

# The reliability of the MMSE and MoCA cognitive screening tools in the population aged 55 years and older

George M Savva<sup>1</sup>, Claire O Regan<sup>2</sup>, Bellinda King-Kallimanis<sup>2</sup>, Hilary Cronin<sup>2</sup>, Rose Anne Kenny<sup>2</sup>

<sup>1</sup> School of Health Sciences, University of East Anglia, Norwich, Norfolk, NR4 7TJ, United Kingdom. Contact: g.savva@uea.ac.uk

<sup>2</sup> The Irish Longitudinal Study on Ageing, Trinity College Dublin, Republic of Ireland



We present estimates of the reliability and minimum detectable change of the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA).

## Background

Measures of cognitive function are not perfectly reliable and it is important to know how likely it is that an observed change in an individual is due to chance

In research settings unreliable measures can lead to biased estimates of effects, and correcting for these requires estimates of the standard error of measurement.

Here we use repeat assessment data from a representative sample of the Irish population aged 55 and older to report the standard error of measurement of MMSE and MoCA, their intra-class correlation and minimum detectable change.

We also estimate inter-rater reliability and the effect of the time between assessments.

## Methods

The Survey for Health, Ageing and Retirement in Europe (SHARE) / Irish Longitudinal Study on Ageing (TILDA) collaboration was established to determine the measurement properties of cognitive and physical health measures among nationally representative samples of the European population.

A random sample of the SHARE-Ireland cohort underwent a health assessment including the MMSE and MoCA within the TILDA health assessment centre, conducted by one of two research nurses.

130 participants then completed an identical repeat assessment, scheduled to be conducted after approximately one month (for 50% of participants) or three months (for the remainder).

Among half of the participants the nurse conducting the assessment was changed for the repeat.

Change of nurse, change of time of day (afternoon or morning assessment) and delay between assessments were balanced along with age group and sex of the participants.

MMSE and MoCA scores were graphically compared between the initial and repeat assessments, between each rater, and time of day.

A mixed effects regression model was used to estimate the variation in test scores between and within participants. Rater, time of day and whether the observation was a baseline or repeat were added as fixed effects.

Within-person and between-person standard deviations arising from these models were used to estimate the residual intra-class correlation coefficients (ICC). The ICC is the proportion of total variance not accounted for by within person variation, that is  $ICC = SD_{Between}^2 / (SD_{Between}^2 + SD_{Within}^2)$ .

Minimum detectable change was estimated by  $MDC = SD_{Within} \times \sqrt{2} \times Z$  where  $Z=1.96$  for the 95% MDC and  $Z=1.64$  for the 90% MDC.

Confidence intervals for reliability statistics and differences in reliability across lag times were estimated using non-parametric bootstrapping, based on 1000 replicates sampled at the participant level.

## Sample

A total of 128 participants underwent repeat assessments (58 men) of both MMSE and MoCA. The median age of the sample was 66 (range 55-89, IQR 61-71). The median delay between assessments was 88 days (range 28-141 days, IQR 70-104 days).

## The Reliability of MMSE and MoCA

The table below shows baseline mean, repeat mean, and inter-rater and reliability statistics for MoCA and MMSE.

	MMSE	MoCA
Baseline mean (s.d.)	28.4 (2.06)	25.2 (3.56)
Repeat mean (s.d.)	28.1 (2.06)*	24.8 (3.61)*
Rater 1 mean (s.d.)	28.0 (2.22)	24.8 (3.68)
Rater 2 mean (s.d.)	28.2 (2.24)	25.1 (3.95)
Test AM mean (s.d.)	28.5 (1.60)	25.3 (3.30)
Test PM mean (s.d.)	28.6 (1.97)	25.8 (3.12)
SD <sub>between</sub> (95% CI)	1.77 (1.39-2.15)	3.22 (2.79-3.65)
SD <sub>within</sub> (95% CI)	1.02 (0.91-1.13)	1.52 (1.32-1.72)
ICC (95% CI)	0.75 (0.67-0.83)	0.81 (0.77-0.87)
90% MDC (95% CI)	2.38 (2.12-2.63)	3.54 (3.08-3.99)
95% MDC (95% CI)	2.83 (2.53-3.13)	4.21 (3.67-4.75)

Scores for MMSE and MoCA were both significantly higher at the repeat indicating a practice effect.

