Comparison of novel tools with traditional cognitive tests in detecting delirium in elderly medical patients

David J Meagher, Henry O’Connell, Maeve Leonard, Olugbenga Williams, Fahad Awan, Chris Exton, Michael Tenorio, Margaret O’Connor, Colum P Dunne, Walter Cullen, John McFarland, Dimitrios Adamis

BACKGROUND
Efficient detection of delirium and comorbid delirium-dementia is a key diagnostic challenge. Development of new, efficient delirium-focused methods of cognitive assessment is a key challenge for improved detection of neurocognitive disorders in everyday clinical practice.

AIM
To compare the accuracy of two novel bedside tests of attention, vigilance and visuospatial function with conventional bedside cognitive tests in identifying delirium in older hospitalized patients.

METHODS
180 consecutive elderly medical inpatients (mean age 79.6 ± 7.2; 51% female) were referred to psychiatry for later life consultation-liaison service with delirium, dementia, comorbid delirium–dementia and cognitively intact controls. Participants were assessed cross-sectionally with conventional bedside cognitive tests [WORLD, Months Backward test (MBT), Spatial span, Vigilance A and B, Clock Drawing test and Interlocking Pentagons test] and two novel cognitive tests [Lighthouse test, Letter and Shape Drawing test (LSD)-4].

RESULTS
Neurocognitive diagnoses were delirium (n = 44), dementia (n = 30), comorbid delirium-dementia (n = 60) and no neurocognitive disorder (n = 46). All
conventional tests had sensitivity of > 70% for delirium, with best overall accuracy for the Vigilance-B (78.3%), Vigilance-A (77.8%) and MBT (76.7%) tests. The sustained attention component of the Lighthouse test was the most distinguishing of delirium (sensitivity 84.6%; overall accuracy 75.6%). The LSD-4 had sensitivity of 74.0% and overall accuracy 74.4% for delirium identification. Combining tests allowed for enhanced sensitivity (> 90%) and overall accuracy (≥ 75%) with the highest overall accuracy for the combination of MBT-Vigilance A and the combined Vigilance A and B tests (both 78.3%). When analyses were repeated for those with dementia, there were similar findings with the MBT-Vigilance A the most accurate overall combination (80.0%). Combining the Lighthouse-SA with the LSD-4, a fail in either test had sensitivity for delirium of 91.4 with overall accuracy of 74.4%.

CONCLUSION
Bedside tests of attention, vigilance and visuospatial ability can help to distinguish neurocognitive disorders, including delirium, from other presentations. The Lighthouse test and the LSD-4 are novel tests with high accuracy for detecting delirium.

Key words: Visuospatial function; Attention; Vigilance; Letter and Shape Drawing test; Lighthouse test; Delirium; Dementia; Phenomenology; Assessment

©The Author(s) 2020. Published by Baishideng Publishing Group Inc. All rights reserved.
across 22 clinical studies. Similarly, a review of 16 studies of the Clock Drawing Test (CDT) in delirium identified 11 different scoring methods\[^{14}\]. As such, more systematized and reliable methods of cognitive testing are needed, either based upon existing tests or by developing novel approaches to testing that draw upon modern (e.g., computer-assisted/smartphone) technologies.

To this end, the Cognitive Impairment Research Group at the University of Limerick developed two novel tests specifically for the purpose of detecting cognitive difficulties that characterize delirium: The Lighthouse test, which focuses on attention/vigilance and the Letter and Shape Drawing test (LSD-4), which focuses upon visuospatial abilities\[^{15-17}\]. This study the accuracy of a series of commonly used conventional tests as well as these novel tests in the detection of delirium in a real world sample of older hospital medical patients. Specifically, the aims were to (1) compare performance on these different tests in identifying delirium in elderly inpatients with varying neurocognitive disorder profiles as well as those with normal cognition; (2) examine how they compare (both individually and in combination) in terms of their specificity, sensitivity, positive and negative predictive value in detecting delirium and comorbid delirium-dementia in older medical inpatients; and (3) explore how accuracy is impacted upon by comorbid dementia.

**MATERIALS AND METHODS**

**Subjects and design**

A cross-sectional study of cognitive performance was conducted in referrals to a consultation-liaison psychiatry service of patients with delirium, dementia, comorbid delirium-dementia, as well as comparison subjects with no neurocognitive diagnosis (NNCD). Consecutive cases with altered mental state were identified on daily rounds by the medical team and referred for assessment and diagnosis by the research team. Assessments were conducted by raters (Leonard M, Awan F, O’Connell H, Williams O, Meagher D) specifically trained in the use of the tests included herein (see below) and to further enhance inter-rater reliability, ratings associated with any uncertainty were discussed and agreed by consensus between raters.

Patients were assessed during the usual working day and in the majority of cases the process of receiving referrals and responding meant that this occurred in the early afternoon when the anchors of the day are thought to be optimally active. The assessments were conducted at the bedside to mimic real world practice.

Delirium was diagnosed according to a cut-off score of ≥ 15 on the severity scale of the Delirium Rating Scale-Revised-98 (DRS-R98)\[^{18}\] and/or presence of DSM IV criteria\[^{19}\] based upon a full clinical assessment. This approach was used because it allows for high diagnostic specificity in populations that include substantial numbers of patients with comorbid dementia. Dementia was defined as a clear history of documented DSM-IV\[^{19}\] dementia (based on all available information at the time of assessment including clinical case notes and collateral history from family and/or carers) or a short Informant Questionnaire on Cognitive Decline in the elderly (IQCODE) score of ≥ 3.5\[^{20}\]. Comorbid delirium-dementia was defined as the presence of both disorders.

