical domains in the respective groups: attention and working memory are cognitive functions specially vulnerable to normal-aging, while functions specially vulnerable to AD-pathology like language, may be correlated in study groups. On the other hand, there were no significant correlations with memory performance which is a clear demonstration of the independence of the CDT relatively to this function.

ROC curve analysis demonstrated a poor discriminative capacity of total scores of the CDT between control and MCI subjects. This is specially true but also predictable for the amnesic sub-type considering that there were no correlations between memory and CDT performance. This evidence reinforces the hypothesis that this instrument, when used individually and based on overall scores, will have a limited clinical interest in the detection of typical MCI due to Alzheimer's disease. Despite the different sensitivity and specificity values between the three systems, their diagnostic accuracy is very similar, slightly higher for the Babins system. According to the literature, the individual use of CDT has good screening precision in moderate and severe dementia, but not in MCI[7,8,35].

For the screening of milder cognitive decline, it is suggested that the CDT may be applied as part of larger assessment protocols[25], and/or in association with other cognitive screening tests like the MMSE[23,36], but this association is still under development and more studies are needed to demonstrate its efficacy.

Acknowledgments

The realization of this project would not be possible without the precious help of some people, to whom I would like to leave my acknowledgement.

First, I deeply thank Professor Isabel Santana, Professor of Neurology of Faculty of Medicine of University of Coimbra and my master in this project. My acknowledgment is for her proposal to make this study, her bibliographic orientation and revision of all texts of this dissertation.

Secondly, I thank Dr. Diana Duro who worked on statistical data and revised the work several times.

A last but most important word for the patients whose data is embodied in this study. It is for them that we will keep on working.

References

1. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. (1999) Mild cognitive impairment: clinical characterization and outcome. Arch Neurol; 56:303-8.

2. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B. (2001) Current Concepts in Mild Cognitive Impairment. Arch Neurol; 58(12):1985-1992.

3. Petersen RC. (2003) Conceptual Overview. In: Mild Cognitive Impairment: Aging to Alzheimer's disease. Oxford University Press. Portuguese Edition: Defeito cognitivo ligeiro: o envelhecimento e a doença de Alzheimer, Lisboa, Climepsi Editores, 2004.

4. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement; 7(3):270–279.

5. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Smith GE, Jack CR. (2009) Mild Cognitive Impairment: Ten Years Later. Arch Neurol; 66(12):1447-1455.

6. Santana I, Duro D, Freitas S, Alves L, Simões MR. (2013) The Clock Drawing Test: Portuguese norms, by age and education, for three different scoring systems. Arch Clin Neuropsychol; 28(4):375–87.

7. Pinto E, Peters R. (2009) Literature review of the Clock Drawing Test as a tool for cognitive screening. Dement Geriatr Cogn Disord; 27(3):201–13.

8. Ehreke L, Luppa M, Konig KK, Riedel-Heller SG. (2010) Is the Clock Drawing Test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. International Psychogeriatrics; 22(1):56–63.

9. Petersen RC. (2007) Mild Cognitive Impairment. In: Continuum Lifelong Learning Neurol; 13(2): 15-38.

10. Folstein M, Folstein S, McHugh P. (1975) Mini mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research; 12:189–198.

11. Guerreiro M, Silva AP, Botelho M, Leitão O, Castro-Caldas A, Garcia C. (1994) Adaptação à população portuguesa da tradução do "Mini Mental State Examination" (MMSE). Revista Portuguesa de Neurologia; 1, 9-10.

12. Guerreiro, M., Silva, A. P., et al. (2008) Avaliação Breve do Estado Mental. In Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), Escalas e Testes na Demência (2ª ed.) (pp. 31-39). Lisboa: Novartis.

13. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. J Am Geriatr Soc; 53(4):695-699.

14. Simões MR, Freitas S, Santana I, Firmino H, Martins C, Nasreddine Z, Vilar M. (2008) Montreal Cognitive Assessment (MoCA): Versão final portuguesa. Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. Coimbra.

15. Freitas S, Simões MR, Alves L, Santana I. (2011) Montreal Cognitive Assessment (MoCA): normative study for the Portuguese population. J Clin Exp Neuropsychol; 33(9):989-996.

16. Rouleau I, Salmon DP, Butters N, Kennedy C, McGuire K. (1992) Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. Brain Cogn; 18(1):70-87.

17. Cahn DA, Salmon DP, Monsch, AU, Butters N, Wiederholt WC, Corey-Bloom J. (1996) Screening for dementia of the Alzheimer type in the community: the utility of the clock drawing test. Arch Clin Neuropsychol; 11:529–539.

18. Freedman MI, Leach L, Kaplan E, Winocur G, Shulman KJ, Delis DC. (1994) Clock drawing: A neuropsychological analysis. Oxford: Oxford University Press.

19. Babins L, Slater M-E, Whitehead V, Chertkow H. (2008) Can an 18-point clock-drawing scoring system predict dementia in elderly individuals with mild cognitive impairment? J Clin Exp Neuropsychol; 30:173–186.

20. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, DeKosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC,

Galasko D, Habert MO, Jicha GA, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL. (2014) Advancing research diagnostic criteria for Alzheimer's Disease: the IWG-2 criteria. Lancet Neurol; 13: 614–29.

21. Silva D, Guerreiro M, Maroco J, Cardoso S, Santana I, Rodrigues A, Mendonça A. (2013) Prediction of long-term (5 years) conversion to dementia using neuropsychological tests. J Alzheimers Dis; 34(3):681-9.

22. Peraita H, García-Herranz S, Díaz-Mardomingo C. (2011) Evolution of specific cognitive subprofiles of mild cognitive impairment in a three-year longitudinal study. Curr Aging Sci; 4(2):171-82.

23. Shulman KI. (2000) Clock-drawing: is it the ideal cognitive screening test? Int J Geriat Psychiatry; 15: 548–561.

24. Jørgensen K, Kristensen MK, Waldemar G, Vogel A. (2014) The six-item Clock Drawing Test - reliability and validity in mild Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn; 30:1-11.

25. Aprahamian I, Martinelli JE, Neri AL, Yassuda MS. (2009) The Clock Drawing Test: a review of its accuracy in screening for dementia. Dement Neuropsychol; 3(2):74-81.

26. Mainland BJ, Amodeo S, Shulman, KI. (2013) Multiple clock drawing scoring systems: simpler is better. Int J Geriatr Psychiatry; 29:127-136.

27. Damian AM, Jacobson SA, Hentz JG, Belden CM, Shill HA, Sabbagh MN, Caviness JN, Adler CH. (2011) The Montreal Cognitive Assessment and the Mini-Mental State Examination as Screening Instruments for Cognitive Impairment: Item Analyses and Threshold Scores. Dement Geriatr Cogn Disord; 31:126-131

28. Martinelli JE, Cecato JF, Bartholomeu D, Montiel JM. (2014) Comparison of the diagnostic accuracy of neuropsychological tests in differentiating Alzheimer's disease from mild cognitive impairment: can the montreal cognitive assessment be better than the cambridge cognitive examination? Dement Geriatr Cogn Disord Extra; 4(2):113–21.

29. Bozikas VP, Giazkoulidou A, Hatzigeorgiadou M, Karavatos A, Kosmidis MH. (2008) Do age and education contribute to performance on the clock drawing test? Normative data for the Greek population. J Clin Exp Neuropsychol; 30(2):199-203.

30. Von Gunten A, Ostos-Wiechetek M, Brull J, Vaudaux-Pisquem I, Cattin S, Duc R. (2008) Clock-Drawing Test Performance in the Normal Elderly and Its Dependence on Age and Education. Eur Neurol; 60:73-78.

31. Lourenço RA, Ribeiro-Filho ST, Moreira IFH, Paradela EMP, Miranda AS. (2008) The Clock Drawing Test: performance among elderly with low educational level. Revista Brasileira de Psiquiatria; 30(4):309-315.

32. Paula JJ, Miranda DM, Moraes EN, Malloy-Diniz LF. (2013). Mapping the clockworks: what does the Clock Drawing Test assess in normal and pathological aging? Arquivos de Neuro-Psiquiatria; 71(10):763–8.

33. Cespón J, Galdo-Alvarez S, Diaz F. (2013) Electrophysiological correlates of amnestic mild cognitive impairment in a simon task. PloS One; 8(12):e81506.

34. Silva D, Mendonça A, Guerreiro M. (2009) The Clock Drawing Test: historical notes followed by a few examples. Sinapse; 9(2):52-57.

35. Nishiwaki Y, Breeze E, Smeeth L, Bulpitt CJ, Peters R, Fletcher AE. (2004) Validity of the Clock-Drawing Test as a screening tool for cognitive impairment in the elderly. American Journal of Epidemiology; 160(8):797–807.

36. Cacho J, Benito-León J, García-García R, Fernández-Calvo B, Vicente-Villardón JL,Mitchell AJ. (2010) Does the combination of the MMSE and Clock Drawing Test (Mini-clock) improve the detection of mild Alzheimer's disease and mild cognitive impairment? J Alzheimers Dis; 22:889-96.

