



Article

Selecting a Bedside Cognitive Vital Sign to Monitor Cognition in Hospital: Feasibility, Reliability, and Responsiveness of Logical Memory

Padraic Nicholas ^{1,†}, Rónán O’Caoimh ^{1,2,3,*}, Yang Gao ¹, Afsana Habib ³,
Thomas Karol Mross ⁴, Roger Clarnette ⁵ and D. William Molloy ^{1,3}

- ¹ Centre for Gerontology and Rehabilitation, University College Cork, St Finbarr’s Hospital, Douglas Road, T12 XH60 Cork, Ireland; 109746021@umail.ucc.ie (P.N.); xhandgy@gmail.com (Y.G.); w.molloy@ucc.ie (D.W.M.)
- ² Clinical Sciences Institute, National University of Ireland Galway, Costello Road, H91 TK33 Galway, Ireland
- ³ Mercy University Hospital, Grenville Place, T12 WE28 Cork, Ireland; afsana.habib@health.wa.gov.au
- ⁴ Geriatisches Zentrum Bethanien, Agaplesion Bethanien Krankenhaus, Rohrbacher Str. 149, 69126 Heidelberg, Germany; thomasmross@googlegmail.com
- ⁵ School of Medicine and Pharmacology, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia; roger.clarnette@health.wa.gov.au
- * Correspondence: rocaoimh@hotmail.com; Tel.: +00353-91-493187
- † These authors contributed equally to the work.

Received: 31 July 2019; Accepted: 19 September 2019; Published: 22 September 2019



Abstract: Although there is a high prevalence of delirium and cognitive impairment among hospitalised older adults, short, reliable cognitive measures are rarely used to monitor cognition and potentially alert healthcare professionals to early changes that might signal delirium. We evaluated the reliability, responsiveness, and feasibility of logical memory (LM), immediate verbal recall of a short story, compared to brief tests of attention as a bedside “cognitive vital sign” (CVS). Trained nursing staff performed twice-daily cognitive assessments on 84 clinically stable inpatients in two geriatric units over 3–5 consecutive days using LM and short tests of attention and orientation including months of the year backwards. Scores were compared to those of an expert rater. Inter-rater reliability was excellent with correlation coefficients for LM increasing from $r = 0.87$ on day 1 to $r = 0.97$ by day 4 ($p < 0.0001$). A diurnal fluctuation of two points from a total of 30 was deemed acceptable in clinically stable patients. LM scores were statistically similar ($p = 0.98$) with repeated testing (suggesting no learning effect). All nurses reported that LM was feasible to score routinely. LM is a reliable measure of cognition showing diurnal variation but minimal learning effects. Further study is required to define the properties of an ideal CVS test, though LM may satisfy these.

Keywords: attention; cognitive screening; cognitive vital sign; logical memory; older adults; delirium; dementia; hospital

1. Introduction

Identifying acute cognitive changes in hospital is challenging [1]. Decline in cognition after admission is multi-factorial and may relate to medical conditions, functional deterioration, pharmacological treatments, or the environment itself [1,2]. Although fluctuating cognition is increasingly considered to be a vital sign, particularly in older adults [3], few short, valid, and reliable cognitive tests are available in hospital to routinely monitor changes in cognitive function [4], where the prevalence of delirium approaches 20% [5] and major neurocognitive disorders such as dementia approach 50% [6,7]. Reflecting this, delirium and dementia are frequently unrecognized in hospitals [2,8].

Although serial assessment is important to monitor cognition in dementia [9], it is challenging as inpatients often experience fluctuations and changes in cognition ranging from normal daily variations to those resulting from delirium [10].

Simple tests, such as the months of the year backwards (MOTYB), focusing on attention may be useful in detecting delirium in hospital [10,11]. Further, short cognitive screening instruments can separate delirium from established cognitive impairment [12]. A brief cognitive or delirium screening instrument could therefore be useful as a “cognitive vital sign” (CVS), which if performed regularly on hospital patients could enable monitoring of cognitive function. Early changes in a CVS, outside the normal or the patients’ previously stable baseline range, might alert staff of an underlying problem or clinical deterioration. An early diagnosis of delirium would prompt the search for a cause of the decline [2], potentially limiting its progression and significant associated costs [13].

Another challenge is how to allocate responsibility for the administration of these instruments in routine clinical practice. While it is suggested that nursing staff may be best placed to monitor a CVS [3], the feasibility of this is unclear. Evidence suggests that healthcare professionals are more receptive to using brief and easy-to-administer tools [14,15]. An ideal CVS should also measure as many cognitive domains as possible and have multiple alternative forms to minimize learning or practice effects [16]. In addition, bias associated with age, education, visual and hearing impairment, manual dexterity, sleep status, and other factors including the time of day may impact upon performance and the reliability of testing and should be considered when selecting a candidate CVS [17–20].

Logical memory (LM) [21], immediate or delayed verbal recall of a short story, is a test of episodic memory originally included as part of the Wechsler Memory Scale [21]. Logically ordered stories test attention and memory and may be easier to encode for those with pre-existing executive dysfunction [22], suggesting LM may be a useful CVS. LM has been included in formal cognitive screening instruments such as the Quick Mild Cognitive Impairment (*Qmci*) screen [23–27], where its immediate version (recall after only a momentary interval), is the most accurate subtest for identifying early cognitive impairment [28]. It is also used in several longitudinal studies [29,30], though like many tests there are concerns over practice effects [22,31] and test-retest reliability [22] even when using alternative forms [31].

The objective of this study was (1) to measure the reliability and feasibility of a selection of potential CVS candidates including LM, MOTYB, and digit-span testing, (2) to examine if normal diurnal fluctuations in these occur, and (3) to assess if the use of LM as a bedside CVS is feasible in hospital. The study also sought to minimize learning effects by developing new alternative LM stories to identify for that are acceptable to nursing staff.

2. Materials and Methods

2.1. Patients

Patients admitted to wards in two university teaching hospitals (one acute geriatric hospital and one adult rehabilitation unit) in Cork City, Ireland, in May 2013 were invited to participate. Suitable patients were selected using convenience sampling directed by a neutral party (clinical nurse manager) to minimise selection bias. Patients were assessed for a minimum of three days and up to a maximum of five consecutive days. Patients were included if they were over 18 years of age, clinically stable at baseline (defined as no change in medication dosage or frequency and no recent change or exacerbation in their medical condition on day 1, not receiving end-of-life care confirmed by the clinical nurse manager), if their remaining admission was expected to be greater than three days, and if they were able to provide informed written consent. The study received ethics approval (April 2013) from the Research Ethics Committee of the Cork teaching hospitals (ECM 4 (gg) 07/05/13).