Each subject was assessed with a battery of nine conventional (Rater A) and 2 novel (Rater B) tests (see below). Standard cut off performances were used to apply a binary (pass/fail) for each test where a fail corresponded with evidence of clinically significant impairment. Assessors were not aware of the patients’ formal neurocognitive diagnoses. The cognitive tests were conducted in a fixed sequence as described below.

**Informed consent**

The procedures and rationale for the study were explained to all patients but because many patients had cognitive impairment at entry into the study it was presumed that many might not be capable of giving informed written consent. Because of the non-invasive nature of the study, University Hospital Limerick Regional Ethics Committee approved (REC 100/12) an approach to establishing consent by virtue of augmenting patient assent with proxy consent from next of kin (where possible) or a responsible caregiver for all participants in accordance with the Helsinki Guidelines for Medical Research involving human subjects\[^{21}\].

**Assessments**

Demographic data and medication at the time of the assessment were recorded. All available information from medical records and collateral history was used. Nursing staff were interviewed to assist rating of symptoms over the previous 24 h.
The DRS-R98 is designed for broad phenomenological assessment of delirium. It is a 16-item scale with 13 severity and 3 diagnostic items with high interrater reliability, sensitivity and specificity for detecting delirium in mixed neuropsychiatric and other hospital populations. Each item is rated 0 (absent/normal) to 3 (severe impairment) with descriptions anchoring each severity level. Severity scale scores range from 0-39 with higher scores indicating more severe delirium. Delirium typically involves scores above 15 points (Severity scale) or 18 points (Total scale) when dementia is in the differential diagnosis.

The IQCODE-SF is a validated screening tool for detecting cognitive impairment. The short version of the IQCODE includes 16 items that rate cognitive change over time, each of which are rated by an informant on a 5 point Likert scale. The total score divided by the number of questions provides a mean item score where ratings ≥ 3.5 are considered indicative of longstanding cognitive difficulties and dementia.

**Cognitive testing**

**WORLD backwards**: The WORLD backwards test was applied according to the Mini-Mental State Examination. Each participant was asked to spell WORLD backwards. Patients who self-corrected their own mistakes without prompting when spelling WORLD backwards were not penalized. Failure to correctly recite all five letters is considered to equate with clinically significant inattention (and thus a failed test).

**MBT**: In this test, the participant was asked to recite the months of the year in reverse order starting from December. Test duration was a maximum of 90 seconds at which point the subject’s best performance was noted. Scoring in subjects over age 60 is that failure to reach July without more than one error of omission equates with clinically significant inattention (and thus a failed test).

**Spatial span forwards**: This was conducted according to the description in the Cognitive test for delirium (CTD). The Spatial span forwards is a visual form of the digit span forwards. The subject is asked to copy the examiner in touching squares on a card (A5 size with 8 cm × 1 cm red squares). Each square represents a number and the test on each occasion requires that the squares corresponding to the digit span code are tapped at one second intervals. Two trials are conducted and the best performance is used. Failure to correctly complete a sequence of 5 or more numbers is considered to equate with clinically significant inattention (and thus a failed test). This was also conducted according to the description in the original description of the CTD.

**Spatial span backwards**: Similarly, the Spatial span backwards uses squares (blue) that are repeated in reverse order to that indicated by the assessor. Two trials are conducted and the best performance is used. Failure to correctly complete a sequence of three or more numbers is considered to equate with clinically significant inattention (and thus a failed test). This was also conducted according to the original description of the CTD.

**Vigilance A test**: The vigilance “A” test was also derived from the CTD scale. A list of 29 letters with the letter “A” included on 11 occasions was presented to the patient and they were asked to indicate each time the letter “A” was mentioned. Scores are calculated by subtracting commissions from correct responses (scored double) and rated as unable to engage with the test (0), score 1-9 (1), score 10-18 (2), score 19-26 (3), score > 27 (4). For the purposes of a binary (pass/fail) cutoff, we used failure to score > 27 to equate with significantly impaired vigilant (or sustained) attention.

**Vigilance B test**: This is similar to the vigilance A test except that there are two required letters (“C” and “E”). Scores are calculated by subtracting commissions from correct responses (scored double) and rated as unable to engage with the test (0), score 1-9 (1), score 10-18 (2), score 19-26 (3), score > 27 (4). For the purposes of a binary (pass/fail) cutoff, we used failure to score ≥ 19 to equate with significantly impaired vigilant (or sustained) attention.

**Global assessment of visuospatial abilities**: Visuospatial ability was rated according to a four point scale based upon DRS-R98 item 13 using both patient and collateral sources regarding ability to perceive differences in shape and distance as well as practical abilities such as navigating the ward environment and response to specific probes of describing how to get somewhere (e.g., bathroom), recognising shapes (“what shape is the window?”) and differences in distances (“which is closer the window or the hallway?”). Patients with moderate or greater impairment in terms of responses to probes and/or reported need for redirection to keep from getting lost in the environment or difficulty locating familiar objects in immediate environment were considered to have pathological impairment of visuospatial abilities (failed test).
**Intersecting Pentagons test:** This geometric copying test is derived from the original Bender Gestalt test\(^\text{[25]}\). The subject is presented with a copy of two intersecting pentagons drawn at angles to one another producing a diamond shape where they overlap. The subject is requested to copy the design on the blank half of the page. For scoring, we applied the six-point hierarchical scoring scale where 6 represents a perfect reproduction and 1 represents the poorest effort with scores < 4 equated with a failed performance\(^\text{[26]}\).

