Submission for MD

University of Limerick, Graduate Entry Medical School

Student name: Henry O’Connell

ID number: 12110213

Title:
Detection and diagnosis of delirium: a study of the application of computer-assisted technologies.
# Table of Contents

List of tables ......................................................................................................................... 4  
List of figures ......................................................................................................................... 5  
Abbreviations ........................................................................................................................ 6  
Acknowledgements ................................................................................................................. 7  
Assessments and scales used ................................................................................................. 8  
    Rater A assessments ........................................................................................................... 8  
    Rater B assessments ........................................................................................................... 8  
Abstract .................................................................................................................................. 9  
    Background ....................................................................................................................... 9  
    Aims .................................................................................................................................. 10  
    Methods ............................................................................................................................ 10  
    Results .............................................................................................................................. 11  
    Conclusions ...................................................................................................................... 12  
Thesis outline .......................................................................................................................... 14  
Chapter 1: Background and hypotheses ............................................................................... 15  
    Delirium: evolution of the concept definition ................................................................... 15  
    Delirium: epidemiology, risk factors and costs ................................................................. 21  
    Delirium: neuropathogenesis ......................................................................................... 28  
    Delirium: assessment ....................................................................................................... 32  
    Android applications (apps) for cognitive assessment: background literature to the development of the Lighthouse app ........................................................................... 37  
    Delirium: management and prevention ......................................................................... 50  
Study hypotheses .................................................................................................................... 53  
Chapter 2: Systematic review of bedside tests of attention .................................................. 54  
Methods .................................................................................................................................. 56  
Results .................................................................................................................................... 57  
    Summaries of studies ....................................................................................................... 64  
Discussion ............................................................................................................................... 72  
    Tests of attention ............................................................................................................ 74  
    Differentiating delirium and dementia ......................................................................... 75  
    Post-operative delirium ................................................................................................. 76  
    Depression ...................................................................................................................... 76
‘Head-to-head’ studies........................................................................................................... 77
Younger populations.................................................................................................................. 78
Acceptability to patients and assessors.................................................................................. 78
Computerised tests.................................................................................................................. 80
Study limitations...................................................................................................................... 80
Conclusion............................................................................................................................... 82
Chapter 3: Methods.................................................................................................................. 95
Subjects and Design................................................................................................................ 95
Ethical approval....................................................................................................................... 96
Definitions for delirium and dementia..................................................................................... 96
Statistical analyses.................................................................................................................. 97
Rater A clinical assessments.................................................................................................... 98
Tests and rating scales............................................................................................................ 99
Cognitive testing...................................................................................................................... 103
Rater B assessments.............................................................................................................. 106
Lighthouse test........................................................................................................................ 107
Letter and Shape Drawing (LSD) Test................................................................................... 108
Chapter 4: Results.................................................................................................................... 109
Description of study subjects.................................................................................................. 109
Performance on ‘gold standard’ tests...................................................................................... 113
Performance on the Lighthouse app tests.............................................................................. 121
Lighthouse identification (LH-ID).......................................................................................... 121
Focussed Attention (FA) component of Lighthouse test ...................................................... 123
Sustained Attention (SA) component of Lighthouse test..................................................... 125
Concordance measures........................................................................................................ 131
Study hypotheses.................................................................................................................... 133
Chapter 5: Discussion.............................................................................................................. 135
‘Gold standard’ tests.............................................................................................................. 136
Lighthouse app tests.............................................................................................................. 139
Study limitations.................................................................................................................... 143
Implications............................................................................................................................ 144
References............................................................................................................................... 145
**List of tables**

Table 1: DSM 5 criteria for delirium (APA, 2013) .......................................................................................... 18
Table 2: DSM guidance notes: assessing attention ......................................................................................... 19
Table 3: ICD-10; F05 (WHO, 1992) ........................................................................................................... 20
Table 4: Risk factors for delirium in hospital ................................................................................................... 26
Table 5: Inflammatory markers and neurotransmitters associated with delirium ............................................. 30
Table 6: Lighthouse test .................................................................................................................................. 43
Table 7: Multicomponent non-pharmacological management of symptoms of delirium .............................. 51
Table 8: Studies categorised according to clinical setting ............................................................................. 58
Table 9: Summaries of cognitive tests used in included publications (tests listed in order of frequency studied) ........................................................................................................................................... 61
Table 10: Summaries of studies ...................................................................................................................... 83
Table 11: Downs and Black Tool parts 1-13 .................................................................................................. 90
Table 12: Downs and Black Tool parts 14-26 ............................................................................................... 92
Table 13: Recommendations for future research .......................................................................................... 94
Table 14: Demographic and medication data for the four neurocognitive groups (mean ± SD) ............... 110
Table 15: CDR scale scores for neurocognitive groups ................................................................................. 112
Table 16: Numbers from each neurocognitive group completing different components of the Lighthouse test ........................................................................................................................................ 112
Table 17: Cognitive Test scores for total group and neurocognitive disorder groups (mean ± SD; % engagement, and pass/fail frequency) ........................................................................................................... 115
Table 18: DRS-R98 Severity and cognitive test scores for patients with delirium (n=85) according to clinical subtype as per the Delirium Motor Subtype Scale (mean ± SD; % engagement, and pass/fail frequency) ........................................................................................................... 117
Table 19: Accuracy of individual tests of attention, vigilance and visuospatial function for the detection of DSM-IV delirium ........................................................................................................... 119
Table 20: Logistic Regression for predictors of DSM-IV delirium amongst tests of attention, vigilance and visuospatial function ........................................................................................................... 120
Table 21: Accuracy of combined tests of attention, vigilance and visuospatial function for the detection of DSM-IV delirium ........................................................................................................... 120
Table 22: LH-ID cross-tabulation for the 4 clinical groups ............................................................................. 121
Table 23: Accuracy of LH-ID test for identification of delirium .................................................................... 122
Table 24: FA test performance by neurocognitive group .............................................................................. 123
Table 25: FA test performance for identification for delirium ...................................................................... 124
Table 26: SA test performance by neurocognitive group ............................................................................. 125
Table 27: SA test performance for identification of delirium ...................................................................... 126
Table 28: Accuracy of different combinations of the individual Lighthouse tests ........................................ 127
Table 29: Overall test accuracy for individual and combined Lighthouse tests ............................................ 128
Table 30: LHIDFASA Failers compared to others (i.e. those who passed at least one component of the Lighthouse test – LHID, FA or SA) ........................................................................................................... 130
Table 31: Kappa values for DRS, CTD and Lighthouse test items ................................................................. 131
Table 32: Concordance for Lighthouse tests and other tests of cognition (kappa values) .......................... 132
Table 33: Future research ............................................................................................................................ 142
List of figures

Figure 1: Lighthouse app screenshot (beginning of flash emerging to right of image) .................. 45
Figure 2: Lighthouse app screenshot (flash further developed to left of image) ....................... 46
Figure 3: Lighthouse app (modified version: for testing in future research; larger flash image may improve clinical utility) ........................................................................................................ 47
Figure 4: Lighthouse app (modified version) .................................................................................. 48
Abbreviations

CSDD: Cornell Scale for Depression in Dementia

CSF: cerebrospinal fluid

CTD: Cognitive Test for Delirium

DRS-R98: Delirium Rating Scale revised version

FA: Focussed attention test

GABA: Gamma-aminobutyric acid

IFN: Interferon

IGF-1: Insulin-like Growth Factor

IL: Interleukin

IL-RA: Interleukin receptor antagonist

LH-ID: Lighthouse identification test

MBT: months backwards test

NICE: National Institute of Health and Clinical Excellence

NNCD: No Neurocognitive Disorder (i.e. individuals without delirium or dementia).

NIQD: Neuropsychiatric Inventory, Distress Scale

NIQS: Neuropsychiatric Inventory, Severity Scale

S100B: Calcium-binding protein.

SA: Sustained attention test

SSB: Spatial Span Forwards test

SSF: Spatial Span Backwards test
Acknowledgements

Dr. Maeve Leonard R.I.P. was the key individual involved in organising and driving the data collection phase of this project in 2014 and 2015. Maeve worked tirelessly to ensure that the project was a success. This thesis and the ensuing papers will form a small part of her vast research legacy. I miss her dearly and I feel very lucky and privileged to have worked with her.

Professor David Meagher was the other key individual involved in devising the research project and ensuring that this thesis reached completion. Since meeting David on my first day as a trainee psychiatrist in 1998 he has become a close friend and my most inspiring colleague.

Sophie, Isabelle and Henry are my shining lights and they helped me through the more tedious parts of this project, when I struggled with statistical analysis late into the night at our kitchen table.

My parents Henry and Nell, my brother Liam and my sister Anna have been my supports and companions through life’s journey and have also helped me through this latest leg.

Geraldine Harrington printed countless articles and drafts of this thesis while at the same time using her imagination, dedication and limitless energy to help keep my ‘day job’ and clinical service afloat during the life of this project.

Most importantly I want to thank my wife Kathy, without whom nothing would be possible and without whom nothing would seem worthwhile.
**Assessments and scales used**

**Rater A assessments**

Delirium Etiology Rating Checklist (Trzepacz and Meagher, 2008, Trzepacz et al., 2011)

Mini-Mental State Examination (Folstein et al., 1975)

Delirium Rating Scale-R-98 (DRS-R98) (Trzepacz et al., 2001)

Cognitive Test for Delirium (CTD) (Hart et al., 1996)

Short IQCODE (Jorm, 1994)

Neuropsychiatric Inventory (Kaufer et al., 2000)

Clinical Dementia Rating Scale (Hughes et al., 1982)

Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988)

Delirium Motor Subtype Scale 4 (DMSS-4) (Meagher et al., 2014)

**Rater B assessments**

Lighthouse Identification (LH-ID) test

Focussed Attention (FA) test

Sustained Attention (SA) test

Bouncing Balls (BB) test

Intersecting pentagons (Folstein et al., 1975, Bender, 1938)

Clock Drawing test (Freedman et al., 1994, Brodaty et al., 2002)

Letter Shape Drawing (LSD) test
Abstract

Background

Delirium is a common, severe and life threatening neuropsychiatric disorder. Encountered in all healthcare settings, delirium affects approximately one in five general hospital inpatients (Ryan et al., 2013, Siddiqi et al., 2006) with delirium incidence and prevalence rates likely to rise substantially in the coming years as healthcare services provide for our increasingly aged society. Delirium impacts negatively upon morbidity, length of stay in hospital, cognitive impairment and mortality (Witlox et al., 2010), along with substantial social and healthcare costs (Leslie and Inouye, 2011).

Despite its clinical importance in terms of frequency and impact, delirium is underdiagnosed and inadequately treated. This may be related to factors such as the complex and heterogenous clinical presentation of delirium, as well as the frequent mis-assumption that cognitive impairment is an expected and normal state for older inpatients, thus missing out on opportunities to identify potentially reversible and modifiable causative factors. This is further compounded by the lack of brief, objective and effective bed-side tests that are acceptable to patients and healthcare workers alike. At a service-planning level, the importance of delirium and cognitive friendly hospital initiatives are frequently neglected (O’Connell et al., 2014).

Prior to DSM III (APA, 1980) and ICD10 (WHO, 1992), the lack of operationalized diagnostic criteria for delirium hampered detection and diagnosis: the emphasis in DSM 5 (APA, 2013) on the critical importance of attentional deficits, especially deficits in sustained attention, has brought clarity to the area of delirium screening, diagnosis and assessment.
Aims

The aims of this study were, firstly, to describe in detail an older and neurocognitively mixed general hospital inpatient population and, secondly, to assess the utility of traditional and novel computerised smartphone based cognitive tests in the diagnosis of delirium. As part of the overall study, a systematic literature review was also conducted, focussing on bedside tests of attention used in the detection and diagnosis of delirium.

Methods

The systematic review was conducted using a structured search strategy in accordance with PROSPERO guidelines. The search focused on studies reporting the use of bedside tests of attention used in the diagnosis of delirium. The Downs and Black tool (Downs and Black, 1998) was used to assess the quality of publications.

For the clinical study, a sample of convenience was assessed, based on referrals to the research team by medical teams at University Hospital Limerick. Each participant underwent a two phase assessment. Rater A recorded medical history and medication and performed a standard battery of cognitive and neuropsychiatric tests. These included the Mini-Mental State Examination (MMSE) (Folstein et al., 1975), Delirium Etiology Checklist (Trzepacz and Meagher, 2008, Trzepacz et al., 2011), Delirium Rating Scale revised version (DRS-R98) (Trzepacz et al., 2001), Cognitive Test for Delirium (CTD) (Hart et al., 1996), Short-IQCODE (Jorm, 1994), Clinical Dementia Rating Scale (CDR) (Hughes et al., 1982), Neuropsychiatric Inventory (NPI-Q) (Kaufman et al., 2000), Cornell Scale for Depression in Dementia (CSDD) (Alexopoulos et al., 1988) and the four-item Delirium Motor Subtype Scale 4 (DMSS-4) (Meagher et al., 2014). Rater B assessments included a novel computerized test of cognitive function on a smartphone application, the Lighthouse test, which includes three elements (recognition, focused attention and sustained attention), the intersecting pentagons test (Folstein et al., 1975, Bender, 1938), clock drawing test (Freedman et al., 1994, Brodaty et al., 2002) and a novel test of visuospatial abilities, the Letter and Shape Drawing (LSD) test.
Results

The systematic review led to the identification of 23 suitable studies. Overall, a wide range of bedside tests were described, with variable levels of utility. The studies were conducted in a wide variety of clinical settings and twelve different neuropsychological tests of attention were included. The majority of tests were pen and paper and there was a notable lack of computerised tests. Likewise, there was a lack of focus on patient factors such as acceptability and clinical coverage of tests.