2.2. Outcomes

LM is a brief test of verbal recall for a short descriptive story; in this study the immediate version (within 30 s) was used. It has 15 items and is scored on a 30-point scale with two points given for every correct answer [32]. Administration takes less than one minute and alternate validated versions are available [33]. LM evaluates working memory, verbal recall, attention, expressive language, and executive function [21]. It is not biased by age or education [28] and is not affected by visual or physical disability. LM is valid in those with different dementia and mild cognitive impairment subtypes [34].

The other rapidly administered measures of attention scored included digit-span testing (auditory presentation of number lists repeated forwards/backwards), MOTYB, and orientation to time. These tests are also presented in the Appendix A. MOTYB is a well-validated test of attention that when used alone is the most accurate, brief bedside screening test for delirium in hospital [10]. In this study it was scored out of 15 points, one point for each month in order and up to three additional points for time of completion (three points for <20, two for 20–25 and one for >25 s). Five-item digit-span forwards and digit-span backwards were each scored out of 10 points, two for each item recalled in order. Digit-span testing is a useful indicator of delirium and cognitive impairment in hospital [35], with the forwards and backwards versions measuring different aspects of cognition; backwards recall draws on visuospatial processing with both versions assessing attention and working memory [36]. Orientation to time, another useful stand-alone guide to the presence and severity of dementia or delirium in hospitalised patients [37], was scored out of five points (a maximum of five if the stated time was within 30 min, four within 30–60 min, three within 1–2 h, two if within three hours and one point if the patient's estimate was more than three hours off the actual time). All scores were converted to percentages.

2.3. Study Measures and Procedures

A pilot study, sample ($n = 10$), was performed to evaluate the feasibility of using LM and to determine administration time. Participants included in pilot testing were excluded from the final research study. Ten alternate forms of LM, three of which had been validated previously [33], were available for repeat testing (see Appendix A). Slight amendments were made to the layout of the new alternative forms and patient instructions as a result of piloting. Data collection questionnaires were piloted simultaneously to ensure they would be time efficient and accurate in recording the study findings in the test environment. Nurses were trained to administer the tests and asked to complete a questionnaire at the end of the study.

Administration and scoring instructions were developed to standardize the testing. A trial demonstration was performed for nursing staff by the principle investigator (DWM). LM was scored as follows: The patient was first asked to concentrate, and any distractions on the ward, were removed (e.g., radios, televisions, hoovers were switched off). On the first consultation an auditory exam, the whispered voice test [38], was performed to test hearing. Patients were warned they were going to hear a short story and that they should repeat back as much of the story as possible, in any order, immediately after the story ending. Nurses were instructed to read the paragraph at about one second for each word unit until complete and wait for a maximum of 30 s for responses. No hints were provided. The patients were read a different version each time. The other tests were scored in the following fixed order (digit-span forwards, digit-span backwards, MOTYB, and orientation to time), see the attached scoring instructions in the Appendix A.

Administration of LM and the battery of other instruments were performed in random order. Independent assessments were performed twice daily (morning/afternoon) by an expert rater (PN). Each assessor was blind to the other's results. In addition, prior to scoring the CVS on patients, nurses were asked to document whether they felt the patient's global cognitive condition was subjectively the same, worse, much worse, better, or much better compared to the previous day, using a simple Likert scale. If patients were deemed to have had a clinical change, further questions were asked to

determine the nature of these variations, e.g., were activities of daily living (ADLs) affected? Questions were derived following consultation with dementia specialists (Appendix A).

2.4. The Questionnaire

A questionnaire was formulated and administered to nurses to identify their view of the time needed to administer the test, their comprehension of the standardised instructions, and the feasibility of performing the CVS testing routinely on the ward. Each questionnaire was pre-assigned an identification number. It was completed by all 14 nurses who performed CVS testing. The questionnaire consisted of a short demographic page and a number of questions using a “Likert scale” format. Nurses were questioned on time needed to administer, comprehension of the standardised instructions, feasibility to perform on the ward, and their willingness to perform a CVS.

2.5. Statistical Analysis

Data were analysed using SPSS version 18.0 (IBM, Armonk, NY, USA). Inter-rater reliability (IRR) between raters (expert rater and nursing staff) and the relationship between LM scores and other tests were evaluated using bivariate Pearson’s correlation analysis; IRR was compared for each patient for each day of the study. Canonical correlation and linear mixed modelling were also used. Daily fluctuations in the score for each test were calculated by establishing each patient’s mean score over the testing period. The largest variation around this mean was then taken (upper limits of normal fluctuation) for each patient and an overall mean variance established for this sample. Variability was explored in a number of ways. Variance components analysis was used to explore total variability of scores and the proportion of that variability attributed to patient factors and day-to-day variability; the proportion explained by the patient is the variability of scores attributed to differences between individuals. The remaining proportion is defined as day-to-day variability, the difference in scores for the same patient between different days. Variation from day to day was also examined using maximum difference from patient-level mean. Learning effects were explored using linear mixed modelling (LMM). This regression technique can be applied where observations are not independent from each other (i.e., observations belonging to the same patient). The study day (number) was used as a predictor of the score as both a categorical and continuous variable with significant results indicating changes in scores over time and suggesting learning effects. Subgroup analysis was performed for age and level of formal education to investigate if greater fluctuations existed in scores within these subgroups.

3. Results

3.1. Baseline Characteristics

In total, 100 patients were assessed. Sixteen were excluded from the study because they were discharged less than three days after commencing the study ($n = 12$), or their clinical condition worsened ($n = 4$) during the course of their participation. The remaining 84 adult inpatients (47 males/37 females; Ratio 1.3:1), were included: $n = 38$ from a rehabilitation unit and 46 from an acute geriatric ward. All patients were Caucasian. The majority of patients were aged >65 years ($n = 69$, 82.1%). There was an inverse relationship between increasing age and decreasing mean LM scores (Table 1). Thirty patients (37.5%) did not complete second-level education, 42 (50%) completed secondary school only, and 12 (14.3%) completed third-level education. Educational attainment up to secondary level did not reflect mean LM scores in this study population. Those who completed third level had higher mean LM scores (16.17) than those who did not (9.9), $p < 0.05$.

Table 1. Age make-up of the cohort * and their associated mean logical memory (LM) scores.