**CDT:** The CDT examines visuospatial abilities as well as receptive language, numerical knowledge, working memory, and executive functions. It is widely used in geriatric practice as a cognitive scan. In this study, subjects were provided with a pre-drawn circle onto which the participant was requested to place all the numbers and the large and small hands on the clock face to show the time “ten past eleven”. We used the scoring method of Sunderland et al\(^\text{[27]}\) (1989) rating performance from 0 to 10 according to spatial representation of the numbers and hands of the clock. A score of < 6 equates with a failed performance.

**LSD:** The LSD is a novel test designed to emphasize visuospatial abilities\(^\text{[15]}\). It consists of a series of 4 designs that link 1cm spheres arranged in increasingly complex grids that the subject copies to an adjacent blank grid. The complexity ranges from very simple (copying an “X” on a 3 by 3 grid which is thought to assess awareness and basic understanding of the test procedures as well as physical ability to engage with testing) to increasingly more complex figures (e.g., on a 6 by 6 grid) designed to challenge those with higher levels of cognitive ability (Figure 1). A correct performance requires that all relevant spheres are connected to complete the required shape. Omissions (but not commissions) are rated as errors. Subjects are permitted a single trial of each of the 4 items. Each item is scored 0 or 1 depending on whether all target zones on the grid are completed, allowing for a total score ranging from 0-4. Scores less than 3 are considered to reflect clinically significant impairment of performance and equate with a failed performance\(^\text{[17]}\). The test typically takes 1-2 min to complete. The test can be presented either as pen and paper or digitalised formats. The latter can allow for delivery of a more versatile test (that can be readily adapted to individual characteristics such as impaired visual or motor skills)\(^\text{[16]}\). The LSD thus provides a brief and easily interpreted bedside test of visuospatial function.

**The Lighthouse test:** The Lighthouse test was developed by the Cognitive Impairment Research Group as an objective assessment of awareness, focused and sustained attention. It is administered using an Android smartphone and involves presentation of an image of a flashing lighthouse on a standard screen (3” x 5”) (Figure 2). The test has 3 main sections; (1) assessing whether the subject recognizes the lighthouse as such; (2) assessing the subject’s capacity to focus attention to describe the number of times the lighthouse flashes (×3 sequences; 4, 3, 5). Subjects are requested to identify the number of flashes; and (3) testing the capacity to sustain attention to count sequences of flashes (×3) (i.e. 4-3-2, 3-2-5, 2-4-3) that last 12-15 seconds in duration. Again, subjects are requested to identify the total number of flashes.

**Statistical analyses**

Statistical analysis was conducted using SPSS-19\(^\text{[28]}\). Continuous data are presenting as means plus standard deviation. Categorical data are presented as counts and percentages. When multiple comparisons were conducted (ANOVA) the Bonferroni correction for multiple comparisons was used. The accuracy of tests of cognition (and their combinations), sensitivity and specificity as well as positive and negative likelihood ratio, Positive Predictive Value, and Negative Predictive Value were calculated, with confidence intervals testing significance at 95%. Post hoc power calculation for the main research question (the ability of LSD-4 to detect delirium against no delirium) was performed by using the G*Power v3.1.2. software. With a = 0.05, effect size = 0.5 and df = 3, a sample size of 180 indicated power of almost 1 (0.99).

**RESULTS**

A total of 180 patients were assessed [mean age 79.6 ± 7.2; 91 (51%) female]. The frequencies of neurocognitive diagnoses were; delirium (n = 44), dementia (n = 30), comorbid delirium-dementia (n = 60) and NNCD (n = 46). Demographic, medication, and general clinical data for these four groups are shown in Table 1. There were no significant differences between the four groups in respect of age, gender distribution.
Table 1 compares mean scores for the four groups for the DRS-R98 and IQ-CODE. Both delirium groups were more impaired than the dementia and NNCD groups on total scores for the DRS-R98. For the short IQCODE mean scores both dementia groups scored well above the suggested cut-off score and higher than the delirium-only and NNCD groups.

Table 2 and Table 3 show the performance on the conventional cognitive tests for the four neurocognitive groups, including accuracy for delirium diagnosis in the overall group as well as within the group with diagnosed dementia (n = 90) (Figure 3). Of note, all tests of attention and vigilance had a sensitivity for delirium of > 70% but, in terms of overall accuracy, the Vigilance A and B and MBT were the best performing tests. The tests of visuospatial ability were less sensitive to delirium, with the global assessment of visuospatial abilities (GVS) performing slightly better than the CDT and Intersecting Pentagons test (IPT).

Tables 4 and 5 show the performance on the three components of the Lighthouse test and the LSD-4. Overall, only one third of patients were able to identify the lighthouse correctly, with one third unable to describe it while the other third described it in a variety of ways including as “a lampost”, “a traffic light”, “a chimney” and “the Eiffel tower”. The identification and focused attention components of the Lighthouse had substantially lower sensitivity and overall accuracy compared to the sustained attention (LH-SA) element and when combined with the LH-SA did not substantially add to its accuracy. The LH-SA alone compared favourably with the conventional tests of attention. Similarly, the LSD-4 compared favourably with the conventional tests of visuospatial abilities in terms of sensitivity and overall accuracy for detecting delirium, especially in those patients with dementia.

We examined the accuracy of various combinations of the conventional tests and the LH-SA/LSD in detecting delirium in the overall group and in the dementia group. The better performing combinations (i.e. those with sensitivity > 90% and overall accuracy ≥ 75%) are shown in Table 6. For the overall group, the MBT-GVS and the combined Vigilance A and B tests were the most sensitive combination (93.3%), while the MBT-Vigilance A and the combined Vigilance A and B tests had the highest overall accuracy (78.3%), with the LH-SA/LSD combination demonstrating similar levels of accuracy. When these analyses were repeated for the population with DSM-IV defined dementia, there were similar findings with the MBT-Vigilance A the most accurate overall combination (80.0%), followed by the LH-SA/LSD and combined Vigilance A and B tests (both 77.8%).