For the clinical study, subjects age 60 and over were included. 193 patients [mean age 79.9 ± 7.3; 97 male] were assessed with delirium (n=32), dementia (n=42), comorbid delirium-dementia (n=53) and no neurocognitive disorder (NNCD) (n=66). The ability to meaningfully engage with the tests varied from 84% (DSF) to 57% (Vig B test), and was especially problematic amongst the comorbid delirium-dementia group. The NNCD group was distinguished from the delirium groups for most tests, and from the dementia group for the Vigilance B test and the CDT, while the dementia group differed from the delirium groups in respect of the MBT, Vigilance A and B tests, Global assessment of visuospatial ability and the IPT. Overall, patients with delirium were best identified by three tests – the MBT, Vigilance A test and the Global Assessment of visuospatial function such that a pass on any two of these allowed for correct prediction of delirium status in 77% of cases. There were minimal differences in test performance across clinical (motor) subtypes of delirium.

The clinical utility of the Lighthouse smartphone application (app) was assessed in terms of ability to identify cases of delirium. For this study, the Lighthouse identification (LH-ID) test had sensitivity of 74% and specificity of 37.7% for the presence of delirium. The focussed attention (FA) test had sensitivity of 70% and specificity of 57%. The sustained attention (SA) test had sensitivity of 87.5% and specificity of 52.3%. When all three Lighthouse app tests were combined (LH-ID, FA and SA tests), any error had a sensitivity of 95% and specificity of 23.4%.
Conclusions

Ultimately, a brief portable means of assessing attention at the bedside can assist clinical staff in their efforts to reliably detect delirium in the general hospital setting. The approach to cognitive testing should be objective, user friendly and have high levels of clinical utility. The Lighthouse app could provide such a test, especially if combined with other delirium-relevant enquiry. The Lighthouse test application provided a brief (less than three minutes) and portable bedside test that allows for highly consistent presentation with easy interpretation that is user friendly for patients. Sensitivity of the Lighthouse test for detection of delirium is excellent at 95%, meaning that it has potential to be a useful screening test. However, in view of poor specificity, the Lighthouse test will likely need to be combined with other tests to enhance this aspect. In view of the complex phenomenological nature of delirium above and beyond deficits in attention, along with the fact that attentional impairment may be associated with other neuropsychiatric disorders aside from delirium, a brief test of attention such as the Lighthouse test is likely to serve as the first screening phase followed by more detailed neuropsychological testing in those identified as having deficits. Furthermore, future modifications to the Lighthouse test, in terms of visual presentation, may improve specificity and overall utility of the test. This test may also form a vital part of an overall cognitive friendly hospital programme whereby delirium is prevented, detected and managed promptly and effectively.
‘Were I a Roman Catholic, perhaps I should on this occasion vow to build a chapel to some saint, but as I am not, if I were to vow at all, it would be to build a light-house’.

Benjamin Franklin, in a letter to his wife in 1757, after he had narrowly avoided a shipwreck.
Thesis outline

This thesis is divided into five separate chapters. Chapter 1 covers background and hypotheses. Key background literature on delirium is summarised here, relating to definitions, epidemiology, risk factors, aetiology, neuropathogenesis, costs, assessment, management and prognosis. In addition, the recent literature on android applications in delirium assessment is summarised. The issue of computerised tests of attention in detection and diagnosis of delirium is addressed. The background to the development of the novel Lighthouse application (app) test is described. Chapter 2 involves a PROSPERO registered systematic review of bedside tests of attention used in the detection and diagnosis of delirium. The 23 studies included in this review are summarised and rated objectively using the Downs and Black tool. Specific issues discussed include testing to discriminate delirium from dementia, post-operative delirium, comorbid depression, ‘head to head’ studies comparing different tests of attention, research in younger populations and acceptability of tests to patients and assessors.

Chapter 3 covers study methodology, including descriptions of the study design, participant selection and the assessment protocol. Two separate raters (Rater A and Rater B) were involved in the assessment of all study participants and the individual assessments and tools used by the two raters are described. Rater A assessments are primarily ‘gold standard’ tests of cognition used in the detection and diagnosis of delirium. Rater B assessments include novel tests of cognition, in particular the Lighthouse app. Results are summarised in chapter 4, including the performance of ‘gold standard’ tests of cognition and the different components of the Lighthouse app in the detection and diagnosis of delirium. Key clinical and sociodemographic characteristics of the study population are summarised. Sensitivity, specificity, clinical coverage and overall test performance of the different cognitive tests are described.

Chapter 5 involves a discussion on the implications of the results, including study limitations and implications for future research.
Delirium: evolution of the concept definition

Delirium is a severe and often life-threatening neuropsychiatric disorder involving acute cognitive decompensation and behavioural changes arising in the context of systemic or cerebral illness (APA, 2013; WHO, 1992). The core clinical concept of acute generalised disturbance of brain function or ‘delirium’ has been described in various guises over the past two millennia (Adamis et al., 2007) and from earliest times it has been accepted as a grave condition with a wide range of clinical features and potentially fatal outcome.

The word delirium is derived from the Latin de-lira, i.e. to go out of the furrow or deviate from a straight line. Celsus was the first to use the term delirium, in the first century AD, describing a syndrome of mental disorder during fever or head trauma. Celsus also used the term phrenitis for this purpose, a term introduced by Hippocrates in 500 BC to describe mental disorder caused by fever, poisoning or head trauma. Hippocrates used approximately sixteen different terms for delirium, including leros, mania, paraleros and lethargus (Adamis et al., 2007). Furthermore, his description of phrenitis (acute onset of behavioural problems, sleep disturbances and cognitive deficits usually associated with fever) and lethargus (inertia and dulling of the senses) can be seen as analogous to currently used descriptions of the hyperactive and hypoactive motoric subtypes of delirium (Meagher and Trzepacz, 2000, Meagher et al., 2008, Meagher, 2009). In subsequent centuries, numerous physicians from various cultures described the concept of delirium (Adamis et al., 2007), including Procopius (who provided phenomenological descriptions based on a possible bubonic epidemic in Constantinople in AD 542), the Arab physician Najab ub din Unhammad and the Persian physician Rhazes. The term delirium was probably first used in English medical literature by Cosin in 1592. By the beginning of the 18th century, delirium was the term used to describe short-term madness or raving whereas the terms ‘phrensy’ or ‘phrenesis’ were used to describe a condition caused by fever or other physical illness (Adamis et al., 2007). Within phrensy and phrenesis, a distinction was acknowledged, with phrenesis being associated with brain inflammation and paraphrenesis associated with
inflammation of other organs. This early distinction can again be interpreted to anticipate modern aetiological and pathophysiological models for conceptualizing delirium, whereby delirium precipitants can be categorized as being due to ‘direct brain insults’ (e.g. hypoxia, stroke and drug effects) and ‘aberrant stress responses’ (e.g. referring to activation of the hypothalamic-pituitary-adrenal axis, sympathetic nervous system or inflammatory pathways) (Macullich et al., 2013). In 1794, Erasmus Darwin (grandfather of Charles) introduced disorientation and alteration of consciousness to distinguish delirium from ‘madness’ and in 1817 Greiner introduced the concept of clouding of consciousness as the primary feature of delirium, a concept that was further developed in the 1860s by Hughlings Jackson (Adamis et al., 2007). In the 20th century, the issue of clouding of consciousness was debated and further developed by Manfred Bleuler, Klaus Conrad and Kurt Schneider. Until the present time, a legion of terms and synonyms have been used to denote delirium, including ‘acute brain failure’, ‘acute confusional state’, ‘cerebral insufficiency’, ‘toxic psychosis’ and even ‘subacute befuddlement’ (Meagher, 2001). Rather than characterizing discrete scientific entities, these merely reflect the diverse phenomenology, aetiological factors and variety of clinical settings in which delirium is encountered. The DSM III (APA, 1980) represented an important step in the process of unifying these different terms and clinical features under the umbrella term ‘delirium’, while also operationalizing the definition. DSM-III (APA, 1980) identified the elusive concept of ‘clouding of consciousness’ as being central to delirium diagnosis. Acknowledging the difficulties with definition and clinical description of this concept, DSM-III-R (APA, 1987) and DSM-IV (APA, 1994) omitted the term ‘clouding of consciousness’ and emphasized the need to link impairments in consciousness to objective attentional deficits. In contrast to DSM, ICD10 (WHO, 1992) (Table 1) persists with the term ‘clouding of consciousness’.

DSM-5 (APA, 2013) (Table 1) has gone further again by omitting the term consciousness and emphasizing the importance of attentional impairment and awareness. Furthermore, DSM-5 (APA, 2013) states that delirium is not diagnosed ‘in the context of a severely reduced level of arousal such as coma’. Consciousness can be conceptualized in terms of level (relating to arousal) and content (relating to attention and comprehension). These concepts are hierarchically related in that a minimum level of arousal must be present in order for level of attention and comprehension to be measurable. For example, profound inattention can be
seen with full arousal (e.g. due to hypervigilance) but optimal attention cannot be attained with reduced levels of arousal.

This emphasis in DSM-5 of measurable attentional impairment has led the European Delirium Association (EDA) and American Delirium Society (ADS) to urge caution in the interpretation of DSM-5 guidelines. Criterion D in DSM-5 (Table 1) states that inattention or changes in cognition ‘must not occur in the context of a severely reduced level of arousal such as coma’. The EDA and ADS have called for a more inclusive definition of delirium that also captures those delirious patients with low levels of arousal and subsequent inability to engage with attentional tests, based on the argument that a narrow interpretation of Criterion D could mean that patients too drowsy to undergo cognitive testing cannot fulfil Criterion A or Criterion C (Table 1) (European Delirium Association, 2014). The EDA and ADS suggests that Criterion D should include ‘all states of altered arousal (except coma) in the spectrum of delirium on scientific, practical and clinical safety grounds’. 
Table 1: DSM 5 criteria for delirium (APA, 2013)

A. Disturbance in attention (i.e. reduced ability to direct, focus, sustain, and shift attention) and awareness (reduced orientation to the environment).

B. The disturbance develops over a short period of time (usually hours to a few days), represents an acute change from baseline attention and awareness, and tends to fluctuate in severity during the course of a day.

C. An additional disturbance in cognition (e.g. memory deficit, disorientation, language, visuospatial ability, or perception).

D. The disturbances in Criteria A and C are not better explained by a pre-existing, established or evolving neurocognitive disorder and do not occur in the context of a severely reduced level of arousal such as coma.

E. There is evidence from the history, physical examination or laboratory findings that the disturbance is a direct physiological consequence of another medical condition, substance intoxication or withdrawal (i.e. due to a drug of abuse or to a medication), or exposure to a toxin, or is due to multiple etiologies.
### Table 2: DSM guidance notes: assessing attention

Examples of assessments

- **Sustained attention:** Maintenance of attention over time (e.g., pressing a button every time a tone is heard, and over a period of time).

- **Selective attention:** Maintenance of attention despite competing stimuli and/or distracters: hearing numbers and letters read and asked to count only letters.

- **Divided attention:** Attending to two tasks within the same time period: rapidly tapping while learning a story being read. Processing speed can be quantified on any task by timing it (e.g. time to put together a design of blocks; time to match symbols with numbers; speed or serial 3 speed).

Note: The DSM notes above are not consistent with actual clinical practice within the field and seem like arbitrary descriptions of potential tests which lack clarity of guidance.
<table>
<thead>
<tr>
<th>Table 3: ICD-10; F05 (WHO, 1992)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delirium, not induced by alcohol and other psychoactive substances</td>
</tr>
</tbody>
</table>

A. Clouding of consciousness, i.e. reduced clarity of awareness of the environment, with reduced ability to focus, sustain, or shift attention.

B. Disturbance of cognition, manifest by both:
   1) Impairment of immediate recall and recent memory, with relatively intact remote memory;
   2) Disorientation in time, place or person.

C. At least one of the following psychomotor disturbances:
   1) rapid, unpredictable shifts from hypo-activity to hyper-activity;
   2) increased reaction time;
   3) increased or decreased flow of speech;
   4) enhanced startle reaction.

D. Disturbance of sleep or the sleep-wake cycle, manifest by at least one of the following:
   1) insomnia, which in severe cases may involve total sleep loss, with or without daytime drowsiness, or reversal of the sleep-wake cycle;
   2) nocturnal worsening of symptoms;
   3) disturbing dreams and nightmares which may continue as hallucinations or illusions after awakening.

E. Rapid onset and fluctuations of the symptoms over the course of the day.

F. Objective evidence from the history, physical and neurological examination or laboratory tests of an underlying cerebral or systemic disease (other than psychoactive substance-related) that can be presumed to be responsible for the clinical manifestations in A-D.

Comments: Emotional disturbance such as depression, anxiety or fear, irritability, euphoria, apathy or wondering perplexity, disturbances of perception (illusions or hallucinations, often visual) and transient delusions are typical but are not specific indications for the diagnosis.

Use the fourth character to indicate whether the delirium is superimposed on dementia or not:

F05.0 Delirium, not superimposed on dementia
F05.1 Delirium, superimposed on dementia
F05.8 Other delirium
F05.9 Delirium, unspecified
**Delirium: epidemiology, risk factors and costs**

In view of the wide variety of clinical contexts and patient profiles in which delirium is encountered, optimal screening and diagnostic tools should be brief, portable and have broad clinical coverage and penetration of different groups. Development of effective tools should also take into account clinical and patient factors such as frailty, mobility, sensory impairment, language/expression difficulties and neurocognitive problems aside from delirium, especially dementia.

The epidemiology of delirium in various clinical settings (community, palliative care, general hospital, emergency department, post operatively and intensive care unit) is described in the following sections.