Patients (%)	84	5 (6%)	10 (11.9%)	31 (36.9%)	33 (39.2%)	5 (6%)
Age	Total	18–49	50–65	66–79	80–89	≥90
Mean LM score	10.7	17.6	10.2	11.81	9.7	7.20

* Expressed as a percentage of the sample group in question.

3.2. Reliability, Variance Day-to-Day (Diurnal), and Learning Effects

The trained nurses and independent rater examined patients over three to five days. Data sets were analysed and categorised into daily results, (see Table 2). IRR for LM was excellent, increasing from $r = 0.87$ on day 1 ($n = 41$ patients) to $r = 0.97$ by day 4 ($n = 13$ patients), a statistically significant difference, $p < 0.0001$. Linear mixed model regression ($r = 0.93$, $p < 0.001$) and canonical correlation (days 1–3) confirmed overall excellent IRR ($r = 0.92$, $p < 0.001$). Excluding those aged <50 years, did not alter the reliability, with the IRR increasing from $r = 0.84$ to $r = 0.98$ between day 1 and day 4. Using variance components analysis, we found that most variation between scores was due to patient factors; we observed that the largest day-to-day variability was for orientation to time (33.3%) and numbers forwards (17.7%), while the lowest was for MOTYB (4.2%) and LM (9%). We found statistically significant learning effect for numbers forwards and numbers backwards, with scores increasing (improving) significantly in the first three days, indicating learning effects (see Table 2).

Table 2. Variation analysis with mean \pm standard deviation (SD) scores for each test over time including logical memory (LM), numbers forwards and backwards, months of the year backwards (MOTYB), and orientation to time.

Test	LM (Independent Rater)	LM (Trained Nurses)	Numbers Forwards	Numbers Backwards	MOTYB	Orientation (Time)
Overall scores						
Mean \pm SD	10.9 \pm 5.5	10.9 \pm 5.9	9.1 \pm 1.4	5.1 \pm 2.4	11.1 \pm 4.1	4.2 \pm 1.0
Scores by day of administration (Mean \pm SD)						
Day 1	10.9 \pm 5.4	10.9 \pm 5.4	8.8 \pm 1.6	4.8 \pm 2.6	11.1 \pm 4.1	4.2 \pm 1.2
Day 2	10.7 \pm 5.1	11.5 \pm 5.7	9.3 \pm 1.3	5.3 \pm 2.4	11.1 \pm 4.1	4.2 \pm 0.9
Day 3	10.8 \pm 5.5	10.5 \pm 5.9	9.3 \pm 1.4	5.3 \pm 2.3	11.3 \pm 4.1	4.0 \pm 0.9
Day 4	11.5 \pm 7.3	11.5 \pm 8.0	9.7 \pm 0.8	4.3 \pm 2.2	10.3 \pm 4.0	4.3 \pm 0.7
Day 5	12.3 \pm 5.9	9.3 \pm 5.9	10.0 \pm 0.0	5.3 \pm 1.2	11.7 \pm 3.5	4.7 \pm 0.6
Variance components analysis (% of variance explained by various factors)						
Time (day of follow-up)	0.3%	0.8%	3.5%	1.5%	0.3%	1.0%
Patient	90.7%	86.7%	78.7%	88.7%	95.5%	65.7%
Day-to-day variability	9.0%	12.5%	17.8%	9.8%	4.2%	33.3%
Variation from day to day based on max difference from patient-level mean						
Mean/Median	1.8/1.3	2.2/2.0	0.5/0.0	1.0/1.3	0.9/0.7	0.7/0.7
% overall mean	16.6%	20.3%	5.9%	18.8%	7.7%	17.2%
% patients with zero fluctuation	11.9%	14.6%	63.1%	28.9%	32.1%	10.8%
Variation from day to day (based on standard deviation at the patient level)						
Mean/Median	1.5/1.2	1.9/2.0	0.5/0.0	0.8/1.2	0.8/0.6	0.6/0.6
Learning effect (Linear Mixed Model, significance of study day)						
Categorical	$p = 0.70$	$p = 0.54$	$p < 0.001^a$	$p = 0.001^c$	$p = 0.23$	$p = 0.28$
Continuous	$p = 0.30$	$p = 0.33$	$p < 0.001^b$	$p = 0.002^d$	$p = 0.40$	$p = 0.46$

^a There is a significant difference between day 1 and day 2 ($p < 0.001$) and between day 1 and day 3, $p < 0.001$;

^b Numbers forward scores increase significantly by an average of 0.21 points/day (95% CI 0.11–0.31), $p < 0.001$;

^c There is significant difference between day 1 and day 2 ($p = 0.003$) and between day 1 and day 3, $p = 0.003$;

^d Numbers backwards scores increase significantly by an average of 0.20 points/day (95% CI 0.07–0.33), $p = 0.002$.

Mean LM scores ranged from 0 to 28.7/30 according to age. The mean LM score was 10.7/30. The mean variation in the standard deviation was ± 1.5 and 1.9 points for the trained rater and nurses

respectively, indicating that fluctuations of approximately two points from the individual's baseline were outside normal daily changes in this clinically stable sample (see Table 2). LM scores were statistically the same ($p = 0.98$) from days 1 to 3. A slight rise in mean LM scores was evident on days 4 and 5, but these changes were not statistically significant, though sample numbers were small. After removing the results for younger patients (<50 years), LM scores still showed no change over time, reducing from a mean of 10.4 ± 4.9 on day 1 to 10.3 ± 5.0 points by day 3 ($p = 0.98$).

3.3. Comparison of Logical Memory with Short Tests of Attention and Orientation

The correlation between LM and tests of attention (digit-span forwards/backwards, months backwards), working memory (digit-span forwards/backwards) and orientation to time are presented in Table 3 and Figure 1. Data from day 4 and 5 were excluded as sample sizes were considered to be too small. Numbers backwards had the highest consistent correlation with LM ($r = 0.61$). Overall, the correlation between LM and other rapid tests was low. There was an increased correlation between patient orientation to time and LM over time. The correlations were generally lower when younger patients (<50 years old) were excluded, though varied by test. Of note, orientation to time had stronger correlation in older patients >50 years (Figure 2).

Table 3. Correlation coefficients (Spearman's rho) for logical memory and short tests of attention including months of the year backwards (MOTYB). (Note: $n = 83/84$ with complete data included; values day 4 and 5 were excluded).