**DISCUSSION**

Performance on bedside tests of attention and visuospatial ability was compared in elderly medical inpatients with a variety of neurocognitive diagnoses and also with normal cognition. Participants were carefully diagnosed using a full neuropsychiatric assessment with well-validated instruments. Patients with active delirium (both with and without comorbid dementia) were distinguished from patients with dementia-alone in respect of performance on simple bedside tests. Moreover, combining tests of attention with visuospatial ability allowed for greater accuracy of delirium detection. Two novel cognitive tests, the LSD test both compare favourably with conventional tests and may offer advantages for use in everyday practice.
One important implication of this study is that formal testing using any conventional test can assist delirium detection – all tests were quite sensitive to the presence of delirium but the Vigilance A and B and the MBT were the best individual tests in terms of overall accuracy. This is in keeping with previous studies that have included direct comparisons of cognitive tests in the identification of delirium in elderly general hospital inpatients and which have consistently found that bedside tests of attention (including sustained or vigilant attention) are sensitive to the presence of delirium, with the Months Backward Test emerging as the most versatile individual test[10,29-33].

The Lighthouse is a novel test that includes three components designed to assess awareness/comprehension, focused attention and sustained attention. Somewhat surprisingly, only one third of subjects could correctly identify the Lighthouse thus, raising the possibility that the visual graphics are suboptimal. Although the ID and FA components did not individually add to the accuracy of the Lighthouse test for delirium, the testing procedures involved engaging with the stimulus and simple testing and may thus have optimised arousal and attention for the sustained attention component.

The LSD performed well in terms of delirium detection, demonstrating greater accuracy than the conventional visuospatial bedside tests (CDT and IPT), especially in those patients with dementia. Previous work has emphasised visuospatial function as a cognitive function that, along with attention, is particularly affected in delirium[6-10]. However, other work suggests that conventional tests such as the CDT lack specificity for delirium compared to dementia[14,32-34]. In contrast, the LSD which has been designed with the aim of optimising delirium-relevance, evidenced better specificity for delirium.

Efforts to identify optimal bedside cognitive testing for delirium monitoring should recognise that combining two tests that focus upon different aspects of cognition that are impaired in delirium can enhance accuracy of testing as well as inform delirium diagnosis which requires evidence of generalised disturbance to brain function. This should include impaired attention with deficits in at least one other cognitive domain[10] - visuospatial functioning offers a suitable second domain. In terms of accuracy, this work suggests that combining two tests can achieve > 90% sensitivity with high overall test accuracy. Combining the MBT with the Vigilance-A test is a particularly useful approach using conventional tests, while the combination of the LSD-4 with the sustained attention component of the Lighthouse test offers a novel approach that has similar accuracy. The latter has the advantage of being delivered by smartphone/tablet technology which can allow for enhanced consistency and reliability in test administration and interpretation. Moreover, digital technology offers the prospect of developing testing procedures that can be readily adapted according to individual patient characteristics such as visual acuity, frailty and motor dexterity – for example by altering the dimensions of presented material and/or the size of target zones on the LSD-4.

The LSD-4 and the Lighthouse are designed to emphasize consistency of administration and ease of interpretation. The methods applied in this study allowed for highly consistent administration procedures and detailed scoring systems that are not typically applied in everyday practice when using conventional bedside tests. Recent reviews[15,34] emphasize that conventional tools such as the MBT and the CDT are subject to considerable variability in use, with a lack of consensus as to optimal methods of administration and interpretation. As such, the accuracy of the conventional tests is likely to be lower in real world use. In contrast, the Lighthouse
Table 1  Demographic and clinical data for the total group and four neurocognitive groups (mean + SD)

|                          | Total group (n = 180) | Delirium (n = 44) | Comorbid delirium-dementia (n = 60) | Dementia (n = 30) | No neurocognitive disorder (n = 46) |
|--------------------------|-----------------------|-------------------|-------------------------------------|------------------|-----------------------------------|
| Female (%)               | 51%                   | 53%               | 56%                                 | 37%              | 48%                               |
| Age                      | 79.6 ± 7.2            | 77.7 ± 8.4        | 80.1 ± 7.0                          | 81.8 ± 4.6       | 79.3 ± 7.5                        |
| Total number of medications | 10.1 ± 4.7          | 10.0 ± 4.2        | 10.3 ± 5.0                          | 9.7 ± 4.4        | 10.3 ± 5.1                        |
| Number of psychotropics  | 1.7 ± 1.6             | 1.9 ± 1.6         | 2.2 ± 1.7                           | 1.7 ± 1.6        | 0.8 ± 1.1                         |
| DRS-R98 total 12         | 17.0 ± 9.2            | 22.9 ± 5.7        | 24.0 ± 6.0                          | 10.9 ± 4.4       | 6.5 ± 3.3                         |
| Short IQCODE3            | 3.7 ± 0.7             | 3.1 ± 0.1         | 4.4 ± 0.5                           | 4.0 ± 0.5        | 3.1 ± 0.1                         |

1Delirium and comorbid delirium-dementia > dementia at P < 0.001.
2Delirium-only, dementia-only and comorbid delirium-dementia groups > no neurocognitive diagnosis (NNCD) at P < 0.001.
3Dementia and comorbid delirium-dementia groups > delirium and NNCD groups at P < 0.001. DRS-R98: Delirium Rating Scale-Revised-98; IQCODE: Informant Questionnaire on Cognitive Decline in the elderly.