**Delirium in community and palliative care settings**

Although little is known about delirium prevalence in community populations, reported prevalence rates are high in nursing homes (8.9%) and residential homes (8.2%) (Boorsma et al., 2012). Of all clinical settings, delirium is most prevalent in palliative care populations. Delirium is particularly common in palliative care settings, with reported prevalence rates of 13.3–88%, increasing from admission through to terminal stages (Hosie et al., 2013, Boettger et al., 2014). A cross-sectional study of 111 cancer patients with delirium found that severity of medical illness was not associated with severity of delirium. However, the more severely medically ill had increased disturbance of consciousness, disorientation and an inability to maintain and shift attention. The aetiologies found to be most commonly associated with delirium in severe medical illness were hypoxia and infection (Boettger et al., 2014).

A comprehensive review of epidemiology and risk factors for delirium (Vasilevskis et al., 2012) described four important clinical settings: the general hospital; the emergency department (ED); postoperative delirium (POD) and delirium in the intensive care unit (ICU) and the results of the review are summarised below.
Delirium in medical inpatients is common and is associated with multiple negative outcomes, including increased mortality, dementia, functional impairment, increased length of stay and increased levels of institutionalisation that are relatively independent of confounders such as age, sex, comorbid illness and baseline dementia (Siddiqi et al., 2006, Witlox et al., 2010).

A recent delirium point prevalence study in a general hospital (Ryan et al., 2013) demonstrated a rate of 20.7% for full delirium and 8.6% for subsyndromal delirium, in keeping with a generally accepted norm of 1 in 5 general hospital inpatients having delirium.

In their review, Vasilevskis et al (2012) report prevalent delirium for hospitalised older patients as 11-25% with incident delirium (i.e. delirium with onset after hospital admission) of 29-31%. A systematic review (Siddiqi et al., 2006) for medical inpatients reported prevalence at admission of 10-31%, an incidence rate of new delirium per admission ranging from 3-29% with occurrence rate per admission varying from 11-42%. Patient vulnerability factors for delirium in the general hospital setting include dementia, low education attainment, advanced age, high comorbidity burden, visual impairment, depression, history of alcohol abuse, malnutrition and prior home opioid or benzodiazepine use. Potentially modifiable factors in the general hospital include higher severity of illness, infections e.g. urinary or respiratory tract, dehydration, electrolyte abnormalities, acute kidney injury, liver failure, alcohol or benzodiazepine withdrawal, central nervous system insults, seizures, congestive heart failure and acute myocardial infarction. In addition, iatrogenic events associated with delirium include physical restraints, bladder catheters, polypharmacy and drugs (opioids, benzodiazepines and drugs with anticholinergic action). Lower levels of delirium detection are likely to be associated with increased mortality, with one study demonstrating an 11% increased mortality rate for every 48 hours of delirium (Gonzalez et al., 2009).
Delirium in the ED

Because of the rapidly changing nature of delirium, with resolution for up to half of patients within 24 hours, those classified as not having delirium in hospital may have had delirium in the ED that has since resolved. One study reported 19% of older ED patients discharged still having delirium (Han et al., 2011). A prospective study of older emergency department patients (Kakuma et al., 2003) found a statistically significant association between delirium and mortality after adjustment for age, gender, functional level, cognitive status, morbidity and number of medications for the first 6 months of follow-up (hazard ratio 7.24). Mortality rates were highest among those subjects whose delirium was not detected by emergency department staff (30.8%) whereas there was no significant difference in mortality for delirious versus non-delirious subjects when delirium was detected in the emergency department. Reported risk factors for delirium in the ED include dementia, functional impairment, hearing impairment, age, cerebrovascular disease, seizure disorder, increased respiratory rate, urinary tract infection and intracranial haemorrhage.

Postoperative delirium (POD)

The type of surgery has a bearing on delirium risk, with the following reported risks: otolaryngological (12%); general surgery (13%); aortic (up to 29%); major abdominal (up to 50%) and cardiac (up to 51%). Increased levels of surgical burden and associated risk of delirium are also seen in open versus endovascular surgery; emergency versus elective; increased blood transfusion and increased surgical duration. Specifically in cardiac surgery, balloon pump support, valve surgery and prolonged cardiopulmonary bypass all increase delirium risk while there is conflicting evidence regarding off-pump versus on-pump coronary bypass. It has also been reported that delirium most commonly appears and resolves within one week of surgery (Galanakis et al., 2001) and so this represents a key time for clinical vigilance with careful monitoring for emerging delirium and pre-emptive treatment strategies.

Patient vulnerability factors for POD include increasing age, pre-existing cognitive impairment, increased cardiovascular co-morbidities (including diabetes) and severity of
illness, male sex, smoking, obstructive sleep apnoea, alcohol excess, atrial fibrillation, renal dysfunction, heart failure, and postoperative ICU admission. In contrast, younger age and higher education have been reported to be protective factors in the development of POD and potentially modifiable risk factors include the use of benzodiazepines (reduced with use of dexmedetomidine versus propofol or midazolam). There is conflicting evidence on risks associated with the use of opioids in the postoperative period and on type of anaesthetic (i.e. general versus regional).

**Delirium in the ICU**

Reported prevalence rates for delirium in the ICU vary from 20-80% or more, with reports on incident delirium ranging from 22-83%. Variations in rates may be related to study design factors, patient characteristics, delirium measurement instruments, site-specific factors and coma (as delirium cannot be assessed or diagnosed in comatose patients).

Predisposing factors for delirium in the ICU again include cognitive impairment and increasing age. Precipitating factors include acute physiological derangements, exposure to benzodiazepines and opioid use while early mobility has been reported as a protective factor. Reduction of sedation and improved mobility may be achieved in the ICU through use of the ABCDE bundle, relating to **Awakening and Breathing Coordination, Choice of Sedation, Delirium monitoring and Management and Early Mobility** (Balas et al., 2012).

Inflammatory markers (procalcitonin and C-reactive protein) and metabolites of tryptophan have been reported to be associated with incident delirium in the ICU, along with environmental factors such as inadequate access to light and interactions with friends or family.

In considering risk factors and aetiology for delirium, a useful general model involves the interaction of patient vulnerability factors (e.g. older age, presence of pre-existing cognitive impairment or dementia and higher comorbidity burden) with illness factors (ranging from e.g. mild peripheral infections to sepsis). This concept of ‘delirium readiness’ has been well described for decades (Henry and Mann, 1965). Bearing this in mind, relatively milder illness
factors may be required in the aetiology of delirium in those with high vulnerability factors, and vice versa (Inouye and Charpentier, 1996, O'Keeffe and Gosney, 1997a, Kalisvaart et al., 2006). Furthermore, delirium risk factors can be considered in terms of the following factors: patient-related, illness-related, medication-related, environment-related and intervention-related. These factors are summarised and listed in Table 4 (O'Connell et al., 2014).
## Table 4: Risk factors for delirium in hospital

From O’Connell et al (2014)

### Patient related
- Age ≥ 70 years
- Pre-existing cognitive impairment
- Previous episode of delirium
- CNS disorder
- Increased blood-brain barrier permeability
- Poor nutritional status
- Number and severity of comorbid illnesses

### Illness-related
- Illness severity
- Dehydration
- Infection, e.g. urinary tract infection
- Fracture
- Hypothermia or fever
- Hypoxia
- Metabolic/electrolyte disturbances, e.g. low sodium
- Pain
- HIV/AIDS
- Organ insufficiency
- Burns
- Nicotine withdrawal

### Medication-related
- Polypharmacy
- Drug or alcohol dependence
- Benzodiazepine use
- Addition of ≥ 3 new medications
- Psychoactive drug use
- Certain drugs: e.g. anticholinergics

### Environment-related
- Social isolation
- Sensory extremes
- Visual deficit
- Hearing deficit
- Immobility
- Use of restraints
- Novel environment
- Stress

### Intervention-related
- Peri-operative
- Type of surgery, e.g. hip, cardiac
- Duration of operation
- Catheterisation
- Emergency procedure
Costs of delirium

The economic costs of delirium are extensive and can be conceptualised broadly in terms of immediate and long-term human and economic costs to the individual, to healthcare systems and to society in general and overall healthcare costs have been demonstrated as being doubled in delirious versus non-delirious patients (Leslie et al., 2008). The economic cost of a single admission of a patient with delirium in the UK has been estimated at £13,000 (Akunne et al., 2012) and between $16,000-64,000 in the US (Leslie et al., 2008). The estimated direct 1-year healthcare cost of delirium in the US is 143-152 billion US Dollars, second only to the cost of cardiovascular disease and ahead of the combined costs for diabetes, non-fatal falls and hip fracture (Leslie and Inouye, 2011).
Delirium: neuropathogenesis

Although phenomenologically complex and heterogenous in clinical presentation, delirium can be thought of as a final common pathway signifying brain failure as a result of various illnesses and physiological insults.

A recent review (Maclullich et al., 2013) outlined a number of current areas of enquiry in the pathogenesis of delirium: how peripheral infections elicit brain dysfunction; the dysfunction causing delirium itself and the process underpinning the association of delirium with long-term cognitive decline. Aetiological factors for delirium are considered in two broad categories: ‘direct brain insults’ (e.g. stroke, drug effects, metabolic abnormalities and hypoxia) and ‘aberrant stress responses’. For the latter group of factors, delirium is seen as arising as a result of the exaggerated and harmful impact of physiological systems such as the systematic nervous system, hypothalamic-pituitary-adrenal axis, inflammatory pathways and other systems involved in response to acute threat, including relatively minor conditions such as mild peripheral infection. The review summarises the range of potential inflammatory markers involved in delirium in human studies (see Table 5).

The use of neuropsychological markers (such as inattention) as measures of delirium proneness is a relatively understudied area. Such measures may be useful not only in the prediction of incident delirium but also in monitoring the response to interventions of established cases. Computerised measures offer advantages over traditional pen and paper tests, in view of increased objectivity. A prospective study evaluating pre- and post-operative neuropsychological performance in 100 dementia-free elective orthopaedic admissions age 70 and over (Lowery et al., 2007) found that pre-operative attentional deficits were closely associated with post-surgical delirium. Therefore, sensitive tests of attentional impairment are likely to have a key role in the early detection of incident delirium.

Pre-existing cognitive impairment is a well-recognised risk factor for delirium, with comorbid dementia evident in more than 50% of delirium in the hospitalised elderly. Furthermore, more subtle impairments of neuropsychological functioning are associated with increased delirium risk in those without dementia. These include deficits in tests of attention,
vigilance, visuospatial function, graphomotor speed and executive function (Greene et al., 2009, Lowery et al., 2007).

Cognitive performance in the immediate post-operative period also predicts delirium likelihood (Sharma et al., 2005). Furthermore, Lowery et al (2008) found that impaired cognitive performance is common in the postoperative period in those who do not develop delirium, although impairments differed with non-delirious patients experiencing a decrease in vigilance whereas delirious patients were impaired in the accuracy of attention.

These findings emphasise the need to specify the character and extent of acute cognitive deterioration that equates with delirium since impairment of cognitive performance is common during periods of high morbidity or exposure to interventions that confer increased delirium risk, including many who do not develop syndromal delirium. These findings also highlight how formal neuropsychological testing can identify patients at risk of delirium for whom preventative actions are especially recommended.
**Table 5: Inflammatory markers and neurotransmitters associated with delirium**

**From Maclullich et al (2013)**

Exogenous administration of cytokines

- Elevated cortisol (in blood and CSF) (delirium possibly mediated by pathologically elevated cortisol in context of acute stress from illness or surgery)
- Elevated endogenous levels of serum interleukin (IL)-6 and IL-8 (serum and CSF)
- Levels of IFN- (pro-inflammatory), IL-RA and IGF-1 (anti-inflammatory) altered during delirium.
- Elevated CSF S100B: a marker of CNS damage derived from astrocytes

Neurotransmitters (implicated due to drug effects):

- Dopamine (increased dopaminergic signalling in CSF of delirium patients with psychotic features)
- Cholinergic system
- Adrenergic system
- GABAergic system

Abbreviations used in Table 5:

CSF: cerebrospinal fluid; IL: Interleukin; IFN: Interferon; IL-RA: Interleukin receptor antagonist; IGF-1: Insulin-like Growth Factor; GABA: Gamma-aminobutyric acid; S100B: Calcium-binding protein.
The review by Maclullich et al (2013) also describes animal studies in delirium. Animal studies have considerable potential because delirium is acute, severe and measurable and we have established information on precipitants and vulnerability factors.

Animal studies of delirium have involved induction of delirium using atropine (Leavitt et al., 1994) and lipopolysaccharide (LPS) (Semmler et al., 2008) and proposed areas of future research will focus on interleukins, gamma-aminobutyric acid (GABA) and acetylcholine.

More recent studies have used animal models involving relatively minor inflammatory insults superimposed on prior brain vulnerability. This reflects the clinical reality whereby delirium in older patients involves underlying dementia in 50% of cases. Therefore, studies have been conducted in aged mice (Chen et al., 2008), in animals with pre-existing neurodegenerative pathology (Murray et al., 2012), in animals with microglia primed by primary pathology (Griffin et al., 2013) and in animals with selective lesions of the basal forebrain cholinergic system (Field et al., 2012). Overall, delirium animal research is pointing towards a role in delirium pathogenesis for altered cholinergic activity on a background of neurodegenerative pathology.
Delirium: assessment

Delirium is frequently misdiagnosed, detected late or totally missed across clinical settings (Ryan et al., 2013, Inouye et al., 2001, Elie et al., 2000). In a point-prevalence study of 311 patients in a general hospital, the treating medical team missed over half of delirium cases (Ryan et al., 2013) and 46% of delirium diagnoses were missed by the referring team in a study of general hospital referrals to liaison psychiatry (Kishi et al., 2007). Detection rates are lower in older patients (Cole, 2004), those with premorbid dementia (Inouye et al., 2001) and in hypoactive cases (Fang et al., 2008, Voyer et al., 2007). This contributes to additional suffering, increased medical and cognitive morbidity and additional healthcare costs associated with increased length of stay and subsequent institutionalisation. An 11% increase in mortality with every additional 48 hours of delirium has been demonstrated by Gonzalez et al (2009). Therefore, effective screening and assessment systems are vital and should be an essential component of cognitive friendly delirium oriented hospital systems (O’Connell et al., 2014).

