Employing Machine Learning in the development of a self-administered, computerised cognitive assessment battery for the assessment of neurodegeneration

Dr Chris Kalafatis, Consultant in Old Age Psychiatry, South London & Maudsley NHS Foundation Trust, King's College London, UK
Dr Seyed Khaligh-Razavi, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

We employed Machine Learning principles to develop a novel, computerised cognitive assessment test that aims at screening for cognitive impairment in a way that can simplify and accelerate that diagnosis of Mild Cognitive Impairment (MCI) and Dementias. The test is primarily focuses on rapid categorisation of natural images of varied statistical properties.

Background

Screening for neurodegeneration-related cognitive impairment in non-specialist environments is disadvantaged as cognitive tests tend to be crude with low sensitivity for milder but clinically significant cognitive deficits while the more accurate batteries take too long to administer.

The Integrated Cognitive Assessment (ICA) is self-administered, very short in duration i.e.. approximately 5mins and it is independent of culture, language, education while it allows for high-resolution monitoring.

Methods

We conducted four different experiments, as summarized in Table 1. The first three experiments were designed to measure the ICA correlation with a wide range of routinely used reference cognitive tests. The goal was to investigate whether the speed and accuracy of visual processing in a rapid visual categorization task is correlated with subject's cognitive performance.

In the first and second experiments, we tested ICA's ability in assessing cognitive performance in older adults. Therefore, we used MoCA and/or ACE-R as reference cognitive tests, both of which are routinely used to screen for Mild Cognitive Impairment (MCI) and dementia in older adults.

In the first experiment, 212 volunteers participated; MoCA was used as the reference cognitive test. The second experiment included 58 participants; the ICA was delivered and both MoCA and ACE-R were used as reference tests in this experiment.

The third experiment had SDMT, BVMT-R and CVLT-II as the reference cognitive tests, measuring speed of information processing, visuo-spatial memory and verbal learning, respectively. These three tests together form the BICAMS battery, which requires about 15 to 20 minutes to administer, and is primarily used to detect cognitive dysfunction in younger adults who suffer from Multiple Sclerosis (MS). 166 participants took part in this experiment. Forty-four of them were selected for a re-test as part of a second visit to assess ICA test-retest reliability.
All the pen-and-paper cognitive tests were administered by a healthcare professional. The administration order for ICA vs. reference cognitive tests was randomised.

Finally, experiment 4 was designed to study whether the ICA test had a learning bias if taken multiple times in short intervals. 12 young volunteers participated in this study. For convenience, the ICA was delivered remotely via a web platform. Participants took the ICA test every other day for two weeks.

Experiment	Number of Participants	Age mean years, SD, [min, max]	Education mean years, SD, [min, max]	Gender (#female)	Cognitive Tests	ICA Platform
1	212	74,10, [46, 98]	9, 6, [0, 23]	110 (51%)	MoCA, ICA	Raspberry Pi (RaPi)
2	58	62, 6, [54, 79]	14, 5, [3, 24]	33 (56%)	MoCA, ACE-R, ICA	iPad
3	166	37,10, [19, 65]	14, 3, [1, 20]	125 (75%)	SDMT, BVMT-R, CVLT-II, ICA	iPad
3' (re-test)	44	38, 12 [18, 64]	14, 2 [8, 20]	29 (66%)	ICA, SDMT	iPad
4	12	29, 3, [20, 36]	19, 4, [15, 24]	5 (41%)	ICA	Web

Table 1: Summary of all the experiments

Results

.Here in three different experiments (see Table 1, experiment 1 to 3), we show that the ICA test is significantly correlated with six standard neuropsychological tests (Figure 2 and Table 2).

We show that the ICA score is significantly correlated with MoCA, tested on two different hardware platforms (RaPi and iPad). ICA correlation with MoCA varies from 0.46 and 0.55 (Figure 2D and 2E) that is within the range for determining construct validity.

The ICA test had a slightly higher correlation with ACE-R (r=0.60, p <10-6), compared to MoCA. ACE-R provides a more comprehensive assessment of cognitive abilities and takes a longer time to administer and score (~ 30 minutes. The ICA correlation with ACE-R (Figure 2F) and its different sub-sections are shown in Table 2. Subjects' MMSE scores can also be extracted from the ACE-R test (see Table 2).