Days	Digit-Span Forwards	Digit-Span Backwards	MOTYB	Orientation to Time
Day 1	0.58	0.54	0.43	0.27
Day 2	0.60	0.48	0.42	0.36
Day 3	0.55	0.61	0.46	0.40

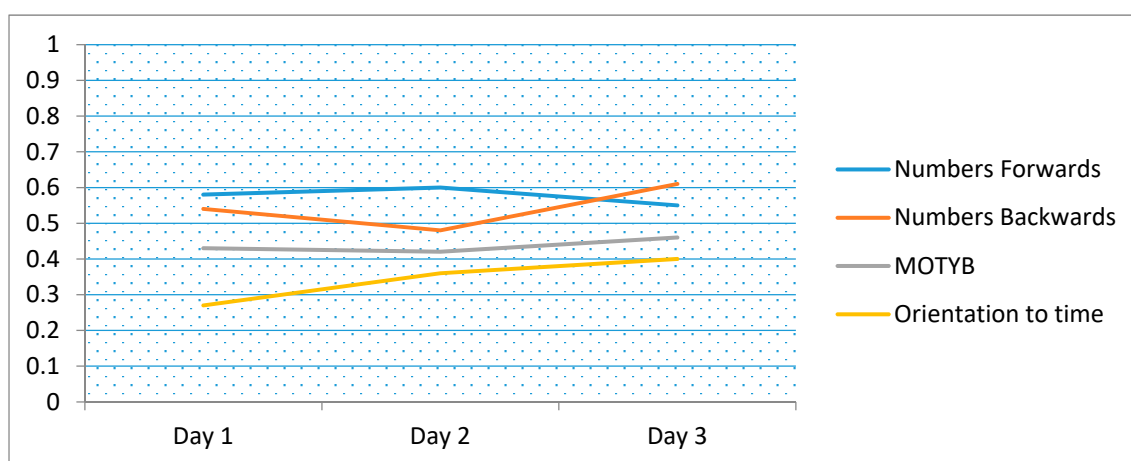


Figure 1. Correlation coefficients for logical memory compared to short tests of attention including months of the year backwards (MOTYB) and orientation (time of the day), over time (days 1 to 3). (Note: 83/84 with complete data included).

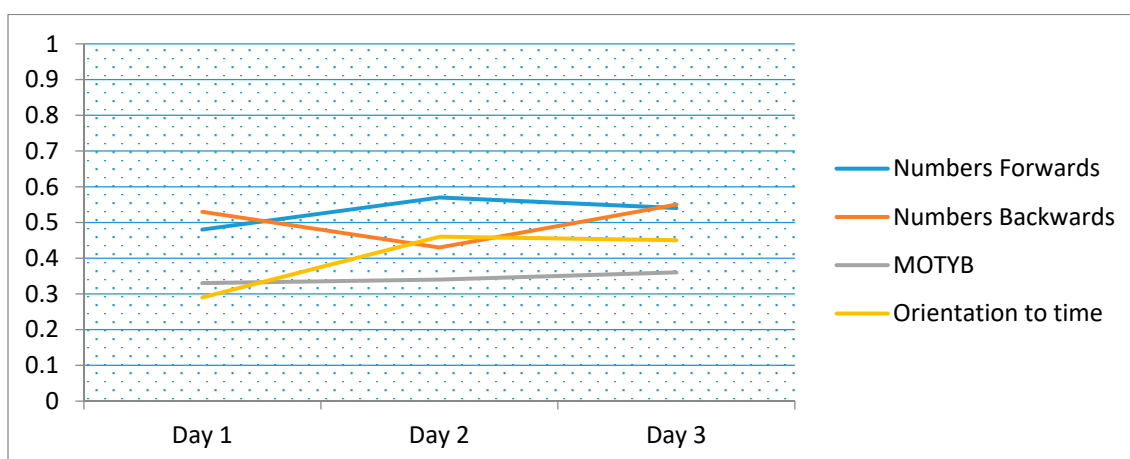


Figure 2. Correlation coefficients for logical memory compared to short tests of attention including months of the year backwards (MOTYB) and orientation (time of the day), over time (days 1 to 3). (Note: $n = 79$, excluding those <50 years).

3.4. Clinical Correlation

Fluctuations in LM scores were then compared to subjective clinical changes in the patients as documented by nursing staff. Nurses observed a global clinical change in 37 patients (44%). In 23 (62%) of these cases, LM score changes reflected nurses' opinions on the patient's clinical condition. For the remainder, LM scores did not correlate with their observations. In nine cases, nurses felt the patient's cognitive condition had become "worse or much worse" as the week progressed. Of these, decreased conversation and social withdrawal were recorded as the primary manifestations ($n = 6$) followed by changes in ADLs ($n = 2$), sleep pattern ($n = 1$), drowsiness ($n = 1$) and anxiety ($n = 2$). Some patients displayed more than one sign of cognitive change.

3.5. Feasibility of Using Logical Memory as a Cognitive Vital Sign

A majority of the 14 trained nurses, 86% ($n = 12$), reported that they administered and scored the LM in less than 3 min. Two nurses, aged >60 years, reported taking 3–5 min, although the age of nurse raters and administration time were not statistically significantly associated. All nurses felt the instructions for testers could be understood without difficulty, that LM was a feasible CVS to perform in a busy inpatient environment, and that they would be willing to use it in the future. None of the 14 nurses suggested changes to improve the scoring.

4. Discussion

This study presents the reliability, diurnal fluctuations, test, and clinical correlations and perceived feasibility of using LM as a CVS on a daily basis in a stable and predominantly older inpatient population. The results show that LM has high levels of IRR that improved with repeated administration (correlation coefficients range 0.87–0.97). These support other studies showing that the reliability of LM is reasonable [39]. Normal daily fluctuations in LM scores were found, supporting evidence that test-retest reliability is variable [40], though LM scores were statistically the same from days 1 to 3 indicating that unlike other studies [22], there were no statistically significant learning effects evident with repeated LM scoring as the week progressed. The results also show that there was moderate correlation between LM and a battery of short tests of attention, working memory, and orientation (time of the day) that was strongest for numbers forwards and backwards. This is expected as LM tests a wider range of cognitive domains including episodic and working memory [21] compared to simple tests of attention and orientation alone.

Variations in LM scores (fluctuations) correlated with the nurses' opinion of clinical change in only two-thirds of cases, suggesting that sensitivity was modest. Fluctuations of approximately two

out of 30 points were found to be outside of the diurnal variation. However, while fluctuations of this magnitude may be abnormal, meriting further investigation, the likely relatively low sensitivity indicates that changes in LM scores should not be considered in isolation, but viewed in the context of patients' overall clinical and cognitive status.