Routine cognitive and delirium screening should be performed on all older patients in the general hospital, especially those with background cognitive impairment, dementia or previous episodes of delirium. The guidelines of the National Institute of Healthcare Excellence (NICE, 2010) advocates routine daily screening for delirium in hospital inpatients, even in those not deemed to be at risk. An efficient approach would involve an initial screening phase using brief and sensitive tools such as the Confusion Assessment Method (Inouye et al., 1990) or the Nursing Delirium Screening Scale, NuDESC (Gaudreau et al., 2005). The development of computerised screening instruments that can be used at the bedside would also be very useful in this phase as they can allow for more accurate and efficient assessment of key delirium-relevant aspects of brain function, such as attention. Such instruments should have high levels of clinical utility and acceptability to patients and clinical staff.

Following this screening phase, diagnosis is clarified by more detailed and expert assessment, if necessary involving the opinion of Geriatric or Old Age Psychiatry Consultation-Liaison services. This two-step approach to detection and diagnosis has been advocated by a number of expert groups in recent years (Meagher and Leonard, 2008, Hall
et al., 2012). Management should then be instituted rapidly, with a focus on establishing and treating underlying medical causes such as infection or cerebral problems. Collateral history should be obtained to clarify baseline cognitive function. Once delirium is identified and appropriate non-pharmacological and pharmacological measures have been instituted, the patient should have brief cognitive assessments performed regularly in order to monitor response to interventions, with the NICE guidelines recommending daily assessment.

There are a variety of factors relevant to the poor detection of delirium. Firstly, the clinical and neuropsychiatric features of delirium are poorly understood by many healthcare professionals with a tendency to think of delirium as a mere symptom of other illness that does not therefore warrant attention in its own right. In addition, delirium is sometimes considered as synonymous with ‘Delirium Tremens’ (akin to hyperactive delirium), thus de-emphasising the hypoactive presentations that form the majority of delirium cases. There is also a tendency to consider delirium as an inevitable consequence of ageing and frailty and not in itself a key target for prevention and assertive treatment. Other common neuropsychiatric conditions in older people, such as dementia and depression, may also overlap and mask delirium and vice versa. Detection rates are lower in older patients, those with comorbid dementia, hypoactive presentations, a previous history of psychiatric disorder and those receiving care in surgical settings (O’Hanlon et al., 2014).

In addition, another key factor in delirium under detection is the lack of brief, reliable and objective assessment tools (Kean and Ryan, 2008, Lowery et al., 2010). In contrast, many bedside tests for delirium require considerable training and expertise and rely on subjective measures of cognition and behaviour. Furthermore, while cases of hyperactive delirium are relatively easy to identify due to agitation and behavioural disturbance, the more common presentation of hypoactive delirium is often missed. As a result, it has been suggested that regular cognitive assessment become a ‘vital sign’ in basic hospital care (Maclullich et al., 2013). The new National Early Warning Score in the UK incorporates a 4-point level of consciousness measure (RCo., 2012) and there is evidence suggesting that reduced level of consciousness is a sensitive (but not specific) sign of delirium (Tieges et al., 2013b, Chester et al., 2012).
An ideal assessment tool should thus be brief, portable, easy to use for assessor and patient alike and give objective and reproducible results. The tool should also focus on key cognitive domains that are strongly associated with delirium and that have relative specificity for delirium over other neuropsychiatric conditions such as dementia and depression.

In a prospective study of 37 patients age 60-85 years undergoing cardiac surgery (Brown et al., 2011b), cognitive assessments were performed 0-8 days before surgery and again 2-9 days after surgery. The concepts of crystallized and fluid cognition were explored, with fluid cognition relating to the active processing of mental representations and crystallized intelligence relating to well-learnt word pronunciation knowledge. Fluid intelligence measures (involving digit span test, verbal fluency and Stroop tests) were impaired in the patients with incident postoperative delirium. In contrast, the measure of crystallized intelligence (National Adult Reading Test, NART) was unchanged in those with incident postoperative delirium. The authors conclude that, despite showing extensive deficits of fluid cognitive processing, crystallized cognition is preserved in delirium. Therefore, screening and detection tests for delirium should focus on tests of fluid cognition.

Attention has been highlighted as a key cognitive domain in delirium assessment, especially the area of sustained attention (O'Keeffe and Gosney, 1997a, Hart et al., 1996). When developing a specific test for that cognitive domain, interference from other cognitive domains, e.g. executive function and memory, should be minimised as much as possible.

The work of the Cognitive Impairment Research Group (CIRG) has led to important conclusions on the primary role of attentional impairment in the phenomenology of delirium. In a study of 100 people with delirium (Meagher et al., 2007) assessed using the Delirium Rating Scale-Revised-98 (DRS-R98) (Trzepacz et al., 2001) and the Cognitive Test for Delirium (CTD) (Hart et al., 1996), sleep-wake cycle abnormalities and inattention were the most frequent while disorientation was the least frequent cognitive deficit in delirium. Inattention was also found to be associated with severity of other cognitive disturbances but not with non-cognitive items. CTD comprehension correlated most closely with non-cognitive features of delirium. The authors conclude that delirium phenomenology is consistent with broad dysfunction of higher cortical centres, characterised in particular by
inattention and sleep-wake cycle disturbance with attention and comprehension together being the cognitive items that best account for the syndrome of delirium.

Further CIRG research compared delirium phenomenology occurring in patients with delirium referred to an Old Age Psychiatry Consultation-Liaison setting to a population of palliative care patients with delirium (Jabbar et al., 2011). Assessments were conducted using the DRS-R98 and CTD. 80 patients were included, 40 of whom had comorbid dementia. Again, inattention (in 100%) was found to be the most prominent cognitive disturbance, while sleep-wake cycle disturbance (98%), altered motor activity (97%) and thought process abnormality (96%) were the most frequent DRS-R98 non-cognitive features. The authors concluded that delirium is a complex neuropsychiatric syndrome with generalized cognitive impairment and disproportionate inattention.

The CIRG has also conducted longitudinal research on delirium, with one study focusing on twice weekly assessments of 100 consecutive patients with DSM-IV delirium in a palliative care setting (Leonard et al., 2015). Delirium was assessed using the DRS-R98 (Trzepacz et al., 2001) and CTD (Hart et al., 1996) and a mixed-effects regression model was employed to estimate changes in severity of individual symptoms over time, resulting in a total of 323 assessments (range of 2-9 per case). Frequency and severity of individual DRS-R98 symptoms was very consistent over time, with DRS-R98 items for attention (88-100%), sleep-wake cycle disturbance (90-100%) and any motor disturbance (87-100%) and CTD attention and vigilance most frequently and consistently impaired. Again the authors conclude that attention is disproportionately impaired during the entire episode of delirium, consistent with thalamic dysfunction underlying both an impaired state of consciousness and well-known EEG slowing. All individual symptoms and the three core domains remained relatively stable despite small fluctuations in symptom severity for a given day, supporting the idea of a consistent state of impaired higher cortical functions throughout an episode of delirium.

More recent research from the CIRG examined the discriminating properties for patients with delirium versus those with dementia and/or no neurocognitive disorder on four objective tests of attention: digit span, vigilance “A” test, serial 7s subtraction and months of the year backwards together with global subjective rating of attention (Adamis et al.,
The study focussed on 200 consecutive older patients admitted to a general hospital, 34 (17%) with delirium. Participants were assessed using the Confusion Assessment Method (CAM) (Inouye et al., 1990), Delirium Rating Scale-98 Revised (DRS-R98) (Trzepacz et al., 2001) and Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) scales. The five approaches to assessing attention had statistically significant correlations (P < 0.05). Discriminant analysis showed that clinical subjective rating of attention in conjunction with the months of the year backwards had the best discriminatory ability to identify Confusion Assessment Method-defined delirium, and to discriminate patients with delirium from those with dementia and/or normal cognition. Both of these approaches had high sensitivity, but modest specificity. The authors concluded that objective tests are useful for prediction of non-delirium, but lack specificity for a delirium diagnosis. Global attentional deficits were more indicative of delirium than deficits of specific domains of attention.
Android applications (apps) for cognitive assessment: background literature to the development of the Lighthouse app

Smartphones and tablet computers are widely used among hospital doctors and other healthcare professionals, with estimates suggesting that over 80% of healthcare professionals use a smartphone in their daily clinical activities (Weir et al., 2014). Smartphone apps offer a number of potential advantages over traditional pen and paper and bedside cognitive tests. These include mobility and flexibility in multiple clinical situations, consistency of test presentation, reproducibility of results, user-friendliness for patients and healthcare professionals, objectivity, relative ease of training in their use, relatively low set-up and maintenance costs and the ability to record and analyze test performance and results. The lack of such testing can be seen as a potential factor in the ongoing and significant underdetection of delirium in general hospital settings, where at least 50% of delirium cases go undetected (Kishi et al., 2007, Ryan et al., 2013).

As the availability and versatility of smartphone apps has grown, a number have been specifically designed for use in the assessment of common neurocognitive disorders. An ideal delirium screening instrument would provide an objective, sensitive, reliable, reproducible and recordable result that would discriminate delirium from dementia and other neuropsychiatric disorders. Furthermore, such an instrument would have broad coverage across all clinical populations and be effective in monitoring change in attentional status or performance.

The Edinburgh Dementia App (Weir et al., 2014)

In view of the significant clinical overlap between delirium, dementia and mild cognitive impairment, a comprehensive delirium screening and diagnostic battery should include tests for dementia. One such test is the Edinburgh Dementia App. This assesses two cognitive domains that may help discriminate patients with Alzheimer’s Dementia (AD) from those with Mild Cognitive Impairment (MCI) and normal controls. The Dual-Task (DT) performance test assesses the ability to perform two independent tasks simultaneously. In AD, deficits in
DT performance are more pronounced with increasing disease severity (Baddeley, 2007, Baddeley et al., 2001). Working Memory Binding (WMB), the temporary retention in working or short-term memory of the shape and/or colour of complex objects is also sensitive to AD and may be a pre-clinical marker for familial AD (Parra et al., 2014, Parra et al., 2010).

The WMB and DT tests are both available as part of the main menu of the Edinburgh Dementia App. The app presents these tests on a 10-inch Android tablet computer running the Google Android OS version 4.0 or above. However, this is not an ‘app’ in the strict definition, as it has not been developed for a smartphone device. Therefore, there is a need to validate and develop delirium assessment smartphone apps, such as the Lighthouse test.

Based on previous studies (Parra et al., 2014, Foley et al., 2011, Parra et al., 2010, Parra et al., 2009, Kaschel et al., 2009), it has been established that the DT is an objective measure that does not show practice based improvements, is unaffected by age and is adjusted to each individual’s level of ability. The WMB test was developed to improve on sensitivity deficits of the DT test. The WMB test has been shown to clearly differentiate deficits associated with AD and has shown promise in observing selective deficits in WMB of asymptomatic carriers of the E280A preselinin 1 AD gene mutation, the cause of the most common form of dominant early onset Familial AD.

**The Edinburgh Delirium Test Box (EDTB)**

This was evaluated in a study of patients with delirium, dementia and cognitively normal controls, with the three groups ranging between 18 and 20 people (Brown et al., 2011a). The EDTB was designed to be a portable, robust, objective and user-friendly device that could be used at the bedside or in clinic settings. The test involves eight sustained attention tasks of differing complexity. All tasks require subjects to focus attention on one or more target locations for periods of up to 72 seconds at a time. Complexity was manipulated by use of one or two target locations and by adding task irrelevant distracting stimuli for some tasks. Patients with delirium had significantly lower scores on all 8 tests of sustained attention used. Furthermore, dementia patients had a generally similar performance to
controls. Receiver operating characteristic analyses demonstrated good or excellent accuracy in discriminating delirium from dementia (area under curve values 0.80-0.94) and between delirium and control patients (area under curve values of 0.89-0.99). The authors noted that patients with delirium had marked deficits in sustained visual attention, deficits that were mainly absent in patients with dementia and controls. The authors concluded that objective testing of sustained visual attention has promise in detecting delirium and in discriminating delirium from dementia. However, the sustained attention tasks in the EDTB take a considerable length of time (4-5 minutes per task and 8 tasks in all). In view of the time involved, this assessment is unlikely to be of use in everyday clinical settings.

The Delirium App (DELAPP)

The DELAPP provides objective measurement of level of arousal (LoA) and attention. It is designed for a 5 inch Android smartphone running the Google Android OS version 2.3 and above and was developed based on methodology using the Edinburgh Delirium Test Box MkII (Tieges et al., 2013a). With the smartphone held at a comfortable distance, the patient observes stimulus on the screen and responds to verbal cues from the assessor.

The DELAPP covers a number of simple cognitive tests, including a preliminary visual acuity test followed by word building and counting tasks (Tieges et al., 2015b). The counting test involves tests of alertness, sustained and divided attention. The full test was reported as typically taking less than 5 minutes to complete. The DELAPP counting task demonstrated comparably with the EDTB2 in an initial feasibility study of 20 hospital inpatients.

In a further study in 156 general ward patients (Tieges et al., 2015b), those with delirium scored significantly lower than others and based on the Receiver Operating Characteristic (ROC) measure, the DELAPP achieved test accuracy of 0.99 in distinguishing delirium from controls and 0.93 in distinguishing delirium from dementia.

A notable deficiency in the study was the lack of patients with comorbid dementia and delirium and this clinically important group should be included in future similar research. The authors concluded that the DELAPP test showed good performance, thus supporting the
further development of mobile computerised tests of attention in the assessment and diagnosis of delirium. A potential weakness of the DELAPP is in those with visual impairment where a simpler visual test may be more acceptable to patients and thus more effective as a screening tool. Furthermore, use of the DELAPP is likely to be problematic in those with more severe cognitive impairment. Finally, input from the tester is considerable, which adds to rater ‘noise’.