Test and the LSD-4 are more likely to maintain the accuracy evident herein due to their presentation in computerized format which enhances consistency of delivery and scoring and which may be associated with relatively less reduction in accuracy when used in everyday practice. In addition, we expect that the Lighthouse and LSD will be less subject to language-related inaccuracies than many other tests because they do not emphasize verbal skills. Future work can examine these issues, including the relative accuracy of computerized forms.

The combination of simple tests can allow for rapid and efficient assessment of delirium-relevant cognitive domains and achieved a sensitivity of almost 90% for delirium presence with these cross-sectional assessment methods. Serial monitoring of performance on these tests as a “cognitive vital sign” could allow for highly consistent detection of delirium in real world practice. Moreover, presentation in computerized formats could make for highly systematized assessment procedures that, given the modest specificity of 55%, would ideally be enhanced by a second phase of assessment for patients who identify as positive. This two-step approach to delirium detection is increasingly advocated as an effective means of improving detection rates in everyday clinical practice[5,36]. It is important to note that although identifying cognitive impairment is central to delirium diagnosis, actual diagnosis requires that the timing (relatively acute onset) and context (a deterioration from usual baseline, not better explained by another neuropsychiatric condition and due to a physical etiology) also be determined. Tools such as the confusion assessment method[37] and DRS-R98[18] incorporate these additional considerations to allow for formal diagnosis. Ultimately, systematized cognitive testing is key to delirium screening efforts and can also be used to support the cognitive assessment that is inherent to formal diagnosis. Psychometric data to guide the choice of test in particular settings is relatively lacking but ultimately the choice of cognitive testing tool is determined by a variety of factors that relate to patient, tester and other resource issues that are particular to the healthcare environment. Further work exploring the impact of these factors on the efficiency of providing cognitive-friendly healthcare is needed to guide choice of testing methods across settings.

Study limitations
This work has some notable shortcomings which include (1) We studied consecutive referrals to a consultation-liaison service for assessment of neuropsychiatric status. As such, these patients are likely to have a heightened symptom burden and are not representative of elderly inpatients in general; (2) We applied binary cut off ratings for each of the tests based upon best convention but for many tests a clear and consistently agreed pass/fail distinction is lacking; (3) We used a fixed order for presentation of the tests which may have influenced performance due to changing levels of arousal during the testing process and with the competing effects of practice versus fatigue[38]; and (4) We did not specify the stage or primary cause of dementia or take account of clinical subtypes of delirium (i.e. hypoactive, hyperactive and mixed motor subtype) although evidence indicates that neurocognitive disturbance varies across dementia types and severity[39].

Implications
Improved identification of major neurocognitive disorders is a key healthcare
Table 2. Performance on the Conventional bedside cognitive tests for the overall population and for each of the neurocognitive diagnostic groups (number completing correctly and %)

| Test                              | Overall group (n = 180) | Delirium (n = 44) | Comorbid delirium-dementia (n = 60) | Dementia (n = 30) | NNCD (n = 46) |
|-----------------------------------|------------------------|-------------------|-------------------------------------|-------------------|--------------|
| World backwards test              | 42 (23)                | 6 (14)            | 5 (8)                               | 6 (20)            | 25 (54)      |
| Months backwards test             | 74 (41)                | 12 (28)           | 8 (13)                              | 20 (66)           | 34 (74)      |
| Spatial span forwards             | 68 (38)                | 14 (32)           | 12 (20)                             | 13 (43)           | 29 (63)      |
| Spatial span backwards            | 67 (37)                | 15 (34)           | 8 (13)                              | 15 (50)           | 29 (63)      |
| Vigilance A test                  | 88 (49)                | 16 (36)           | 10 (17)                             | 23 (77)           | 39 (85)      |
| Vigilance B test                  | 49 (27)                | 3 (7)             | 3 (5)                               | 13 (43)           | 30 (65)      |
| Global visuospatial test          | 72 (40)                | 9 (20)            | 15 (25)                             | 12 (40)           | 37 (80)      |
| Clock drawing test                | 78 (43)                | 16 (36)           | 13 (22)                             | 14 (47)           | 35 (76)      |
| Interlocking pentagons test       | 87 (48)                | 17 (39)           | 14 (23)                             | 18 (60)           | 38 (83)      |

NNCD: No neurocognitive diagnosis.

Challenge. In particular, accurate and consistent detection of delirium is a priority because evidence indicates that more than half of cases are missed or detected late in everyday practice, with implications for morbidity, length of stay in hospital and mortality. A fundamental factor in enhancing recognition rates is to identify simple and brief methods for establishing the presence of clinically significant cognitive impairment at the bedside. Although both delirium and dementia involve generalised disturbance of cognitive function, delirium can be distinguished by virtue of the disproportionate impairment of attention and visuospatial ability. These cognitive domains can be readily assessed in everyday clinical practice using simple bedside tests. Both the Lighthouse and the LSD-4 provide accurate and delirium-oriented means of assessing cognitive function in delirium and in combination achieve a sensitivity of over 90% for delirium detection. Their impact upon delirium detection in everyday practice warrants further study as we seek to develop more efficient delirium monitoring in everyday practice.
**Table 3** Accuracy of conventional bedside tests of cognition for delirium diagnosis in the overall population (n = 180) and for those with dementia (n = 90)