The study also compared LM to short tests of attention and orientation often used for detecting delirium and dementia. There was generally only moderate correlation between the LM and these tests, and it was highest for digit-span forwards/backwards. Orientation to time had stronger correlation with LM in older patients, i.e., when those <50 years were excluded. This may be because both orientation [37] and LM [28] are useful tests of cognitive impairment in older patients. Many patients persistently scored towards the upper limits of these other tests, indicating possible ceiling effects. For example, very little variability between individuals was detected for digit-span forwards/backwards suggesting that these tests would not be suitable for a general inpatient population who typically would have a higher cognitive status than the sample tested in this study. This study found that LM and MOTYB had the least tendency towards a learning effect, while digit-span forwards/backwards, possibly due to the repetitive nature of this test, showed evidence of significant learning effects. These highlight potential problems using these other tests in detecting cognitive change when applied on a daily basis, strengthening the case for using LM as a CVS.

Trained nurses perceived that LM was feasible and acceptable to administer as a CVS. The significant increase in IRR as the study progressed may relate to increased assessor and patient familiarity with the test, providing more stable results over time. This suggests that it is beneficial to perform an initial recording (trial run), to improve patient concentration and understanding of the screen, before taking a baseline score. The small variation between the expert and nurse raters shows that scoring LM is not dependent on experience once the standardised instructions are followed. As several nurses scored the same patients (attending nurses changed between shifts and sections), the findings also indicate that LM, when used as a CVS, can be scored by multiple raters, enabling continuity of cognitive monitoring. The CVS was administered in less than three minutes by 86% of nurses. Few brief tests are as time efficient in what is a recognised time efficacy trade-off [41].

Strengths and Limitations

Strengths include that all cognitive tests were performed concurrently and that training and standardisation were conducted thoroughly before testing. There are a number of limitations. The sample included in this study was small (no sample size calculation was performed) and selected by convenience, which may have introduced bias. However, this was a feasibility study and patient sampling in this way was required to ensure that inappropriate patients, e.g., those with severe illness, active delirium, or those receiving end-of-life care at baseline, were not included. While this limits the study findings to those who were stable, it was required to measure normal day-to-day (diurnal) variability in a steady state and show the potential of these instruments to be reliable. Further study to correlate change outside of this normal variability with the onset of delirium is now required. Similarly, no gold standard measure of delirium or cognition was used in this study, meaning it was not possible to correlate or associate scores with confirmed diagnostic states. Baseline cognition was not assessed, and it is possible that some variation in scores may have related to pre-existing cognitive disorders, e.g., there was some indication that a "sundowning" effect may have influenced evening testing, but it was not possible to confirm. No follow-up of patients was conducted, which also reduces the ability to interpret the findings. Likewise, no independent measure of change was recorded. Instead, LM was correlated with established bedside tests of attention and orientation such as MOTYB [10] and staff nurses' global subjective opinion of change, recorded prior to each assessment. Although this may have created bias, nurses were instructed to record this before and not after administering the CVS. The repetitive nature of many brief bedside cognitive or delirium tests is a fundamental weakness in conducting daily testing [42,43] and monitoring for cognitive change. It facilitates learning effects with repeated use. To minimise this, we used multiple formats for LM (10 variations) and ensured

patients were given different forms for each administration. However, only three had previously been validated, potentially resulting in bias. This said, all 10 variations had the same identical structure only varying slightly and were developed by the designers of the *Qmci* screen of which LM is a subset. All participants in this study were Irish nationals of Caucasian ethnicity; this limits external validity and generalisability. There were also potential environmental influences, which may have affected testing and created bias; patients may have scored better for example (e.g., in orientation) if they were rehabilitating rather than in an acute ward or had been listening to the news, had breakfast, or recently checked the time in anticipation of cognitive testing or attending scheduled therapy appointments. This however, reflects real-life practice, arguably a strength of the study.

5. Conclusions

In summary, LM appears to be a reliable and efficient candidate for a CVS for use in routine ward practice, a brief test that could detect changes in cognition while in hospital to indicate delirium and those who require specialist referral and assessment. It was also deemed acceptable, easy, and quick to administer by nurses working in clinical practice. It showed moderate correlation with established tests of attention, working memory, and orientation. However, this study was not designed to show superiority of one instrument over another or predictive accuracy for delirium or other acute mental status changes. Instead, this pilot study focused on the reliability and feasibility of LM. Further research is needed to compare it with other short delirium and cognitive screens such as the 4AT and 6-Item Cognitive Impairment Test [12] as well as established standards for diagnosing cognitive impairment and delirium, in a less select sample of consecutive admissions with and without established cognitive impairment and to investigate whether incorporating it into assessments such as the modified early warning score [44] could promote time-efficient detection of early cognitive changes, leading to better use of resources, rapid treatment if required, and improved patient outcomes.


Author Contributions: P.N., R.O. and D.W.M. conceived and designed the experiments; P.N. and D.W.M. performed the experiments; P.N. and Y.G. analysed the data; P.N., R.C., A.H., T.K.M., and R.O. wrote the paper.

Funding: This research was supported by a Health Research Board of Ireland student summer scholarship.

Acknowledgments: The authors thank Anton Svendrovski who provided statistical support for this analysis.


Conflicts of Interest: The authors declare no conflict of interest. The funding organisation had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Cognitive Vital Sign Study



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Pt. Study No.	Pt. Initials	Gender	Pt. Age	Diagnosis	Presenting Complaint	Location

Clinical global impression

Please tick which box best describes your overall impression of the patient's cognitive state in the past 24 hour period. Please score in terms of confusion and ability to communicate in comparison with yesterday.

Day	Date	Time	Signature	Much worse	Worse	No change	Better	Much better
<u>Day 1</u>				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<u>Day 2</u>				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<u>Day 3</u>				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<u>Day 4</u>				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<u>Day 5</u>				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Administration and scoring guidelines for the Logical Memory test.

The Logical Memory test is administered twice daily, once AM once PM. Please make sure the patient is comfortable and hearing aid is in place if applicable. Test should be performed in a quiet area with no distractions. Test hearing by asking, "Are you able to do a short test for me?" If the patient consents please proceed with logical memory test.

Figure A1. Cont.




Instructions to the patient:

"I am going to read you a short story. After I have finished reading I want you to tell me as much as you can. OK?" (Patient signifies agreement, then begin reading the paragraph at about 1 second for each word unit) "The red...fox...ran across...the ploughed...field".