Annex 1 – ROC curve analysis: CNT vs. MCI

CNT vs MCI				
Test Result Variable(s)	Coordinates of the Curve Positive if Less Than or Equal To	Sensitivity	1 - Specificity	**Youden**
Rouleau	1,00	0,000	0,000	0,000
	2,50	0,006	0,022	-0,017
	3,50	0,050	0,044	0,006
	4,50	0,073	0,067	0,006
	5,50	0,179	0,167	0,012
	6,50	0,246	0,200	0,046
	7,50	0,369	0,222	0,146
	8,50	0,547	0,378	0,170
	9,50	0,855	0,656	0,199
	11,00	1,000	1,000	0,000
Cahn	-3,00	0,000	0,000	0,000
	-1,50	0,006	0,000	0,006
	-0,50	0,011	0,000	0,011
	0,50	0,039	0,033	0,006
	1,50	0,067	0,033	0,034
	2,50	0,117	0,133	-0,016
	3,50	0,184	0,156	0,029
	4,50	0,257	0,200	0,057
	5,50	0,318	0,200	0,118
	6,50	0,413	0,244	0,169
	7,50	0,553	0,344	0,209
	8,50	0,799	0,600	0,199
	9,50	0,866	0,711	0,155
	11,00	1,000	1,000	0,000
Babins	2,00	0,000	0,000	0,000
	3,50	0,006	0,000	0,006
	4,50	0,022	0,000	0,022
	5,50	0,045	0,033	0,011
	6,50	0,084	0,044	0,039
	7,50	0,128	0,078	0,051
	8,50	0,145	0,144	0,001
	9,50	0,196	0,167	0,029
	10,50	0,212	0,178	0,035
	11,50	0,223	0,178	0,046
	12,50	0,296	0,211	0,085
	13,50	0,358	0,256	0,102
	14,50	0,525	0,300	0,225
	15,50	0,626	0,444	0,181
	16,50	0,821	0,644	0,177
	17,50	0,978	0,867	0,111
	19,00	1,000	1,000	0,000

Annex 2 – ROC curve analysis: CNT vs. aMCI

CNT vs aMCI				
Test Result Variable(s)	Coordinates of the Curve Positive if Less Than or Equal To	Sensitivity	1 - Specificity	**Youden**
Rouleau	1,00	0,000	0,000	0,000
	2,50	0,000	0,022	-0,022
	3,50	0,034	0,044	-0,011
	4,50	0,045	0,067	-0,022
	5,50	0,135	0,167	-0,032
	6,50	0,191	0,200	-0,009
	7,50	0,303	0,222	0,081
	8,50	0,449	0,378	0,072
	9,50	0,775	0,656	0,120
	11,00	1,000	1,000	0,000
Cahn	-1,00	0,000	0,000	0,000
	0,50	0,034	0,033	0,000
	1,50	0,045	0,033	0,012
	2,50	0,067	0,133	-0,066
	3,50	0,146	0,156	-0,009
	4,50	0,191	0,200	-0,009
	5,50	0,236	0,200	0,036
	6,50	0,326	0,244	0,081
	7,50	0,461	0,344	0,116
	8,50	0,730	0,600	0,130
	9,50	0,798	0,711	0,087
	11,00	1,000	1,000	0,000
Babins	3,00	0,000	0,000	0,000
	4,50	0,022	0,000	0,022
	5,50	0,045	0,033	0,012
	6,50	0,090	0,044	0,045
	7,50	0,146	0,078	0,068
	8,50	0,169	0,144	0,024
	9,50	0,180	0,167	0,013
	11,00	0,180	0,178	0,002
	12,50	0,225	0,211	0,014
	13,50	0,281	0,256	0,025
	14,50	0,427	0,300	0,127
	15,50	0,551	0,444	0,106
	16,50	0,742	0,644	0,097
	17,50	0,955	0,867	0,088
	19,00	1,000	1,000	0,000

Annex 3 – ROC curve analysis: CNT vs. mdMCI

CNT vs mdMCI				
Test Result Variable(s)	Coordinates of the Curve Positive if Less Than or Equal To	Sensitivity	1 - Specificity	**Youden**
Rouleau	1,00	0,000	0,000	0,000
	2,50	0,011	0,022	-0,011
	3,50	0,067	0,044	0,022
	4,50	0,100	0,067	0,033
	5,50	0,222	0,167	0,056
	6,50	0,300	0,200	0,100
	7,50	0,433	0,222	0,211
	8,50	0,644	0,378	0,267
	9,50	0,933	0,656	0,278
	11,00	1,000	1,000	0,000
Cahn	-3,00	0,000	0,000	0,000
	-1,50	0,011	0,000	0,011
	-0,50	0,022	0,000	0,022
	0,50	0,044	0,033	0,011
	1,50	0,089	0,033	0,056
	2,50	0,167	0,133	0,033
	3,50	0,222	0,156	0,067
	4,50	0,322	0,200	0,122
	5,50	0,400	0,200	0,200
	6,50	0,500	0,244	0,256
	7,50	0,644	0,344	0,300
	8,50	0,867	0,600	0,267
	9,50	0,933	0,711	0,222
	11,00	1,000	1,000	0,000
Babins	2,00	0,000	0,000	0,000
	3,50	0,011	0,000	0,011
	4,50	0,022	0,000	0,022
	5,50	0,044	0,033	0,011
	6,50	0,078	0,044	0,033
	7,50	0,111	0,078	0,033
	8,50	0,122	0,144	-0,022
	9,50	0,211	0,167	0,044
	10,50	0,244	0,178	0,067
	11,50	0,267	0,178	0,089
	12,50	0,367	0,211	0,156
	13,50	0,433	0,256	0,178
	14,50	0,622	0,300	0,322
	15,50	0,700	0,444	0,256
	16,50	0,900	0,644	0,256
	17,50	1,000	0,867	0,133
	19,00	1,000	1,000	0,000