| Test          | Overall population (n = 180) | Dementia population (n = 90) |
|---------------|-------------------------------|------------------------------|
|               | Sensitivity (95%CI) | Specificity (95%CI) | Positive predictive value (95%CI) | Negative predictive value (95%CI) | Overall accuracy (95%CI) |
| WORLD         | 89.4% (81.9–94.6) | 40.8% (29.7–52.7) | 67.4% (62.9–71.6) | 73.8% (60.2–84.0) | 68.9% (61.6–75.6) |
| MBT           | 80.8% (71.9–87.4) | 71.1% (59.5–80.9) | 79.3% (72.8–84.6) | 73.0% (64.0–80.4) | 76.7% (69.8–82.6) |
| SSF           | 75.0% (65.6–83.0) | 55.3% (43.4–67.0) | 69.6% (64.0–75.1) | 61.8% (52.3–70.5) | 66.7% (59.3–73.5) |
| SSB           | 77.9% (68.7–85.4) | 57.9% (46.0–69.1) | 71.7% (65.6–77.1) | 65.7% (56.0–74.2) | 69.4% (62.2–76.1) |
| Vigilance A   | 75.0% (65.6–83.0) | 81.6% (71.0–89.6) | 84.8% (77.4–90.1) | 70.5% (62.7–77.2) | 77.8% (71.0–83.6) |
| Vigilance B   | 94.2% (87.9–97.9) | 56.6% (44.7–67.9) | 74.8% (69.6–79.4) | 87.8% (78.3–94.1) | 78.3% (71.8–84.1) |
| GVS           | 76.9% (67.6–84.6) | 63.2% (51.3–73.9) | 74.1% (67.6–79.6) | 66.7% (57.5–74.7) | 71.1% (63.9–77.6) |
| CDT           | 72.1% (62.5–80.5) | 64.5% (52.7–75.1) | 73.5% (66.7–79.4) | 62.8% (54.3–70.6) | 68.9% (61.6–75.6) |
| IPT           | 70.2% (60.4–78.8) | 73.7% (62.3–83.1) | 78.5% (71.1–84.4) | 64.4% (56.6–71.4) | 71.7% (64.5–78.1) |
|               |                  |                    |                  |                  |                  |
|               |                  |                    |                  |                  |                  |

MBT: Months Backward test; SSF: Spatial span forwards; SSB: Spatial span backwards; GVS: Global assessment of visuospatial abilities; CDT: Clock Drawing test; IPT: Intersecting Pentagons test.

**Table 4** Performance on Letter and Shape Drawing test and Lighthouse components (number completing correctly and %)

| Test          | Overall group (n = 180) | Delirium (n = 44) | Comorbid delirium-dementia (n = 60) | Dementia (n = 30) | NNCD (n = 46) |
|---------------|-------------------------|-------------------|-------------------------------------|-------------------|---------------|
|               | Sensitivity (95%CI) | Specificity (95%CI) | Positive predictive value (95%CI) | Negative predictive value (95%CI) | Overall accuracy (95%CI) |
| LH-identification | 60 (33)                | 16 (36)            | 11 (18)                            | 10 (33)           | 24 (52)       |
| LH-FA         | 85 (46)                | 18 (41)            | 14 (23)                            | 20 (66)           | 31 (67)       |
| LH-SA         | 64 (36)                | 9 (21)             | 7 (12)                             | 18 (60)           | 30 (65)       |
| LSD           | 84 (48)                | 17 (39)            | 10 (17)                            | 19 (63)           | 38 (83)       |

NNCD: No neurocognitive diagnosis; LSD: Letter and Shape Drawing test; LH: Lighthouse.

**Table 5** Accuracy of Lighthouse components and Letter and Shape Drawing test for delirium diagnosis in the overall population (n = 180) and in those with dementia (n = 90)

| Test          | Overall population (n = 180) | Dementia population (n = 90) |
|---------------|-------------------------------|------------------------------|
|               | Sensitivity (95%CI) | Specificity (95%CI) | Positive predictive value (95%CI) | Negative predictive value (95%CI) | Overall accuracy (95%CI) |
|               |                  |                  |                  |                  |                  |
| LH-ID         | 74.0% (64.5–82.1) | 44.7% (33.3–6.6) | 64.7% (59.2–69.8) | 55.7% (45.5–65.5) | 61.7% (54.1–68.8) |
| LH-FA         | 69.2% (59.4–77.9) | 67.1% (55.4–77.5) | 74.2% (67.1–80.3) | 61.5% (53.4–68.9) | 68.3% (61.0–75.1) |
| LH-SA         | 84.6% (76.2–90.9) | 63.2% (51.3–73.9) | 75.9% (69.8–81.0) | 75.0% (64.9–82.9) | 75.6% (68.6–81.6) |
| LSD           | 74.0% (64.5–82.1) | 75.0% (63.7–84.2) | 80.2% (73.0–85.9) | 67.9% (59.8–75.0) | 74.4% (67.4–80.6) |
|               |                  |                    |                  |                  |                  |
|               |                  |                    |                  |                  |                  |

Meagher DJ et al. New tools for delirium detection
LH-FA 76.7% (64.0-86.6) 66.7% (47.2-82.7) 82.1% (73.1-88.6) 58.8% (45.8-70.7) 73.3% (63.0-82.1)
LH-SA 88.3% (77.4-95.2) 60.0% (40.6-77.3) 81.5% (73.8-87.4) 72.0% (54.7-84.6) 78.9% (69.0-86.8)
LSD 83.3% (71.5-91.7) 63.3% (43.9-80.1) 82.0% (73.7-88.1) 65.5% (50.4-78.1) 76.7% (66.6-84.9)

LSD: Letter and Shape Drawing test; LH: Lighthouse.