Scoring: For each word unit in **bold** the patient repeats, award **two marks**. One mark cannot be given. For a mark to be given the patient must state the **exact word in any order. Do not prompt**. If the patient states the test was too hard encourage them to recite anything they remember. Each day (1-5) the subsequent story will be read to avoid learning.

Day 1		Day 2		Day 3		Day 4		Day 5	
Date	Time	Date	Time	Date	Time	Date	Time	Date	Time
The red	/2	The brown	/2	The white	/2	The brown	/2	The blue	/2
Fox	/2	Dog	/2	hen	/2	Robin	/2	Duck	/2
Ran across	/2	ran across	/2	Walked across	/2	Hopped along	/2	Swam across	/2
The ploughed	/2	The metal	/2	The concrete	/2	The tin	/2	The shallow	/2
Field.	/2	Bridge.	/2	Road.	/2	Roof.	/2	Lake.	/2
It was chased	/2	It was hunting	/2	It was followed	/2	It was watched	/2	It was following	/2
by		by		by		by			
A brown	/2	A white	/2	A black	/2	A grey	/2	A black	/2
Dog.	/2	Rabbit.	/2	Cat.	/2	Cat	/2	Swan	/2
It was a hot	/2	It was a cold	/2	It was a warm	/2	It was a cold	/2	It was a bright	/2
May	/2	October	/2	September	/2	January	/2	Spring	/2
Morning.	/2	Day.	/2	afternoon.	/2	Morning.	/2	Evening.	/2
Fragrant	/2	Ripe	/2	Dry	/2	Icey	/2	Tall	/2
Blossoms	/2	Apples	/2	Leaves	/2	berries	/2	Reeds	/2
Were forming	/2	Were hanging	/2	Were blowing	/2	Were hardening	/2	Were swaying	/2
On the bushes.	/2	On the trees.	/2	In the wind.	/2	In the frost.	/2	In the breeze.	/2
Score	/30	Score	/30	Score	/30	Score	/30	Score	/30


Figure A1. Cont.



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Rapid

Comparative Tests



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Test 1. Numbers to repeat forwards

27, 381, 7361, 39472, 681579

	<u>Day 1</u>	<u>Day 2</u>	<u>Day 3</u>	<u>Day 4</u>	<u>Day 5</u>
<u>Date</u>					
<u>Score</u>	___/10	___/10	___/10	___/10	___/10

Test2. Numbers Backwards

37, 935, 4175, 83519, 725163

	<u>Day 1</u>	<u>Day 2</u>	<u>Day 3</u>	<u>Day 4</u>	<u>Day 5</u>
<u>Score</u>	___/10	___/10	___/10	___/10	___/10

Test 3. Months of the year

<20 Seconds, 20-25 Seconds, 25-30 Seconds


	<u>Time</u>	<u>D</u>	<u>N</u>	<u>O</u>	<u>S</u>	<u>A</u>	<u>J</u>	<u>J</u>	<u>M</u>	<u>A</u>	<u>M</u>	<u>F</u>	<u>J</u>	<u>Score</u>
<u>Day 1</u>														___/15
<u>Day 2</u>														___/15
<u>Day 3</u>														___/15
<u>Day 4</u>														___/15
<u>Day 5</u>														___/15

Test 4. Time < 30mins=5, 30-60mins=4, 1-2hrs=3,2-3hrs=2, >3hrs 1

	<u>Day 1</u>	<u>Day 2</u>	<u>Day 3</u>	<u>Day 4</u>	<u>Day 5</u>
<u>Score /5</u>					


Combined Score: ___/30

Figure A1. Outcome measures including logical memory test score sheets with alternative versions.



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Primary Investigator sheet



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Pt. Study No.	Pt. Initials	Gender	Pt. Age	Diagnosis	Presenting Complaint	Location

Discuss with Carer: Changes in past 24 hours

	Day 1	Day 2	Day 3	Day 4	Day 5
<u>Date</u>					
<u>Time</u>					
<u>Sleep</u>					
<u>Drowsy</u>					
<u>Withdrawn</u>					
<u>Anxious</u>					
<u>Score</u>	__/20	__/20	__/20	__/20	__/20
<u>New Meds</u>					


1. A minimal change
2. A small change
3. Moderate change
4. Large change
5. Extreme change

Modified ADLS

0-5	Day 1	Day 2	Day 3	Day 4	Day 5
Conversation					
Groom					
Bath					
Walk					
Toilet					
Feed Themselves					
Transfer					
Total	__/35	__/35	__/35	__/35	__/35


0=Independently, 1= Prompting, 2=Set up required, 3=Supervision Required,4= Assistance required, 5= Complete Care Required

Figure A2. Investigator score sheets documenting clinical status.



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Administrator's feasibility Questionnaire



UCC
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

PLEASE TICK YOUR RESPONSE IN THE APPROPRIATE BOX

Your Gender: Male Female

Your Age: 18-24 25-39 40-49 50-64 65+

Is English your first language: Yes No

1. The length of time taken to read the short story and score the patient was:

Less than 2 Minutes Less than 3 Minutes Less than 5 Minutes Greater Than 5 Minutes

2. I found the standardised "instructions to patient" and "scoring" instructions:

Easy to understand Required time to understand Very difficult to understand

3. It is hoped in the future a rapid cognitive screen would be performed with other vital signs as an early indicator of subclinical cognitive change. I would be:

Willing to perform this rapid test Unwilling to perform such a test

3b. If not please briefly explain why:

4. Overall I found this rapid screen to be;

Feasible to perform in a Hospital setting Unfeasible to perform in hospital setting

4b. If not please briefly explain why:

5. Other problems I encountered (If any)

Figure A3. Feasibility questionnaire.