**Cognitive Drug Research computerised assessment system**

For this assessment, information is presented on a high resolution computer screen with responses recorded via a module containing two buttons, marked ‘NO’ and ‘YES’. The test battery takes approximately 5 minutes to perform. In a prospective study of 100 elective orthopaedic patients over 70 years old without dementia, participants were assessed using the Digit Vigilance (DV) and Choice Reaction Time (CRT) tasks (Lowery et al., 2008). The DV task is designed to measure ability to sustain attention through accuracy of response. The CRT measures mean reaction time and intra-trial variability of reaction time. The latter has been used as a measure of fluctuating cognition in dementia. The test battery took approximately 5 minutes to perform. Over the first post-operative week after people with delirium scored lower on the MMSE and performed less accurately, slower and with greater variability of reaction time than people without delirium.
The Lighthouse test: background to development

There are numerous definitions of attention in the psychological literature. However, intact attention is generally taken as the ability to focus on a selected stimulus, sustain that focus and shift it at will, i.e. the ability to focus, sustain and shift attention (Cohen, 1993). These three components of attention are reflected in the definitions of inattention in both DSM 5 (APA, 2013) and ICD-10 (WHO, 1992), where the central importance of inattention in diagnosis of delirium is highlighted.

One neuropsychological model of attention involves four component processes: working memory; competitive selection; top-down sensitivity control and filtering for stimuli that are likely to be behaviourally important (salience filters) (Knudsen, 2007). According to this model, information about the environment is transduced and is processed by salience filters that respond differentially to infrequent or important stimuli (bottom-up processing or stimulus driven access). Neural representations in various hierarchies then encode information about the environment and a competitive process selects the representation with the highest signal strength for entry into working memory. Working memory then directs top-down bias signals that modulate the sensitivity of representations that are being processed in working memory. Working memory and competitive selection direct eye movements and other orienting behaviours and associated discharges associated with gaze control modulate sensitivity control.

There is also overlap between attention and other areas of cognition, such as working memory and executive control and definitions for working memory and executive control also vary. This has implications for testing attention, with some attentional tasks (e.g. digit span) being used as a test of attention but also involving aspects of working memory and executive functioning (Iverson, 2001). Therefore, deficits on tests of attention may be related to deficits in other areas of cognition. Furthermore, conditions such as dementia may in themselves involve attentional deficits (especially e.g. Lewy body variants) and this can lead to challenges in the clinical distinction of delirium from dementia and other cognitive problems.
Subjective assessments of attention are commonly used in delirium assessment scales, such as the Delirium Rating Scale-Revised-98 (DRS-R98; (Trzepacz et al., 2001) and Confusion Assessment Method (Inouye et al., 1990). However, subjective assessments of attention are problematic due to problems with inter-rater agreement and the need for high levels of training and expertise.

Bearing these factors in mind, the Lighthouse test was developed by the Cognitive Impairment Research Group (CIRG) in collaboration with the Department of Computer Science and Information Systems (CSIS) at the University of Limerick as an objective assessment of focussed and sustained attention, along with a measure of comprehension and awareness. A key driver for development of the Lighthouse app was the need for a computerised, brief, objective and user friendly test. Computerised tests already developed, such as the Edinburgh DeliriumTest Box and the DelApp (Tiegys et al, 2015) tend to take several minutes to complete and involve multiple stages. The Lighthouse was designed to function as a very brief delirium screening test, involving tests of identification, focussed and sustained attention. It is administered using an Android smartphone with details summarised in Table 6. Although not as yet developed for iPhone platforms, this may be considered for future versions of the Lighthouse app, in order to maximise rates of uptake and use by healthcare professionals.
<table>
<thead>
<tr>
<th>Table 6: Lighthouse test</th>
</tr>
</thead>
</table>

**Xperia SP model: c5303 specifications:**

- Operating system: Google™ Android™ 4.1.2 (Jelly Bean)
- OS Build number 12.0.A.2.254
- Processor 1.7 GHz Qualcomm MSM8960Pro Dual Core
- GPU Adreno 320
- Size 130.6 x 67.1 x 9.98 mm
- Weight 155 grams

**Screen:**

- Colours 16,777,216 colour TFT
- Resolution 1280 x 720 pixels
- Size (diagonal) 4.6 inches
- Scratch-resistant Corning™ Gorilla™ glass

**Input mechanisms:**

- Text input on-screen QWERTY keyboard
- Touch screen Capacitive
- Touch gesture: multi-touch, up to 10 fingers supported

**Memory:**

- RAM 1 GB
- IDE and Language for App developed
- Java on Android Studio 1.0.1
The Lighthouse test: operation and pilot study

The Lighthouse test involves presentation of an image of a flashing lighthouse on a standard smartphone screen (3” x 5”).

Screenshots from the Lighthouse app test, as used in this study, are included as Figures 1 and 2 below. Screenshots of modified versions of the Lighthouse app can be seen in Figures 3 and 4, for use in future studies. The utility of the modified version of the Lighthouse app in Figures 3 and 4 will be assessed in future research projects. Modifications include a clearer and more accurate image of the actual Lighthouse and a clearer flash. These modifications were informed by the results of this project, whereby a qualitative analysis of 100 participants revealed that 35 correctly identified the object in the app as a Lighthouse, 35 identified it as something else and 30 had no response.

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 (x 3 sequences; 4, 3, 5). Subjects are requested to identify the number of flashes.

(3) Testing the capacity to sustain attention to count sequences of flashes (x3) (i.e. 4-3-2, 3-2-5, 2-4-3). Again, subjects are requested to identify the number of flashes.

Moreover, it is also expected that the Lighthouse test would be brief; easy to administer; acceptable to older persons; minimally affected by education, sex, age and other factors unrelated to dementia; and have high sensitivity and specificity for major neurocognitive disorders.
Figure 1: Lighthouse app screenshot (beginning of flash emerging to right of image)
Figure 2: Lighthouse app screenshot (flash further developed to left of image)
Figure 3: Lighthouse app (modified version: for testing in future research; larger flash image may improve clinical utility)
Figure 4: Lighthouse app (modified version)
In a study of 51 medical outpatients over 65 years of age (Rice et al., 2015), an earlier version of the Lighthouse test was applied along with other standard tests of cognition: months of the year backwards, WORLD backwards, serial sevens, digit span forwards and digit span backwards. 64.7% of participants expressed a preference for the Lighthouse test over standard tests, with 15.6% expressing a preference for the standard tests and the remainder being ambivalent. There was limited relationship between performance on the two parts of the Lighthouse test (i.e. focused and sustained attention tasks). A significant correlation was found between those scoring less than 3 on the sustained attention task of the Lighthouse test and those scoring 23 or less on the MMSE ($X^2 = 8.1$, $p = 0.004$, Kappa = 0.004).

An earlier version of the Lighthouse test was also piloted in a general medical inpatient population in University Hospital Limerick (O'Regan, 2015). A random sample of 50 inpatients over age 65 years were asked to complete the Lighthouse test along with other brief tests of attention and general cognition: months of the year backwards test (MBT) (Meagher et al., 2015); serial sevens; WORLD backwards (Folstein et al., 1975); Digit Span Forward (DSF); Digit Span Backwards (DSB) and MMSE (Folstein et al., 1975). The Lighthouse test was presented on a Samsung Nexus-10 tablet. Cognitive status was determined by MMSE score and participants were categorized as being cognitively impaired (MMSE ≤ 23) ($N = 15$, 30%) or cognitively normal (MMSE ≥ 24) ($N = 35$, 70%). Overall, 32 patients (64%) were able to identify the image as a lighthouse. This included 26 (74%) of the cognitively normal group, in comparison to only 6 (40%) of those in the cognitively impaired group ($p = 0.03$).

In terms of performance, 11 participants (73.3%) of the cognitively impaired group were able to complete all three elements of the focused attention (FA) task and 2 participants (13%) completed the sustained attention (SA) task. The corresponding figures for the cognitively intact group were 77% and 43% respectively. The two main components (FA and SA tests) of the Lighthouse test had modest concordance (kappa = 0.42; $p < 0.001$). While the Lighthouse test had only modest concordance with other tests of attention, it proved to be the most preferred test by patients, when asked directly to comment on this issue.
The version of the Lighthouse app used for this study (Figures 1 and 2) has been modified further, with a clearer lighthouse image and flash (Figures 3 and 4). This modification is likely to help with identification and engagement with the test, hence improving overall acceptability, performance and utility of the test.

**Delirium: management and prevention**

**Non-pharmacological**

Non-pharmacological management and prevention of delirium in the general hospital involves the application of principles of good nursing and medical care for individual patients: Table 4 (O’Connell et al., 2014). There is also a clear need for the development of cognitive friendly delirium-oriented hospital-wide programmes that would address among all staff members (clinical and support staff) aspects of general awareness, education, training, use of evidence-based treatment and preventative approaches and the use of effective discharge planning (O’Connell et al., 2014). In a recent narrative review, we concluded that the success of complex delirium programmes in the general hospital setting is linked to key systems factors that include:

- Involvement of clinical leaders
- Support from senior management
- Linking the implementation of programmes to periods of systems change (e.g. realignment of care pathways)
- Educational elements that are sustained and engaging
- Mechanisms to support decision-making that are integrated into everyday routines (e.g. electronic care pathways)
- Monitoring procedures to promote continued adherence
| Table 7: Multicomponent non-pharmacological management of symptoms of delirium  
(O'Connell et al., 2014) |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Educate patient and family/carer on delirium and prognosis</td>
</tr>
<tr>
<td>• Involve family/carer in hospital care routine</td>
</tr>
<tr>
<td>• Repeatedly reorient and reassure the patient</td>
</tr>
<tr>
<td>• Normalise sleep patterns</td>
</tr>
<tr>
<td>• Prevent complications, e.g. falls, constipation</td>
</tr>
<tr>
<td>• Ensure adequate hydration</td>
</tr>
<tr>
<td>• Ensure adequate pain relief</td>
</tr>
<tr>
<td>• Encourage general activity, mobility and activities of daily living</td>
</tr>
<tr>
<td>• Use visual/hearing aids to facilitate communication</td>
</tr>
<tr>
<td>• Nurse with familiar staff in relaxed environment</td>
</tr>
</tbody>
</table>
Pharmacological management of delirium

Low dose atypical and typical antipsychotics are commonly used to treat delirium - primarily the hyperactive presentation of delirium, especially when there are behavioural problems, psychotic symptoms or associated distress. Open-label studies and long-standing clinical experience support the use of antipsychotic agents, but there is limited evidence available from randomised controlled trials. However, a large study of over 2,000 general hospital patients concluded that supervised and controlled use is safe (Hatta et al., 2014). The NICE guidelines support the targeted use of haloperidol or olanzapine (NICE, 2010) and there is also evidence from placebo-controlled studies to support the use of quetiapine (Devlin et al., 2010, Tahir et al., 2010). Despite the theoretical potential for procholinergic agents in the treatment of delirium, the evidence base for their use in clinical studies is negative (van Eijk et al., 2010).

Evidence supports the prophylactic use of low dose haloperidol, olanzapine and risperidone in high risk populations (e.g. older people undergoing orthopaedic surgery) but routine use is not yet justified (Larsen et al., 2010, Wang et al., 2012).

A Cochrane review found insufficient evidence to support the use of benzodiazepines in delirium which, along with their deliriogenic potential, suggests that they should be used with caution and limited to delirium in specific circumstances such as withdrawal states or seizure-related delirium (Lonergan et al., 2009).
Study hypotheses

1. The clinical utility of the novel Lighthouse tests (Lighthouse identification, focussed attention and sustained attention) in terms of accuracy for distinguishing delirium from non-delirium in patients with mixed neuropsychiatric presentations will compare favourably with commonly used ‘gold-standard’ cognitive screening tests.

2. The coverage, i.e. the proportion of patients with various neurocognitive disorders who are able to engage meaningfully (i.e. are rateable) with the Lighthouse test will be at least equivalent to that evident for standard tests.
Chapter 2: Systematic review of bedside tests of attention

Delirium is a common neuropsychiatric syndrome with cerebral decompensation caused by an underlying physical stressor and characterised by disturbances of attention, arousal and behaviour (APA, 2013). Delirium is independently associated with increased risk of functional loss and morbidity but is commonly undetected or missed in the hospital setting (Ryan et al., 2013). As such, early detection and treatment of delirium are crucial to minimising these adverse consequences of delirium.

Inattention is a mandatory component for DSM-5 diagnosis of delirium defined by ‘the reduced ability to direct, focus, sustain and shift attention’ (Tables 1 and 2) (APA, 2013). However, inattention can also be impaired in a variety of disorders other than delirium (O'Regan et al., 2014b) with distinction from dementia and depression especially challenging due to the overlap in clinical features. While the onset of delirium is considered more acute than that of dementia, dementias such as vascular type and Lewy Body type can also present acutely and with fluctuation of symptoms. Furthermore, dementia may be present in up to two thirds of delirium occurring in elderly populations (Meagher et al., 2010).

In order to improve management and outcomes for patients with delirium there is a need for a reliable, objective and convenient test that can be applied to the efficient assessment of delirium-relevant cognitive impairment. Bettin et al (1998) describe how a suitable tool should be repeatable, easily administered and reflect the severity of delirium. Moreover, the prognostic gravity of delirium is such that test sensitivity must be emphasised over specificity in order to minimise the risk of missed cases (O'Regan et al., 2014b).

At present, many different methods are used to determine the presence of delirium (Meagher and Leonard, 2008). These include a variety of subjective and objective tests that are administered by a range of health care providers including physicians and nurses. However, there is a lack of consensus as to which tool is most useful in the detection and assessment of delirium.