| Overall population (n = 180) | Sensitivity (95% CI) | Specificity (95% CI) | Positive predictive value (95% CI) | Negative predictive value (95% CI) | Overall accuracy (95% CI) |
|-----------------------------|---------------------|---------------------|-------------------------------|-------------------------|------------------------|
| MBT plus GVS                | 93.3% (86.6-97.3)   | 48.7% (37.0-60.4)   | 71.3% (66.5-75.7)             | 84.1% (71.4-91.8)       | 74.4% (67.4-80.6)      |
| Vig A plus GVS              | 92.3% (85.4-96.6)   | 57.9% (46.0-69.1)   | 75.0% (69.6-79.7)             | 84.6% (73.3-91.7)       | 77.8% (71.0-83.6)      |
| Vig A plus Vig B            | 93.3% (86.6-97.3)   | 57.9% (46.0-69.1)   | 75.2% (69.9-79.9)             | 86.3% (75.0-93.0)       | 78.3% (71.6-84.1)      |
| MBT plus Vig A              | 90.4% (83.0-95.3)   | 61.8% (50.0-72.8)   | 76.4% (70.7-81.3)             | 82.5% (71.8-89.7)       | 78.3% (71.6-84.1)      |
| LSD-4 plus Lighthouse-SA    | 91.4% (84.2-96.0)   | 51.3% (39.6-63.0)   | 72.0% (66.9-76.5)             | 81.3% (69.1-89.4)       | 74.4% (67.4-80.6)      |

| Dementia population (n = 90) | Sensitivity (95% CI) | Specificity (95% CI) | Positive predictive value (95% CI) | Negative predictive value (95% CI) | Overall accuracy (95% CI) |
|------------------------------|---------------------|---------------------|-------------------------------|-------------------------|------------------------|
| MBT plus GVS                | 93.3% (83.8-98.1)   | 30.0% (14.7-49.4)   | 72.7% (67.6-77.3)             | 69.2% (43.0-87.0)       | 72.2% (61.8-81.2)      |
| Vig A plus GVS              | 95.0% (86.1-98.9)   | 33.3% (17.3-52.8)   | 74.0% (68.7-78.7)             | 76.9% (49.8-91.8)       | 74.4% (64.2-83.1)      |
| Vig A plus Vig B            | 95.0% (86.1-98.9)   | 43.3% (25.5-62.6)   | 77.0% (70.9-82.2)             | 81.3% (57.2-93.4)       | 77.8% (67.8-85.9)      |
| MBT plus Vig A              | 95.0% (86.1-98.9)   | 50.0% (31.3-68.7)   | 79.2% (72.6-84.5)             | 83.3% (61.1-94.1)       | 80.0% (70.3-87.7)      |
| LSD-4 plus LH-SA            | 95.0% (86.1-99.0)   | 43.3% (25.5-62.6)   | 77.0% (70.9-82.2)             | 81.3% (57.2-93.4)       | 77.8% (67.8-85.9)      |

MBT: Months Backward test; GVS: Global assessment of visuospatial abilities; CDT: Clock Drawing test; IPT: Intersecting Pentagons test; LH: Lighthouse; LSD: Letter and Shape Drawing test.

Figure 3 Receiver operating characteristic analyses depicting the accuracy of conventional bedside tests of cognition for delirium diagnosis in the overall population (n = 180). ROC: Receiver operating characteristic analyses.

ARTICLE HIGHLIGHTS

Research background
Efficient detection of delirium and comorbid delirium-dementia is a key diagnostic challenge. It’s a key challenge of developing of new, efficient delirium-focused methods of cognitive assessment for improved detection of neurocognitive disorders in everyday clinical practice.

Research motivation
This study the accuracy of a series of commonly used conventional tests as well as these novel tools for delirium detection.
tests in the detection of delirium in a real world sample of older hospital medical patients.

**Research objectives**

The authors aimed to compare the accuracy of two novel bedside tests of attention, vigilance and visuospatial function with conventional bedside cognitive tests in identifying delirium in older hospitalized patients.

**Research methods**

This cognitive performance study was conducted in referrals to a consultation-liaison psychiatry service of patients with delirium, dementia, comorbid delirium-dementia, as well as comparison subjects with no neurocognitive diagnosis. Altered mental state consecutive cases were identified on daily rounds.

**Research results**

All conventional tests had sensitivity of > 70% for delirium, with best overall accuracy for the Vigilance-B, Vigilance-A and Months Backward tests. The sustained attention component of the Lighthouse Test was the most distinguishing of delirium.

**Research conclusions**

Vigilance and visuospatial ability can help to distinguish neurocognitive disorders, including delirium, from other presentations. The Lighthouse test, Letter and Shape Drawing test are novel tests with high accuracy for detecting delirium.

**Research perspectives**

Lighthouse test, Letter and Shape Drawing tests’ impact upon delirium detection in everyday practice warrants further study.

**ACKNOWLEDGEMENTS**

This work was supported by a research project grant from the Health Research Board (HRA 2011/48).