References

1. Mathews, S.B.; Arnold, S.E.; Epperson, C.N. Hospitalization and cognitive decline: Can the nature of the relationship be deciphered? *Am. J. Geriatr. Psychiatry* **2014**, *22*, 465–480. [[CrossRef](#)] [[PubMed](#)]
2. Jackson, T.A.; Gladman, J.R.; Harwood, R.H.; MacLulich, A.M.; Sampson, E.L.; Sheehan, B.; Davis, D.H. Challenges and opportunities in understanding dementia and delirium in the acute hospital. *PLoS Med.* **2017**, *14*, e1002247. [[CrossRef](#)] [[PubMed](#)]

3. Fick, D.M. The Critical Vital Sign of Cognitive Health and Delirium: Whose Responsibility Is It? *J. Gerontol. Nurs.* **2018**, *44*, 3–5. [[CrossRef](#)] [[PubMed](#)]
4. Voyer, P.; Champoux, N.; Desrosiers, J.; Landreville, P.; McCusker, J.; Monette, J.; Savoie, M.; Carmichael, P.H.; Richard, H.; Richard, S. RADAR: A measure of the sixth vital sign? *Clin. Nurs. Res.* **2016**, *25*, 9–29. [[CrossRef](#)] [[PubMed](#)]
5. Ryan, D.J.; O'Regan, N.A.; O'Caomh, R.; Clare, J.; O'Connor, M.; Leonard, M.; McFarland, J.; Tighe, S.; O'Sullivan, K.; Trzepacz, P.T.; et al. Delirium in an adult acute hospital population: Predictors, prevalence and detection. *BMJ Open* **2013**, *3*, e001772. [[CrossRef](#)]
6. Sampson, E.L.; Blanchard, M.R.; Jones, L.; Tookman, A.; King, M. Dementia in the acute hospital: Prospective cohort study of prevalence and mortality. *Br. J. Psychiatry* **2009**, *195*, 61–66. [[CrossRef](#)]
7. Timmons, S.; Manning, E.; Barrett, A.; Brady, N.M.; Browne, V.; O'shea, E.; Molloy, D.W.; O'regan, N.A.; Trawley, S.; Cahill, S.; et al. Dementia in older people admitted to hospital: A regional multi-hospital observational study of prevalence, associations and case recognition. *Age Ageing* **2015**, *44*, 993–999. [[CrossRef](#)]
8. Clegg, A.; Westby, M.; Young, J.B. Under-reporting of delirium in the NHS. *Age Ageing* **2011**, *40*, 283–286. [[CrossRef](#)]
9. McKhann, G.M.; Knopman, D.S.; Chertkow, H.; Hyman, B.T.; Jack, C.R., Jr.; Kawas, C.H.; Klunk, W.E.; Koroshetz, W.J.; Manly, J.J.; Mayeux, R.; et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dementia* **2011**, *7*, 263–269. [[CrossRef](#)]
10. O'Regan, N.A.; Ryan, D.J.; Boland, E.; Connolly, W.; McGlade, C.; Leonard, M.; Clare, J.; Eustace, J.A.; Meagher, D.; Timmons, S. Attention! A good bedside test for delirium? *J. Neurol. Neurosurg. Psychiatry* **2014**, *85*, 1122–1131. [[CrossRef](#)]
11. Adamis, D.; Meagher, D.; Murray, O.; O'neill, D.; O'mahony, E.; Mulligan, O.; McCarthy, G. Evaluating attention in delirium: A comparison of bedside tests of attention. *Geriatr. Gerontol. Int.* **2016**, *16*, 1028–1035. [[CrossRef](#)] [[PubMed](#)]
12. O'Sullivan, D.; Brady, N.; Manning, E.; O'shea, E.; O'grady, S.; O'Regan, N.; Timmons, S. Validation of the 6-Item Cognitive Impairment Test and the 4AT test for combined delirium and dementia screening in older Emergency Department attendees. *Age Ageing* **2018**, *47*, 61–68.
13. Weinreb, W.; Johannsdottir, E.; Karaman, M.; Füsgen, I. What does delirium cost? *Z. Gerontol. Geriatr.* **2016**, *49*, 52–58. [[CrossRef](#)] [[PubMed](#)]
14. Ismail, Z.; Mulsant, B.H.; Herrmann, N.; Rapoport, M.; Nilsson, M.; Shulman, K. Canadian academy of geriatric psychiatry survey of brief cognitive screening instruments. *Can. Geriatr. J.* **2013**, *16*, 54. [[CrossRef](#)] [[PubMed](#)]
15. Martin, S.; Kelly, S.; Khan, A.; Cullum, S.; Denning, T.; Rait, G.; Fox, C.; Katona, C.; Cosco, T.; Brayne, C.; et al. Attitudes and preferences towards screening for dementia: A systematic review of the literature. *BMC Geriatr.* **2015**, *15*, 66. [[CrossRef](#)] [[PubMed](#)]
16. Goldberg, T.E.; Harvey, P.D.; Wesnes, K.A.; Snyder, P.J.; Schneider, L.S. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimer's Dementia Diagn. Assess Dis. Monit.* **2015**, *1*, 103–111. [[CrossRef](#)]
17. O'Caomh, R.; Gao, Y.; Svendovski, A.; Gallagher, P.; Eustace, J.; Molloy, D.W. Comparing approaches to optimize cut-off scores for short cognitive screening instruments in mild cognitive impairment and dementia. *J. Alzheimer's Dis.* **2017**, *57*, 123–133. [[CrossRef](#)]
18. Blatter, K.; Cajochen, C. Circadian rhythms in cognitive performance: Methodological constraints, protocols, theoretical underpinnings. *Physiol. Behav.* **2007**, *90*, 196–208. [[CrossRef](#)]
19. Gildner, T.E.; Liebert, M.A.; Kowal, P.; Chatterji, S.; Snodgrass, J.J. Associations between sleep duration, sleep quality, and cognitive test performance among older adults from six middle income countries: Results from the Study on Global Ageing and Adult Health (SAGE). *J. Clin. Sleep Med.* **2014**, *10*, 613–621. [[CrossRef](#)]
20. Pye, A.; Charalambous, A.P.; Leroi, I.; Thodi, C.; Dawes, P. Screening tools for the identification of dementia for adults with age-related acquired hearing or vision impairment: A scoping review. *Int. Psychoger.* **2017**, *29*, 1771–1784. [[CrossRef](#)]
21. Wechsler, D. A standardized memory scale for clinical use. *J. Psychol.* **1945**, *19*, 87–95. [[CrossRef](#)]