In addition to these, we compared ICA against another set of tests, including SDMT, BVMT-R and CVLT-II (Figure 2A, 2B, 2C, and Table 2) that are more often used in younger individuals to assess cognitive performance. For example, all these three tests are included as part of a larger battery of tests that assess cognitive impairment in individuals with MS, such as the 'minimal assessment of cognitive function in MS (MACFIMS) battery and the 'brief international cognitive assessment for MS' (BICAMS).

ICA had the highest correlation with SDMT (r=0.80, p<10-7), which measures the speed of information processing. CVLT-II, measuring verbal learning, and BVMT-R, measuring visual memory, had correlations of 0.66 and 0.54 with the ICA score, respectively. It is worth noting that a correlation of one is not desirable between the ICA test and any of these cognitive tests, as none of these standard tests are considered the ground truth (or gold standard) in detecting cognitive impairment.
The majority of cognitive tests (Table 2) were more correlated with the accuracy component of the ICA test, except for SDMT and CVLT-II, both of which have got a significantly higher correlation with speed compared to that of accuracy (p<0.001; bootstrap resampling of subjects).

Each reference cognitive test used in this study (shown in Table 2) measured different domains of cognition. The ICA score had a significant correlation with all of these tests, suggesting that it can be effectively used as one integrated test to provide insights about different cognitive domains (e.g. speed of processing, memory, verbal learning, attention, and fluency).

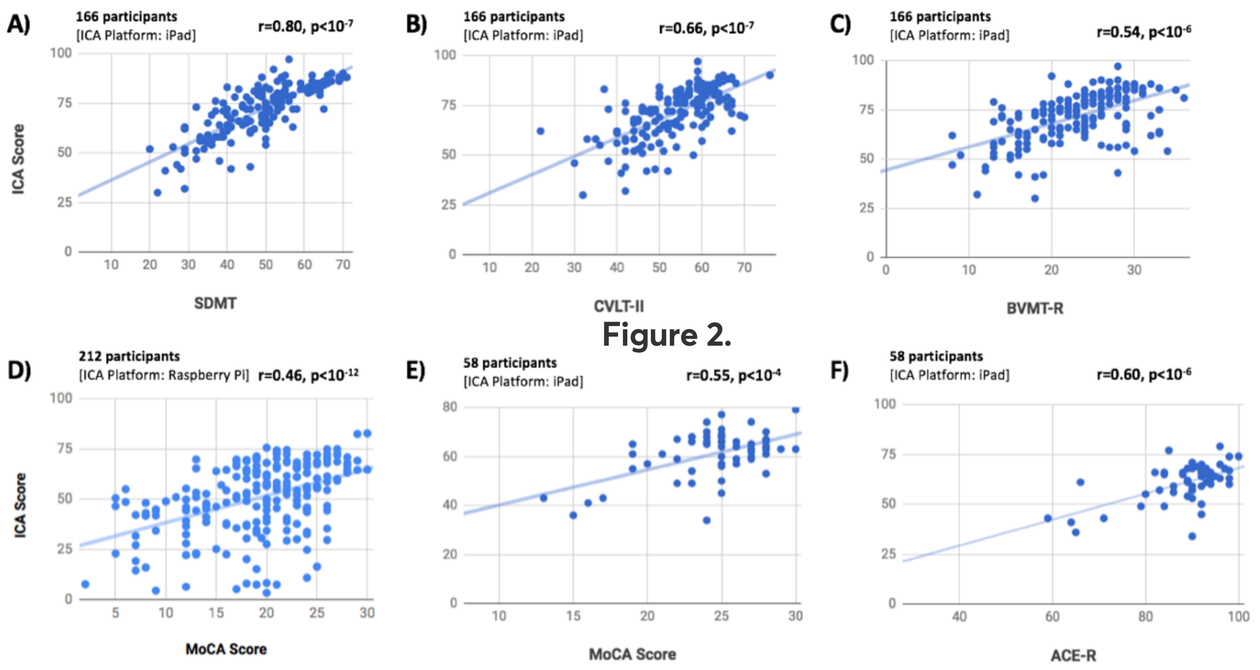

Figure 2.

To assess the reliability of the ICA test, a subgroup of 44 participants from experiment 3 (see Table 1) took the ICA test for the second time after about five weeks (+-15 days). Test-retest reliability was measured by computing the Pearson correlation between the two ICA scores [Pearson's r = 0.96 (p<10-7)]. R values for test-retest correlation are considered adequate if >0.70 and good if >0.80 (Anastasi, 1988).