**REFERENCES**

1. Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA* 2010; 304: 443-451 [PMID: 20664035 DOI: 10.1001/jama.2010.1013]

2. Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. *J Am Geriatr Soc* 2011; 59 Suppl 2: S241-S243 [PMID: 22091567 DOI: 10.1111/j.1532-5415.2011.03671.x]

3. Collins N, Blanchard MR, Tookey A, Sampson EL. Detection of delirium in the acute hospital. *Age Ageing* 2010; 39: 131-135 [PMID: 19917632 DOI: 10.1093/ageing/afp201]

4. Ryan DJ, O'Regan NA, Caonih RÔ, Clare J, O'Connor M, Leonard M, McFarland J, Tighe S, O'Sullivan K, Trzepacz PT, Meagher D, Timmons S. Delirium in an adult acute hospital population: predictors, prevalence and detection. *BMJ Open* 2013; 3: e001772 [PMID: 23299110 DOI: 10.1136/bmjopen-2012-001772]

5. O'Hanlon S, O'Regan N, MacLullich AM, Cullen W, Dunne C, Exton C, Meagher D. Improving delirium care through early intervention: from bench to bedside to boardroom. *J Neurol Neurosurg Psychiatry* 2014; 85: 207-213 [PMID: 23355807 DOI: 10.1136/jnnp-2012-304334]

6. Meagher DJ, Leonard M, Donnelly S, Conroy M, Saunders J, Trzepacz PT. A comparison of neuropsychiatric and cognitive profiles in delirium, dementia, comorbid delirium-dementia and cognitively intact controls. *J Neurol Neurosurg Psychiatry* 2010; 81: 876-881 [PMID: 20587481 DOI: 10.1136/jnnp.2009.200956]

7. Jabbar F, Leonard M, Meehan K, O'Connor M, Cronin C, Reynolds P, Meaney AM, Meagher D. Neuropsychiatric and cognitive profile of patients with DSM-IV delirium referred to an old age psychiatry consultation-liaison service. *Int Psychogeriatr* 2011; 23: 1167-1174 [PMID: 21253153 DOI: 10.1017/S1041610210002333]

8. Brown LJ, Fordyce C, Zaghdani H, Starr JM, MacLullich AM. Detecting deficits of sustained visual attention in delirium. *J Neurol Neurosurg Psychiatry* 2011; 82: 1334-1340 [PMID: 20587493 DOI: 10.1136/jnnp.2009.200956]

9. Mattoo SK, Grover S, Chakravarty K, Trzepacz PT, Meagher DJ, Gupta N. Symptom profile and etiology of delirium in a referral population in northern india: factor analysis of the DR5-R9. *J Neuropsychiatry Clin Neurosci* 2012; 24: 95-101 [PMID: 22450619 DOI: 10.1176/appi.neuropsych.11001009]

10. Leonard M, Mclnerney S, McFarland J, Condon C, Awan F, O'Connor M, Reynolds P, Meaney AM, Adams D, Dunne C, Cullen W, Trzepacz PT, Meagher DJ. Comparison of cognitive and neuropsychiatric profiles in hospitalised elderly medical patients with delirium, dementia and comorbid delirium-dementia. *BMJ Open* 2016; 6: e009212 [PMID: 26956160 DOI: 10.1136/bmjopen-2015-009212]

11. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3rd ed. Washington: American Psychiatric Association, 1980

12. Leonard M, O'Connell H, Awan F, Exton C, O'Connor M, Adams D, Dunne C, Cullen W, Meagher DJ. Attention, vigilance and visuospatial function in hospitalized elderly medical patients: Relationship to neurocognitive diagnosis. *J Psychosom Res* 2016; 80: 84-90 [PMID: 27772564 DOI: 10.1016/j.jpsychores.2016.09.011]

13. Meagher J, Leonard M, Donoghue L, O'Regan N, Timmons S, Exton C, Cullen W, Dunne C, Adams D, MacLullich AJ, Meagher D. Months backward test: A review of its use in clinical studies. *World J...
Meagher DJ et al. New tools for delirium detection

Psychiatry 2015; 5: 305-314 [PMID: 26425444 DOI: 10.5498/wjp.v5.i3.305]

Meagher D, Williams OA, O'Connell H, Leonard M, Cullen W, Dunne CP, Mulligan O, Adamis D. A systematic review and meta-analysis of the accuracy of the clock drawing test (CDT) in the identification of delirium in older hospitalised patients. Aging Ment Health 2020; 1-10 [PMID: 32991236 DOI: 10.1080/13607863.2020.1727849]

White D, Williams OA, Leonard M, Exton C, Adamis D, Hannigan A, Cullen W, Dunne CP, Meagher D. A pilot study of performance among hospitalised elderly patients on a novel test of visuospatial cognition: the letter and shape drawing (LSD) test. Ir J Psychol Med 2017; 34: 169-175 [PMID: 30151345 DOI: 10.1017/ipm.2016.15]

Tenorio M, Williams OA, Leonard M, White D, McKenna F, Dunne CP, Meagher D, Exton C. The Letter and Shape Drawing (LSD) test: An efficient and systematised approach to testing of visuospatial function. Conf Proc IEEE Eng Med Biol Soc 2016; 2016: 2523-2526 [PMID: 26268791 DOI: 10.1109/EMBC.2016.7591195]

Williams OA, O'Connell H, Leonard M, Awan F, White D, McKenna F, Hannigan A, Cullen W, Exton C, Emadi W, Dunne C, Adamis D, Meagher D. Development of the four-item Letter and Shape Drawing test (LSD-4): A brief bedside test of visuospatial function. Psychiatry Res 2017; 247: 317-322 [PMID: 27951480 DOI: 10.1016/j.psychres.2016.12.004]

Trezpaz PT, Mittal D, Torres R, Kanary K, Norton J, Imerson N. Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. J Neuropsychiatry Clin Neurosci 2001; 13: 229-242 [PMID: 11449630 DOI: 10.1176/jnp.13.2.229]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. Washington: American Psychiatric Association, 2000

Jorm AF. The Informant Questionnaire on cognitive decline in the elderly (IQCIDE): a review. Int Psychogeriatr 2004; 16: 275-293 [PMID: 15559753 DOI: 10.1017/S1041610204000390]

World Medical Association. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. Tokyo, Japan: 55th WMA General Assembly, 2004. Available from: http://www.wma.net/en/30publications/10policies/b3/...