22. Gavett, B.E.; Gurnani, A.S.; Saurman, J.L.; Chapman, K.R.; Steinberg, E.G.; Martin, B.; Chaisson, C.E.; Mez, J.; Tripodis, Y.; Stern, R.A. Practice effects on story memory and list learning tests in the neuropsychological assessment of older adults. *PLoS ONE* **2016**, *11*, e0164492. [[CrossRef](#)] [[PubMed](#)]
23. O’Caoimh, R.; Gao, Y.; McGlade, C.; Healy, L.; Gallagher, P.; Timmons, S.; Molloy, D.W. Comparison of the quick mild cognitive impairment (Qmci) screen and the SMMSE in screening for mild cognitive impairment. *Age Ageing* **2012**, *41*, 624–629. [[CrossRef](#)] [[PubMed](#)]
24. O’Caoimh, R.; Svendrovski, A.; Johnston, B.C.; Gao, Y.; McGlade, C.; Eustace, J.; Timmons, S.; Guyatt, G.; Molloy, D.W. The Quick Mild Cognitive Impairment screen correlated with the Standardized Alzheimer’s Disease Assessment Scale–cognitive section in clinical trials. *J. Clin. Epidemiol.* **2014**, *67*, 87–92. [[CrossRef](#)] [[PubMed](#)]
25. O’Caoimh, R.; Sato, S.; Wall, J.; Igras, E.; Foley, M.J.; Timmons, S.; Molloy, W. Potential for a “Memory Gym” intervention to delay conversion of mild cognitive impairment to dementia. *J. Am. Med. Directors Assoc.* **2015**, *16*, 998–999.
26. Bunt, S.; O’Caoimh, R.; Krijnen, W.P.; Molloy, D.W.; Goodijk, G.P.; van der Schans, C.P.; Hobbelen, H.J. Validation of the Dutch version of the quick mild cognitive impairment screen (Qmci-D). *BMC Geriatr.* **2015**, *15*, 115. [[CrossRef](#)]
27. Clarnette, R.; O’Caoimh, R.; Antony, D.N.; Svendrovski, A.; Molloy, D.W. Comparison of the quick mild cognitive impairment (Qmci) screen to the Montreal cognitive assessment (MoCA) in an Australian geriatrics clinic. *Int. J. Geriatr. Psychiatry* **2017**, *32*, 643–649. [[CrossRef](#)]
28. O’Caoimh, R.; Gao, Y.; Gallagher, P.F.; Eustace, J.; McGlade, C.; Molloy, D.W. Which part of the Quick mild cognitive impairment screen (Qmci) discriminates between normal cognition, mild cognitive impairment and dementia? *Age Ageing* **2013**, *42*, 324–330. [[CrossRef](#)]
29. Chapman, K.R.; Bing-Canar, H.; Alosco, M.L.; Steinberg, E.G.; Martin, B.; Chaisson, C.; Kowall, N.; Tripodis, Y.; Stern, R.A. Mini Mental State Examination and Logical Memory scores for entry into Alzheimer’s disease trials. *Alzheimer’s Res. Ther.* **2016**, *8*, 9. [[CrossRef](#)]
30. Weintraub, S.; Salmon, D.; Mercaldo, N.; Ferris, S.; Graff-Radford, N.R.; Chui, H.; Cummings, J.; DeCarli, C.; Foster, N.L.; Galasko, D.; et al. The Alzheimer’s Disease Centers’ Uniform Data Set (UDS): The neuropsychologic test battery. *Alzheimer Dis. Assoc. Disord.* **2009**, *23*, 91–101. [[CrossRef](#)]
31. Dikmen, S.S.; Heaton, R.K.; Grant, I.; Temkin, N.R. Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. *J. Int. Neuropsychol. Soc.* **1999**, *5*, 346–356. [[CrossRef](#)]
32. O’Caoimh, R.; Molloy, D.W. The Quick Mild Cognitive Impairment Screen (Qmci). In *Cognitive Screening Instruments; A Practical Approach*, 2nd ed.; Larner, A., Ed.; Springer-Verlag: London, UK, 2017; Volume 1, pp. 255–272.
33. Cunje, A.; Molloy, D.W.; Standish, T.I.; Lewis, D.L. Alternate forms of logical memory and verbal fluency tasks for repeated testing in early cognitive changes. *Int. Psychogeriatr.* **2007**, *19*, 65–75. [[CrossRef](#)]
34. O’Caoimh, R.; Timmons, S.; Molloy, D.W. Screening for mild cognitive impairment: Comparison of “MCI specific” screening instruments. *J. Alzheimer’s Dis.* **2016**, *51*, 619–629. [[CrossRef](#)]
35. Leung, J.L.; Lee, G.T.; Lam, Y.H.; Chan, R.C.; Wu, J.Y. The use of the Digit Span Test in screening for cognitive impairment in acute medical inpatients. *Int. Psychogeriatr.* **2011**, *23*, 1569–1574. [[CrossRef](#)]
36. St. Clair-Thompson, H.L.; Allen, R.J. Are forwards and backwards digit recall the same? A dual task study of digit recall. *Mem. Cognit.* **2013**, *41*, 519–532. [[CrossRef](#)]
37. O’Keeffe, E.; Mukhtar, O.; O’Keeffe, S.T. Orientation to time as a guide to the presence and severity of cognitive impairment in older hospital patients. *J. Neurol. Neurosurg. Psychiatry* **2011**, *82*, 500–504. [[CrossRef](#)]
38. Pirozzo, S.; Papinczak, T.; Glasziou, P. Whispered voice test for screening for hearing impairment in adults and children: Systematic review. *BMJ* **2003**, 327. [[CrossRef](#)]
39. Sullivan, K. Estimates of interrater reliability for the Logical Memory subtest of the Wechsler Memory Scale-Revised. *J. Clin. Exp. Neuropsychol.* **1996**, *18*, 707–712. [[CrossRef](#)]
40. Gavett, B.E.; Ashendorf, L.; Gurnani, A.S. Reliable change on neuropsychological tests in the Uniform Data Set. *J. Int. Neuropsychol. Soc.* **2015**, *21*, 558–567. [[CrossRef](#)]
41. Larner, A.J. Speed versus accuracy in cognitive assessment when using CSIs. *Prog. Neurol. Psychiatry* **2015**, *19*, 21–24. [[CrossRef](#)]

42. Teng, E.L.; Hasegawa, K.; Homma, A.; Imai, Y.; Larson, E.; Graves, A.; Sugimoto, K.; Yamaguchi, T.; Sasaki, H.; Chiu, D.; et al. The Cognitive Abilities Screening Instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia. *Int. Psychogeriatr.* **1994**, *6*, 45–58. [[CrossRef](#)] [[PubMed](#)]
43. Cullen, B.; O'Neill, B.; Evans, J.J.; Coen, R.F.; Lawlor, B.A. A review of screening tests for cognitive impairment. *J. Neurol. Neurosurg. Psychiatry* **2007**, *78*, 790–799. [[CrossRef](#)] [[PubMed](#)]
44. Subbe, C.P.; Kruger, M.; Rutherford, P.; Gemmel, L. Validation of a modified early warning score in medical admissions. *QJM* **2001**, *94*, 521–526. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).