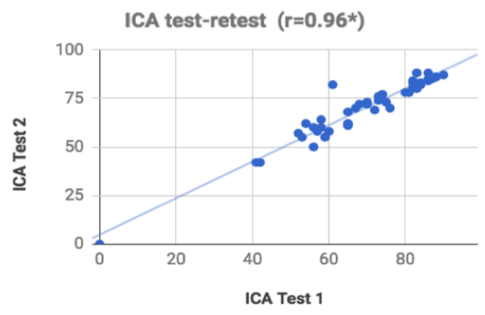

Figure 1.
ICA score, test-retest scatter plot.

Each blue dot shows the ICA score for an individual taken on two different days. The blue line indicates the linear curve fitted to the test-retest data. [Pearson's r = 0.96 (*p<10-7)]

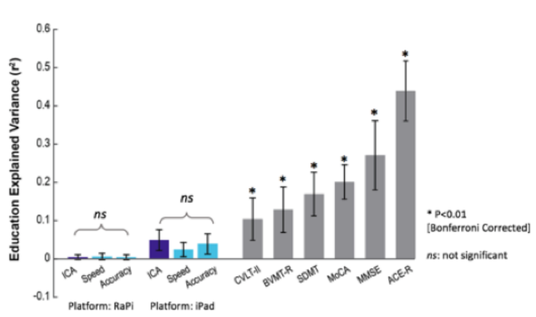

Figure 3.
Dependency of standard-of-care cognitive tests on education.

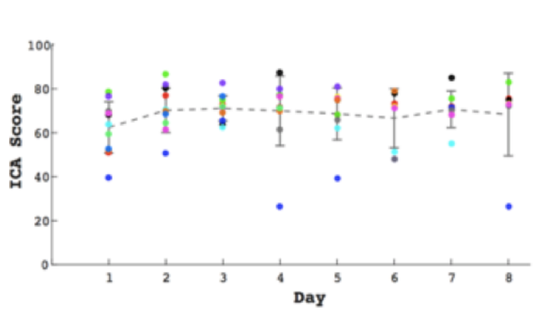

Figure 4.
No significant effect of learning in repeated exposure to the ICA test.

We find no learning bias when the test is taken multiple times. 12 healthy participants (age range = [20,36]) took the ICA test every other day for over two weeks (ANOVA, F(7) = 0.62; P-value = 0.73). From these 12 participants, 7 of them completed all the sessions (8 days); and the rest did the test for at least the first three days.

Cognitive Test	Correlation with			Cognitive Domain
	ICA	Accuracy	Speed	
SDMT	0.80, p<10-7	0.41, p<10-7	0.71, p<10-7	speed of processing
CVLT-II	0.66, p<10-7	0.40, p<10-6	0.56, p<10-14	verbal learning
BVMT-R	0.54, p<10-6	0.40, p<10-8	0.43, p<10-7	visual memory
MoCA (1) -RaPi	0.46, p<10-11	0.48, p<10-12	0.19, p<0.01	general
MoCA (2) -iPad	0.55, p<10-4	0.56, p<10-5	0.19, ns	general
ACE-R total score	0.60, p<10-6	0.52, p<10-4	0.28, p<0.05	total score of 5 domains
ACE | Attention	0.27, p<0.05	0.25, ns	0.11, ns	attention
ACE | Memory	0.47, p<10-3	0.37, p<0.01	0.26, p<0.05	verbal memory
ACE | Fluency	0.45, p<10-3	0.25 ns	0.35, p<0.01	fluency
ACE | Language	0.53, p<10-4	0.62, p<10-6	0.12, ns	language
ACE | Visuospatial	0.42, p<10-3	0.44, p<10-3	0.12, ns	visuospatial
MMSE	0.33, p<0.01	0.36, p<0.01	0.08, ns	general

Table 2.
Correlation of the ICA test scores with various domains of cognition

Furthermore, we formally tested whether the ICA score is independent of education using a non-parametric test of independence (Gretton and L. Gyorfi, 2008). In experiment 1 (i.e. ICA taken on RaPi), the statistical test of independence was positive, showing that ICA score is independent of education (based on 10,000 bootstrap resampling of subjects).

The test data indicate that even in subjects with a high capacity to learn, no learning bias was detected (Figure 4). The ICA score does not increase monotonically, and comparing the mean of the ICA scores across these days no significant difference was observed (ANOVA, F(7) = 0.62; P-value = 0.73).

Conclusions

We established that the Integrated Cognitive Assessment correlation with MoCA is significant to determine construct validity. The ICA is independent of language and education while there is no discernible learning effect. The above attributes will yield more significant clinical benefits in the day-to-day identification of neurodegeneration both in primary care and specialist clinics.