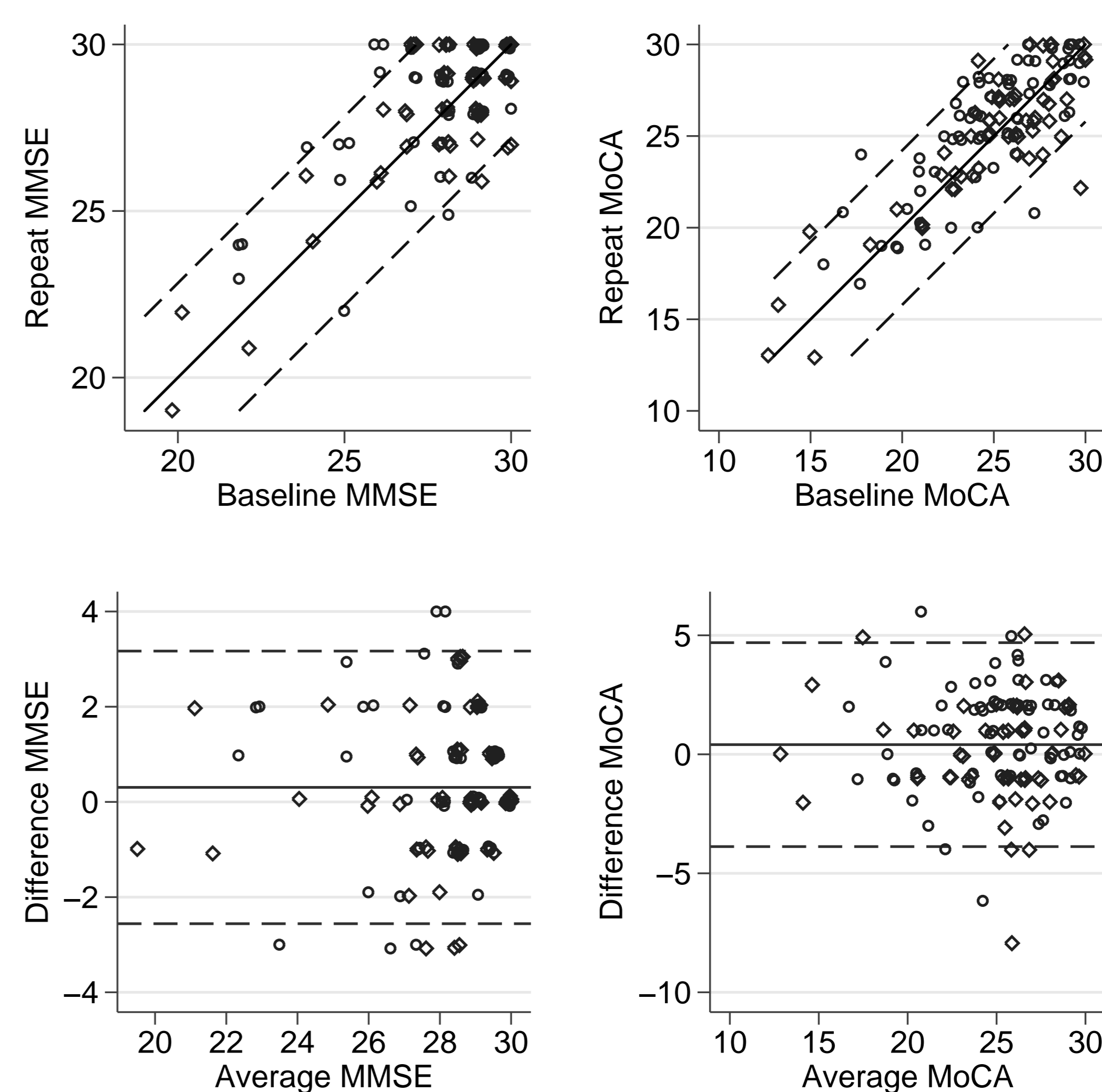
There was no effect of rater or time of day on MMSE or MoCA.

MoCA (ICC=0.81) is more reliable in this relatively healthy sample than MMSE (ICC=0.75).

The lag between assessments makes no difference to any reliability statistic (data not shown)

**The minimum detectable change in MMSE is around 3 points, for MoCA is around 3-4 points.**

## Graphical representation of MoCA and MMSE reliability



**Figure.** (above) Baseline and repeat scores for MoCA and MMSE with dashed lines indicating the 95% minimum detectable change. (below) Bland-Altman plots showing the difference vs average MoCA and MMSE scores. Circles indicate repeat assessments taken by the same rater, diamonds indicate a repeat taken by a different rater to the baseline assessment.

## Conclusions

We have estimated the test-retest reliability and hence the minimum detectable change of two commonly used cognitive tests using a sample representative of a the population aged 55 and older in Ireland. These estimates should be used when assessing changes in cognitive test scores in individuals or when using methods to adjust for measurement error in research studies.

For MMSE, the standard error of measurement is around 1 point meaning that a difference of 3 points between two observations should be considered beyond that expected by chance. For MoCA the SEM is 1.5 points meaning that a difference of 3-4 points could be expected by chance depending on the confidence level used.

TILDA is supported by the Department for Health and Children, Irish Life PLC and the Atlantic Philanthropies. The TILDA/SHARE collaboration was supported by the National Institute for Aging grant number R21 AG040387.