A recent narrative review of objective tests of attention in delirium (Tieges et al., 2014) identified fifteen publications. The attention tests used in the studies were grouped into measures of attention span, vigilance tests, other pen-and-paper tests and computerised
tests of speeded reaction, vigilance and sustained attention. The authors concluded that the existing evidence base on objective attention in delirium is small, that objective testing of attention is underdeveloped but that it shows considerable promise in clinical practice and research.

By now approaching the literature systematically, we explore the use of objective bedside tests of attention in more detail. Furthermore, applying an objective measure of quality for included papers, the Downs and Black Tool (Downs and Black, 1998) will help develop further and more detailed recommendations for future research.
Methods

The search aimed to identify all studies that have examined the use of objective bedside tests of attention for delirium assessment. We performed a comprehensive systematic search of the literature including: Medline Bibliographic Database on Ebsco platform (coverage 1946-present); Psychology and Behavioural Science Collection Database on Ebsco platform (coverage 1952-present); PsycINFO, American Psychological Association database on Ovid platform (coverage 1806 – present); PubMed database (coverage 1946 – present); Cochrane Library of Systematic Reviews and NICE guidelines. A structured search strategy was devised and used which included controlled vocabulary and relevant key terms. Two limiters were used: search results were limited by language (English Language), and year, (1980 to October 2015).

The search was conducted by Librarian staff at Midland Regional Hospital Tullamore as detailed below with the search strategy summaries attached as appendices. The studies included described the use of bedside tests, i.e. brief and mobile tests (pen and paper or computerised) to assess for attentional impairment in the assessment and diagnosis of delirium: titles and abstracts were reviewed and studies were considered for inclusion based on these criteria. One reviewer (HO’C) then reviewed the full-text versions of the papers considered for inclusion and discussed all papers considered with another reviewer (DM). Manual searching of reference lists for included full-text articles was also conducted.

Studies of general hospital in-patients were included (medical, surgical and psychiatric populations) with studies in outpatients and community dwelling populations excluded. Tests that require extended testing time (i.e. more than 10 minutes), that are not mobile or that cannot be conducted at the bedside in a general hospital setting were excluded.

The Downs and Black Tool (Downs and Black, 1998) (see Appendix) was used to assess the quality of included studies. This tool consists of a 27 point study checklist covering the areas of reporting, external validity, bias, confounding and power. In a feasibility study of this tool, the Quality Index has been shown to have high internal consistency (KR-20: 0.89) and Test-retest (r 0.88) and inter-rater (r 0.75) reliability of the Quality Index were also good. Due to the heterogenous nature of the study populations and the variety of assessment tools used, pooling of the data was not considered to be feasible.
Results

The search strategy yielded 272 papers. Abstracts of these 272 papers were reviewed by one of the authors (HOC) and 23 articles which met the inclusion criteria were identified and are summarised in Table 8.

Sample sizes of groups with delirium ranged from 8 (Christensen et al., 1996) to 100 (Meagher et al., 2007) participants. A total of 2,529 subjects were included in all studies, including 831 with delirium and 169 explicitly categorised with comorbid delirium and dementia.

The Digit Span Test was the most frequently studied test of attention (10 studies) and was consistently accurate in the detection of delirium. Spatial span testing was examined in four studies, trail making tests in three, vigilance testing in three and months backwards testing in two (see Table 3 for list of studies including the different tests).
### Table 8: Studies categorised according to clinical setting

<table>
<thead>
<tr>
<th>Clinical setting</th>
<th>Number of studies</th>
<th>Studies</th>
<th>Total N</th>
<th>N with delirium</th>
<th>N with comorbid dementia and delirium</th>
</tr>
</thead>
<tbody>
<tr>
<td>General hospital</td>
<td>8</td>
<td>Pompei et al., 1995, Christensen et al., 1996, O'Keeffe and Gosney, 1997a, Bettin et al., 1998, Brown et al., 2011a, O'Regan et al., 2014a, Tieges et al., 2015b, Adamis et al., 2015(Fick et al., 2015)</td>
<td>1,446</td>
<td>295</td>
<td>4</td>
</tr>
<tr>
<td>Palliative care</td>
<td>5</td>
<td>Bosisio et al., 2006, Meagher et al., 2007, Leonard et al, 2009, Meagher et al., 2010, Meagher et al., 2012</td>
<td>546</td>
<td>291</td>
<td>84</td>
</tr>
<tr>
<td>Intensive care</td>
<td>3</td>
<td>McNicoll et al., 2005, Hart et al., 1997, Hart et al., 1996</td>
<td>199</td>
<td>49</td>
<td>7</td>
</tr>
<tr>
<td>Psychiatry Consultation-Liaison</td>
<td>3</td>
<td>Trzepacz et al., 1988a, Jabbar et al., 2011, Rajlakshmi et al., 2013</td>
<td>202</td>
<td>144</td>
<td>40</td>
</tr>
<tr>
<td>Stem cell transplantation</td>
<td>1</td>
<td>Beglinger et al., 2011</td>
<td>59</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Cardiac surgery</td>
<td>1</td>
<td>Brown et al., 2011b</td>
<td>37</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Liver transplantation</td>
<td>1</td>
<td>Trzepacz et al., 1986)</td>
<td>40</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Orthopaedic surgery</td>
<td>1</td>
<td>Lowery et al., 2008)</td>
<td>100</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>
8 studies (Pompei et al., 1995, Christensen et al., 1996, O’Keeffe and Gosney, 1997a, Bettin et al., 1998, Brown et al., 2011a, O'Regan et al., 2014b, Tiegess et al., 2015b, Adamis et al., 2015, Fick et al., 2015) were based in general hospital settings and included 1,446 patients, including 295 with delirium and 4 explicitly described as having comorbid dementia and delirium.

4 studies (Bosisio et al., 2006, Meagher et al., 2007, Meagher et al., 2010, Meagher et al., 2012) were based in palliative care settings and included a total of 446 patients, including 262 with delirium and 84 explicitly described as having comorbid dementia and delirium.

3 studies (McNicoll et al., 2005, Hart et al., 1997, Hart et al., 1996) were based in intensive care settings and included 199 patients in total, including 49 with delirium and 7 explicitly described as having comorbid dementia and delirium.

3 studies (Trzepacz et al., 1988a, Jabbar et al., 2011, Rajlakshmi et al., 2013) involved referrals to Psychiatry Consultation-Liaison services and included a total of 202 patients, including 144 with delirium and 40 explicitly described as having comorbid dementia and delirium.

One study (Beglinger et al., 2011) focussed on stem cell transplantation patients and included 59 participants, 19 of whom developed delirium.

One study (Brown et al., 2011b) was conducted in a cardiac surgery setting and included 37 participants, 9 with delirium.

One study (Trzepacz et al., 1986) was in a liver transplantation setting and included 40 participants, 12 with delirium.

Finally, one study (Lowery et al., 2008) was based in an orthopaedic surgery setting and included 100 participants, 14 with delirium.

Age ranges varied widely between different studies and within different studies, depending on diagnostic group. Age, gender and diagnostic groupings for all studies are summarised in Table 4. All studies except one (Trzepacz et al., 1986) included at least some patients over age 65.
All studies apart from that by Rajlakshmi et al (2013) included non-delirium control groups or a comparison group with comorbid dementia and delirium. Control groups included: dementia (6 studies), schizophrenia (2 studies) and depression (1 study).

Four studies included patients with comorbid dementia and delirium (O’Keeffe and Gosney, 1997a, McNicoll et al., 2005, Meagher et al., 2010, Meagher et al., 2007).

Thirteen different bedside attentional tests were used in these studies (see Table 9).
### Table 9: Summaries of cognitive tests used in included publications (tests listed in order of frequency studied)

<table>
<thead>
<tr>
<th>Test</th>
<th>Neuropsychological domains</th>
<th>Description</th>
<th>Time</th>
<th>Studies involving this test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit span (Weschler, 1997)</td>
<td>attention span; working memory; arithmetic abilities; executive functioning.</td>
<td>Examiner presents a series of random numbers at a rate of 1 per second. Subject asked to repeat the sequence which they have been presented with. Firstly, a 2 number sequence is presented followed by a 3 number sequence. Subject is asked to repeat the sequence in either the same or reverse order. Test scores based on maximum span length reproduced.</td>
<td>5-10 minutes</td>
<td>Total studies = 10 (Pompei et al., 1995, Christensen et al., 1996, O'Keeffe and Gosney, 1997a, Bettin et al., 1998, McNicoll et al., 2005, Bosiljo et al., 2006, Brown et al., 2011c, Adamis et al., 2015, Leonard et al., 2009, Fick et al., 2015)</td>
</tr>
<tr>
<td>Spatial span (Weschler, 1997)</td>
<td>attention span; visuospatial working memory</td>
<td>Sequence of numbered squares tapped out by examiner. Subject asked to repeat the sequence either in same order (forward span) or in reverse order (backward span). Maximum length of sequence which can be remembered is used to score the participant. The spatial span forwards and backwards tests are used to assess attention in the Cognitive Test for Delirium (CTD) (Hart et al., 1996)</td>
<td>5-10 minutes</td>
<td>Total studies = 4 (Meagher et al., 2007, Meagher et al., 2010, O'Regan et al., 2014a, Hart et al., 1997)</td>
</tr>
<tr>
<td>Trail Making Test (Reitan and Wolfson, 1985)</td>
<td>visual search/scanning; selective, divided and sustained attention; psychomotor speed; set shifting.</td>
<td>Series numbers and letters enclosed by circles on a sheet of paper. Part A: subject required to draw a line to connect the numbers 1 – 25 in ascending order. Part B: subject required to draw a line in order to connect both numbers (1 -13) and letters (A-L), i.e 1-A-2-B-3-C, etc. During Part A and Part B, subject must not lift pencil from paper and they are instructed to connect the sequence as quickly as possible.</td>
<td>5-10 minutes</td>
<td>Total studies = 3 (Trzepacz et al., 1986, Trzepacz et al., 1988b, Beglinger et al., 2011)</td>
</tr>
<tr>
<td>Vigilance 'A' Test (Strub and Black, 2000)</td>
<td>sustained attention.</td>
<td>Subject is read a list of 60 letters and instructed to tap once on the table each time the letter ‘A’ is read by the examiner.</td>
<td>3 minutes or less</td>
<td>Total studies = 3 (Pompei et al., 1995, McNicoll et al., 2005, Adamis et al., 2015)</td>
</tr>
<tr>
<td>Months of the year backwards (Katzman et al., 1983)</td>
<td>focussed and sustained attention; processing speed</td>
<td>Subject asked to recite months of the year in correct order and then in reverse order from December to January. Subjects considered to have passed the test on reaching July without error. This test is used to score the attention item in the DRS-R98 (Trzepacz et al., 2001).</td>
<td>3 minutes or less</td>
<td>Total studies = 3 (Adamis et al., 2015, O'Regan et al., 2014a, Fick et al., 2015)</td>
</tr>
<tr>
<td>Stroop task (Spreen and Strauss, 1998)</td>
<td>selective attention; inhibition of task-irrelevant information; executive function.</td>
<td>Stimuli consist of printed words in different ink colours. Subjects asked to name the colour of the ink in which the word is printed. The printed words may be congruent or incongruent with the ink colour (e.g. the word ‘red’ printed in red or green ink). The ‘Stroop effect’ reflects the degree of difficulty in inhibiting a learned response.</td>
<td>At least 10 minutes</td>
<td>Total studies = 1 (Brown et al., 2011b)</td>
</tr>
<tr>
<td>Edinburgh Delirium Test Box (Brown et al., 2011a)</td>
<td>sustained attention, arithmetic abilities.</td>
<td>Portable electronic device is composed of large and clearly visible illuminable response buttons. Visual target stimuli are presented by the illumination of a series of flashing lights. Subject must either press the buttons as they illuminate or count the number of illuminations. Subject is then scored on the accuracy of their responses to the stimuli. The DelApp (Tiegas et al., 2015) has been developed as a smartphone application based on the EDTB.</td>
<td>At least 10 minutes</td>
<td>Total studies = 1 (Brown et al., 2011b)</td>
</tr>
<tr>
<td>Digit Vigilance Task (Strub and Black, 2000)</td>
<td>sustained attention.</td>
<td>In computerised version, participant required to use the module to identify multiple presentations of ‘target’ digit with ‘distracter’ digits.</td>
<td>At least 5 minutes</td>
<td>Total studies = 1 (Lowery et al., 2008)</td>
</tr>
<tr>
<td>Digit Cancellation Test (Della Sala et al., 1992)</td>
<td>sustained attention, visual search.</td>
<td>Subject is asked to identify target digits among an array of symbols within a specified amount of time.</td>
<td>5 minutes</td>
<td>Total studies = 1</td>
</tr>
<tr>
<td>Test Name</td>
<td>Cognitive Domain</td>
<td>Instructions</td>
<td>Duration</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Choice Reaction Time (Simpson et al., 1991)</td>
<td>Sustained attention, psychomotor speed</td>
<td>The words YES and NO are presented on a computer screen and subjects are asked to press either a YES or No button accordingly. Participant is then scored on the accuracy of responses, late and early responses and response speed.</td>
<td>At least 5 minutes</td>
<td>(O’Keeffe and Gosney, 1997a)</td>
</tr>
<tr>
<td>DelApp (Tieges et al., 2015b)</td>
<td>Derived from Edinburgh Delirium Test Box: sustained attention, arithmetic abilities.</td>
<td>Visual acuity and visual sustained attention counting tasks presented on smartphone app, and level of arousal assessment</td>
<td>At least 5 minutes</td>
<td>(Lowery et al., 2008)</td>
</tr>
<tr>
<td>Serial sevens test (Folstein et al., 1975)</td>
<td>Focussed and sustained attention; processing speed</td>
<td>Patient asked to count back from 100 in sevens</td>
<td>3 minutes</td>
<td>(Tieges et al., 2015b)</td>
</tr>
<tr>
<td>Days of the week backwards</td>
<td>Focussed and sustained attention; processing speed</td>
<td>Subject asked to recite days of the week backwards</td>
<td>3 minutes or less</td>
<td>(Adamis et al., 2015)</td>
</tr>
</tbody>
</table>

Total studies = 1
The Downs and Black tool was applied to the 23 studies. Scores ranged from 11 to 15 with higher scores reflecting higher quality papers: see Tables 11 and 12.

There was a notable lack of studies explicitly focussing on patients with comorbid dementia and delirium. Only three studies included computerised tests of attention.

Areas within the Downs and Black Tool where studies performed consistently poorly included sample selection (items 11 and 12), blinding of patients and assessors (items 14 and 15) and power analysis (item 27). Other areas, such as interventions (item 4), adverse events (item 8) and compliance with interventions (item 19) are most applicable to clinical trials and intervention studies and less so for this selection of predominantly descriptive and phenomenological studies.
Trzepacz et al (1986) explored the use of Trail Making Parts A and B tests (along with EEG, albumin levels and MMSE) in the detection of delirium in adult candidates for liver transplantation. The study included forty consecutive patients (12 with delirium) admitted for liver transplantation. All patients were evaluated twice. Trail Making Part A and B test results were significantly associated with the presence of delirium (p=0.001 for test A, p=0.002 for test B). Trail Making Test Part A was found to be the most specific (82.1%), while Trail Making Test Part B was most sensitive (91.7%) for the presence of delirium. Other neurocognitive disorders were not addressed in this study, with only one patient categorised as having dementia.

Trzepacz et al (1988a) examined the use of Trail Making Part A and Trail Making Part B tests in the detection of delirium among medical and surgical referrals to a psychiatry consultation-liaison service. Twenty patients had delirium, while there were three comparison groups: one group comprised of 9 hospitalised patients with schizophrenia; another group comprised of 9 hospitalised patients with dementia and another group comprised of 9 medical inpatients referred for psychiatric consultation with psychiatric disorders that did not include psychotic symptoms or cognitive impairment. Assessments were conducted independently by two psychiatrists who were unaware of the subject’s diagnosis or ratings of the other psychiatrist. Both the delirium and schizophrenia groups showed severe impairment on Trail Making Parts A and B. Nine participants with delirium were unable to complete Trail Making Part A and B but the reasons for these failures of participation are not given. This can be viewed as a deficit in clinical coverage or penetration for these particular tests. Delirium Rating Scale scores correlated highly with Trail Making Test Part B scores (r= 0.66).

Pompei et al (1995) examined the Digit Span Test and the Vigilance ‘A’ Test in a prospective cohort study of medical and surgical inpatients aged > 65 years, including individuals with delirium (n=61) and control subjects (n=367). Participants were assessed daily during the first week of hospitalisation and then every other day until they were discharged. Sensitivity/specificity for the Digit Span Test was reported as 0.34/0.90. Sensitivity/specificity for the Vigilance A Test was reported as 0.61/0.77. The authors
concluded that both the Digit Span Test and Vigilance ‘A’ Test were more useful for confirming delirium than excluding it.

Christensen et al (1996) examined the Digit Span Forwards (DSF) and Digit Span Backwards (DSB) in monitoring delirium severity in medical and surgical inpatients (10 assessed during a delirium episode and later during recovery). Bedside testing using the Digit Span Forward and Digit Span Backward was conducted at three designated times per day for seven successive interviews. Further testing was carried out at one and two weeks after initial enrolment. They compared scores from early interviews (when participants had delirium) with scores from the later interviews when participants were recovering. Digit Span Forward and Digit Span Backwards scores significantly improved in the recovery period, with mean forward span scores improving from 9.7 to 15.1 and mean backward span scores improving from 0.89 to 4.2.

Hart et al (1997) examined the use of the Visual Attention Span Forward (VASF) and Visual Attention Span Backwards (VASB) tests in the detection of delirium (19 patients with delirium; 26 with dementia; 30 with depression; 25 with schizophrenia). Both of these tests form components of the larger Cognitive Test for Delirium (CTD). Participants included inpatients with delirium, schizophrenia and depressive illness and outpatients with dementia. The VASF was a reliable method of distinguishing those with delirium from all groups, including those with severe dementia (p<0.05). The VASB did not accurately distinguish delirium from dementia. While participants with schizophrenia and depressive illness were included in this study, the specific results of the Visual Attention Span Tests for these groups are not described.

O’Keeffe and Gosney (1997a) examined the use of the Digit Span Forwards (DSF), Digit Span Backwards (DSB), Digit Cancellation Test (DCT) and Vigilance A test in the diagnosis of delirium amongst patients admitted to an acute geriatric unit (18 with delirium; 18 with dementia; 52 controls). Participants were assessed independently by two geriatricians within 90 minutes of each other. Global attentiveness was rated by one physician while a second physician assessed patients to determine whether they met the DSM-III-R criteria for delirium. Following the determination of a diagnosis of delirium, the second physician conducted DSF, DSB, a timed DCT and Vigilance A test. The DCT used was a modified version
which included DCT1 and DCT2; for DCT1 participants were instructed to cross out one target digit, while for DCT2 there were two target digits to be crossed out. All tests of attention were found to correlate significantly with global ratings of attentiveness except the DSF test. Both the DSB and DCT distinguished delirium from dementia. Participants with delirium were significantly more impaired than those with dementia despite similar MMSE scores.

Bettin et al (1998) explored the utility of the Digit Span Forward Test (DSF) in the assessment of delirium severity (22 with delirium and 15 controls). All participants were recruited from acute care wards and control subjects were then matched to subjects with delirium on the basis of age and severity of medical illness. Cognitive assessments were conducted three times daily for seven consecutive assessments and once daily on the 7th and 14th days post-enrolment. Participants with delirium scored significantly lower on the DSF test in comparison to the control group on early testing. At late assessment, the mean scores for DSF were not significantly different between the two groups. Delirium Severity Scale scores at late assessment were also improved. The authors suggest that the enhanced performance on the DSF test can be accounted for by an improvement in delirium status.

McNicoll et al (2005) explored the use of the Vigilance A Task in the diagnosis of delirium in ICU patients (8 with delirium; 3 with dementia; 7 with comorbid dementia and delirium and 4 controls). Patients were recruited and assessed by trained researchers who were blinded to each other’s findings. Using the Confusion Assessment Method (CAM) as the reference standard, the intensive care unit version (CAM-ICU) had a sensitivity of 73% and specificity of 100%. The false negative results with the CAM-ICU were attributed to the more detailed cognitive testing used in the CAM. Regarding inattention, the CAM tasks (MMSE recall and WORLD backwards; Digit Span) and the CAM-ICU tasks (Vigilance A; picture recognition) were combined and had sensitivity of 76% and specificity of 100%.

Bosisio et al (2006) assessed 106 cancer patients (66 with delirium and 40 without) recruited from six centres including four medical oncology units and two palliative care units. The Delirium Rating Scale (DRS) (Trzepacz et al., 1988b) and Memorial Delirium Assessment Scale (MDAS) (Breitbart et al., 1997) were used to assess patients. 100% of the participants with delirium demonstrated impairment on the attention item of the MDAS, with at least
mild (25%) or moderate-to-severe impairment (75%) and this reached statistical significance in comparison to non-delirium participants (p < 0.001).

Meagher et al (2007) conducted a study using the Spatial Span and Vigilance task components of the Cognitive Test for Delirium (CTD). This was a prospective cross-sectional study of delirium symptoms and cognitive performance in consecutive cases of DSM-IV delirium referred from a palliative care in-patient service. 100 participants with delirium were included in the study (83 with delirium and 17 with comorbid dementia and delirium). The attention components of the Delirium Rating Scale-Revised-98 (assessed using the months backwards test) (Trzepacz et al., 2001) and CTD (assessed using spatial span forwards and backwards) (Hart et al., 1996) correlated highly (r= -0.73).

Lowery et al (2008) hypothesised that measuring attention with computerised versions of the Digit Vigilance Task Vigilance and Choice Reaction Time tasks would provide a sensitive marker of delirium during the post-operative period. These tests measure sustained attention and processing speed. Participants who were undergoing orthopaedic surgery (hip and knee replacements) were recruited (14 patients with delirium and 80 controls). Patients identified as meeting criteria for dementia were excluded from the study. Patients with delirium showed a decrease in accuracy on the vigilance task post-operatively. In contrast, those without delirium had no change in accuracy on the choice reaction task but did demonstrate a slowing of reaction time.

Leonard et al (2009) assessed 100 consecutive palliative care admissions immediately after admission and one week later. Overall, 51% experienced either major depression or delirium. Most patients with syndromal delirium met criteria for major depressive illness and 50% of those with depression had delirium or subsyndromal delirium. At initial assessment the CAM attention item was positive in 100% of those with delirium, in comparison to 23% with depression without syndromal depression and 30% in those with neither depression nor delirium. Comparisons of attentional items in the Memorial Delirium Assessment Scale for these three clinical groups showed significant impairment for digit span and ability to maintain and shift attention in the syndromal delirium group in comparison to the other two groups (impaired digit span scores of 2.45; 0.79 and 0.73.
respectively and reduced ability to maintain and shift attention of 2.34; 0.32 and 0.33 respectively).

Meagher et al (2010) explored the use of the Spatial Span Forward (SSF) and Spatial Span Backward (SSB) in the diagnosis of delirium. This was a prospective cross-sectional study of delirium symptoms and cognitive performance in a population with mixed neurocognitive disorders (40 with delirium; 20 with dementia; 40 with comorbid dementia and delirium and 40 controls) receiving care in a palliative care inpatient service. Spatial Span Forward (SSF) scores were significantly worse for delirium versus dementia (p=0.02) and for comorbid delirium-dementia versus dementia (p=0.05). The mean scores for the Spatial Span Backward (SSB) did not distinguish delirium from dementia.

Jabbar et al (2011) explored the phenomenology of delirium in patients referred to an old age psychiatry consultation-liaison service in comparison to palliative care patients. Consecutive cases were assessed using the DRS-R98 and CTD. 80 patients with delirium were included (40 with comorbid dementia). Inattention (100%) was the most prominent cognitive disturbance and was associated with severity of other cognitive disturbances on both the DRS-R98 and CTD but not with DRS-R98 non-cognitive items. Scores on the DRS-R98 and CTD attention items were related. Scores on the attention item of the DRS-R98 increased in parallel with the attention scores on the CTD attention scores (DRS-R98 attention item scores of 1.4 ± 0.9, 2.4 ± 0.7 and 2.7 ± 0.6 corresponding to CTD attention scores of 4-6, 2 and 0, respectively, p < 0.01).

Brown et al (2011b) assessed patients aged over 60 years before and after cardiac surgery using the Victoria version of the Stroop test and Digit Span tests (37 patients were included in the study and 9 developed delirium post-operatively). Patients with dementia were excluded. Pre-operative Stroop and Digit Span Tests were conducted on hospital wards between 0-8 days prior to surgery. These assessments were conducted again in hospital 2-9 days after surgery. Digit Span scores in patients who developed delirium were significantly lower (z= -2.53) compared to those without delirium. Furthermore, Stroop interference was significantly greater in the group who developed delirium (z = -2.02).

Brown et al (2011c) developed the Edinburgh Delirium Test Box to assess sustained visual attention. They hypothesised that deficits of sustained visual attention would be detected in
patients with Delirium using this tool. Three groups of patients over the age of 65 years were recruited for this study; i) patients with delirium and no known dementia (total number 20) who were medical and surgical inpatients of a general hospital ii) patients with Alzheimer’s Dementia who were outpatients of a memory clinic (total number 28) and iii) cognitively normal patients who were inpatients in the same wards as those with delirium (total number 20). Eight tasks measuring sustained visual attention were administered using the Edinburgh Delirium Test Box less than 1 hr after they were assessed for delirium. Patients with delirium had significantly lower scores on all sustained attention tasks than the group with dementia and the control group (p values from 0.003 to <0.001). Across the eight sustained attention tasks, there was either good or excellent discrimination between delirium and cognitively normal patients (AUC values 0.89-0.99) and delirium and dementia (AUC values 0.80-0.94).

Beglinger et al (2011) compared Trail Making Test performance in bone marrow transplant (BMT) patients (total number 52) with a healthy comparison group (total number 10). Of the 52 BMT patients, 19 developed post-transplant delirium. Trail Making Test parts A and B did not differ significantly between BMT patients with delirium and those without delirium. The delirium group performed significantly less well than comparisons in Trails A, Trails B, List Learning and List Recall. The BMT delirium group was worse than the no delirium BMT group on List Learning and List Recall.

Meagher et al (2012) assessed 100 Palliative care in-patients with delirium using the Cognitive Test for Delirium (CTD) and Delirium Rating Scale-Revised-98. Participants underwent twice weekly visits for 3 weeks followed by weekly visits thereafter. The CTD was used to compare subsyndromal with syndromal delirium. The CTD attention component (tested using visual spatial span, forwards and backwards) accurately distinguished persistent from resolving delirium. Furthermore, attention scores on the CTD distinguished full syndromal versus subsyndromal delirium suggesting a spectrum of inattention according to delirium syndromal status.

Rajlakshmi et al (2013) performed a cross sectional study (84 patients with DSM-IVTR delirium (APA, 2000)) of medical and surgical in-patients referred to a consultation-liaison psychiatry service. Patients were assessed with the DRS-R-98 (Trzepacz et al., 2001) and CTD
Poor attention on the CTD was associated with significantly higher motor retardation and higher DRS-R98 severity scores. There were no significant differences between those with and without poor attention. Higher attention deficits were associated with higher dysfunction on all other CTD domains of cognition. There was significant correlation between cognitive functions as assessed on CTD and total DRS-R-98 score, DRS-R-98 severity score and DRS-R-98 severity score without the attention item score. Few correlations emerged between CTD domains and CTD total scores with cognitive symptom total score or DRS-R-98 (items 9-13) and non-cognitive symptom total score of DRS-R-98 (items 1-8). There was no significant correlation found between DRS-R-98 and CTD attention and vigilance domains (r=-0.10 and r=-0.15 respectively).

O’Regan et al (2014b) examined the accuracy of the Months of The Year Backwards (MBT) and Spatial Span Forward (SSF) tests in the detection of delirium in a cross sectional study of hospital inpatients (48 patients with delirium; 217 patients without delirium). SSF and MBT tests were administered to individuals with delirium who had been diagnosed using DSM-IV criteria and these results were compared to control subjects who were also hospital inpatients. The most accurate single test to screen for delirium was the MBT with a sensitivity of 83.3% and specificity of 90.8%. Considering participants over the median age of 69 years, MBT was found to have a sensitivity of 83.8% and specificity of 89.6%. The performance of combined tests for the prediction of delirium was also assessed. The combined application of MBT and SSF4 demonstrated a sensitivity of 93.8% and specificity of 81.1%.

Tieges et al (2015a) conducted a two part study to investigate the use of the DelApp in the assessment of attention (50 patients with delirium; 52 patients with dementia; 54 controls). Study 1 was designed to compare the use of the Edinburgh Delirium Test Box (EDTB) with the DelApp. In-patients were recruited from Medicine for the Elderly and orthopaedic wards. For Study 1, EDTB and DelApp sustained attention counting tasks were completed by participants. Study 2 consisted of a case control study which assessed performance on the DelApp in participants with delirium, dementia and individuals without cognitive impairment. Study 1 identified no significant differences between scores on the DelApp and EDTB. Furthermore, no participant stated that they had difficulty completing tasks on the DelApp. These results suggest that the DelApp may be a suitable alternative to the EDTB. Study 2 demonstrated that participants with delirium scored significantly lower than...
patients with dementia and controls. The DelApp was found to have a sensitivity of 98% with 93% specificity in the detection of delirium. There was a moderate negative correlation between DelApp and DRS-R-98 severity scores.

Adamis et al (2015) prospectively assessed 200 older general hospital patients (34 patients with delirium) using the Confusion Assessment Method, Delirium Rating Scale-98 Revised and Montreal Cognitive Assessment Scales. Attention was assessed using four objective tests (digit span, vigilance ‘A’ test, serial 7s and months of the year backwards) along with global clinical subjective rating. The five approaches to assessing attention had statistically significant correlations (p < 0.05). The area under the receiver operating characteristic curve for each individual test in terms of identifying Confusion Assessment Method-defined delirium was as follows: Serial sevens test (0.738); Digit span (0.703); Vigilance ‘A’ test (0.713); months of the year backwards test (0.740); global clinical subjective rating of attention (0.878). Discriminant functional analysis showed that clinical subjective rating of attention in conjunction with months of the year backwards had the best discriminatory ability to identify Confusion Assessment Method-defined delirium and to discriminate patients with delirium from those with dementia and/or normal cognition.

Fick et al (2015) evaluated a number of delirium tests in a general hospital population of 201 participants who had been prospectively enrolled for the 3D-CAM (Marcantonio et al., 2014) validation study. Participants were 75 years or older with mean age of 84. Test performances for different components of the 3D-CAM were compared, including a number of brief bedside tests of attention. The months of the year backwards test performed best, with sensitivity of 83% and specificity of 69%. In comparison, the four digits backwards test had the same sensitivity (83%) but lower specificity (52%). The days of the week backwards test had sensitivity of 50% and specificity of 94%. The three digits backwards test had sensitivity of 45% and specificity of 92%. The authors looked at various test combinations and found that the best two-item screen was the combination of ‘months of the year backwards’ and ‘what is the day of the week?’ with a sensitivity of 93% and specificity of 64%.
Discussion

This review identified studies using a variety of attentional tests to assist the diagnosis of delirium. The quality of the 23 studies included varied as evidenced by the Downs and Black Tool scores which ranged from 11-15.

Areas within the Downs and Black Tool where studies performed consistently poorly included sample selection (items 11 and 12), blinding of patients and assessors (items 14 and 15) and power analysis (item 27). Other areas, such as interventions (item 4), adverse events (item 8) and compliance with interventions (item 19) are most applicable to clinical trials and intervention studies and less so for this selection of predominantly descriptive and phenomenological studies.

In a previous narrative review, Tieges et al (2014) examined 15 articles relating to the use of objective attentional testing in delirium. Through a systematic search, 7 further studies were included in this review. Librarian staff assisted in the initial search strategy and all key databases were searched. This is likely to have led to the higher yield of studies in this review. Following selection of studies, an objective assessment tool (Downs and Black tool) was applied to all papers, adding to the quality of analysis in this review.

Inattention is an essential component of a DSM-5 diagnosis of delirium, although few studies have analysed the ability of attentional tests to detect delirium. Despite the emphasis in DSM-5 on assessment of attention in delirium diagnosis, there is a lack of clear and practical guidance in both DSM-5 and ICD-10 on how attention should be clinically assessed. DSM and ICD describe the different functional elements of attention and their assessment in general terms and do not indicate specific tests or cut-off values. There is clearly a need in everyday clinical practice for tests with defined cut-off values and pass/fail scores on tests of attention that distinguish delirium from non-delirium cases. Of note, none of the studies gave clearly defined guidance on cutoff scores for delirium.

As this review indicates, delirium affects a wide range of patients and as such, different approaches may be better suited to particular populations or clinical settings. Moreover, confounding factors must be taken into account such as the differentiation of delirium from dementia, depression and post-operative states. Bettin et al (1988) have described how the
ideal delirium detection tool should be repeatable, easily administered and reflect the severity of the individual’s delirium. Further qualities of the ideal tool should include brevity, ease of interpretation of results, good coverage of target populations and clear performance cut-offs.

Despite this emphasis on attentional impairment in delirium, there are a number of complicating issues that must be considered. Firstly, guidance from DSM and ICD on the assessment of attention in delirium diagnosis is vague, referring essentially to potential deficits in sustained, selective and divided attention. Secondly, neuropsychological definitions of attention and its components vary and there are a number of different models for conceptualising this area (Knudsen, 2007). Depending on how it is defined, there may be overlap between the concept of attention and concepts such as working memory, executive functioning and attentional deficits may be closely related to other cognitive deficits and vice versa. Thirdly, it is unclear whether attentional deficits in delirium are globally consistent or whether specific aspects of attention are disproportionately impaired, such as sustained attention (Bhat and Rockwood, 2007). Finally, attentional deficits are not exclusive to delirium and occur in other neuropsychiatric conditions with considerable comorbidity and overlap with delirium, most importantly dementia (Kolanowski et al., 2012, Silveri et al., 2007, McGuinness et al., 2010) and depression (Bunce et al., 2012), but also with anxiety disorders (Beaudreau and O’Hara, 2009) and attention deficit hyperactivity disorder (Semeijn et al., 2015). Therefore, focussing exclusively upon detection of attentional deficits in attempting to detect and diagnose delirium is to oversimplify this complex area.

While objective and computerised tests of attention are likely to have a key role in the development of delirium detection programmes, it should also be outlined that research in the area of bedside testing of attention in delirium has highlighted the importance and indeed superiority of subjective assessment by experienced clinicians. O’Keeffe and Gosney (1997a) et al compared a variety of approaches to assessment of attention in older medical inpatients and found that while objective tests of attention could distinguish delirium from non-delirium, a clinical global assessment of attention was the most distinguishing approach to delirium detection. Likewise, when O’Regan et al (2014a) examined the utility of a number of bedside tests of attention in an older inpatient population, they concluded that
the most precise screening method involved a combination of clinical assessment for subjective/objective confusion along with the MBT test. Adamis et al (2015) also examined a number of bedside tests of attention in an elderly inpatient population and again concluded that clinical subjective rating of attention in conjunction with the MBT had the best discriminatory ability to identify delirium. The authors also reported a high negative predictive value for the bedside tests (i.e. good accuracy in excluding delirium) but low positive predictive value (meaning they are not diagnostic for delirium).

The evidence on objective tests suggests that they are useful for excluding non-delirium cases. However, more detailed assessment by trained clinicians is subsequently required in order to confirm delirium status in those who have screened positive initially with objective tests. This two-phase approach, along with daily screening for all those at risk, is advocated by the National Institute of Health and Care Excellence (NICE) guidelines (Young et al., 2010).

Tests of attention

While all of the tests included in this review can be described broadly as tests of attention, there is a wide variation between the tests in terms of neuropsychological domains covered and level of complexity of tests, as outlined in Table 9. Therefore, while attention span is emphasised in the digit and spatial span tests, sustained attention is emphasised in most of the other tests. Other domains covered include visual search and scanning, inhibition of irrelevant information, processing speed and executive function.

In this review, The Digit Span Test was the most frequently studied test of attention (10 studies) and was consistently successful in the detection of delirium. Spatial span testing was examined in four studies, trail making tests in three, vigilance testing in three and months backwards testing in two (see Table 9 for list of studies including the different tests).

Pompei et al (Pompei et al., 1995) and Christensen et al (Christensen et al., 1996) suggested that patterns of scores on the Digit Span Test could be useful in monitoring the course of delirium. However, as outlined by Pompei et al (Pompei et al., 1995), while performance-based objective tests measure point prevalence of delirium, observer-rated instruments
may have an advantage in that they take into account observations over time, which is essential in view of the transient and fluctuating nature of delirium. Differences in the neurophysiological underpinnings of forward and backward testing with the months backwards test (Meagher et al., 2015) have been described with functional magnetic resonance imaging studies with greater involvement of more complex networks (bilateral middle and inferior frontal gyri, the posterior parietal cortex and the left anterior cingulate gyrus) for backwards cognitive processing. This principle may also apply to digit span testing, with backward testing involving multiple cognitive domains in contrast to a more ‘pure’ emphasis on attention seen in forward testing. Patient factors such as motor and verbal abilities should also be taken into account in choosing the most suitable tests with, for example, the spatial span more suitable than verbal digit span in patients with limited verbal abilities.

**Differentiating delirium and dementia**

Differentiating between delirium and dementia is a key challenge in everyday clinical practice. These efforts are complicated by the phenomenological similarities between the conditions, particularly when attempting to differentiate delirium from dementia complicated by behavioural and psychological symptoms (BPSD). Comorbidity between these conditions is also high, with estimates of as many as two thirds of elderly delirium cases having comorbid dementia (Fick et al., 2002). While 10 studies included in this review included patients with dementia, only 5 examined patients with comorbid dementia and delirium (see Table 10). Furthermore, no study examined levels of dementia severity, dementia subtype or the issue of Mild Cognitive Impairment or Mild Neurocognitive Disorder (APA, 2013).

Many of the studies reviewed herein compared attentional test results in delirium to results in dementia. The attentional tests that proved most effective at distinguishing delirium from dementia included: the attention component of the Cognitive Test for Delirium (including Spatial Span Forward and Backward) (Hart et al., 1997, Jabbar et al., 2011), Digit Span Backward and Digit Cancellation (O’Keeffe and Gosney, 1997a), Spatial Span (Meagher et al.,
2010), The Edinburgh Delirium Test Box (Brown et al., 2011c) and the DelApp (Tieges et al., 2015b). Jabbar et al (Jabbar et al., 2011) found that participants with comorbid delirium-dementia had greater delirium severity compared to those with delirium-alone as evidenced by higher DRS-R98 severity scores and lower CTD scores.

Post-operative delirium

Post-operative delirium is commonly linked to prolonged hospitalisations, increased cost of care, disruptive behaviour and incontinence and occurs frequently even for individuals undergoing elective surgery (O'Regan et al., 2013). Lowery et al (2008) compared neuropsychological test performance post operatively in patients with and without delirium. Patients with delirium showed reduction in accuracy on a vigilance task. In contrast, the patients without delirium showed a slowing of reaction on a choice reaction task but no change in accuracy. Beglinger et al (2011) examined the use of the Trail Making test in a population undergoing stem cell transplantation. There were no differences in performance in the Trail Making Test parts A and B when comparing those who developed delirium to those who did not. However, both groups (delirious and non-delirious) undergoing transplantation performed more poorly than a control group. Brown et al (2011b) found that the Stroop and Digit Span tests could accurately identify delirium in the post-operative period in a study of 37 patients, 9 of whom developed delirium after undergoing cardiac surgery. These studies highlight how disturbances to cognitive function occur in many patients post-operatively, including many who do not develop syndromal delirium. Potential clinical practice implications would include the introduction of routine cognitive assessments pre and post operatively, especially in those with pre-existing cognitive impairment and other risk factors for delirium.

Depression

Depression is an important comorbidity for delirium and overlap of symptomatology, particularly in cases of hypoactive delirium, may complicate detection and delay appropriate
treatment for both conditions. It has been suggested that the two conditions share similar pathophysiological mechanisms, involving disturbances in stress and inflammatory responses, monoaminergic and melatonergic signalling (O’Sullivan et al., 2014). A recent review has reported a specific association between delirium and depression in patients with hip fracture (Nelson et al., 2015). Despite the importance of depression, only two of the studies included in this review included depressed patients (Hart et al., 1997, Leonard et al., 2009). These studies suggest that disturbances of attention are typical of depression but at a considerably lower level of severity than delirium.

‘Head-to-head’ studies

There is a lack of research comparing tests of attention ‘head to head’ in terms of clinical utility and coverage. Adamis et al (2015) examined the performance of 4 different tests of attention in delirium diagnosis (serial sevens test, digit span, Vigilance ‘A’ test and months of the year backwards (MBT) and found area under curve values ranging from 0.70 to 0.74, with MBT having the highest value. They concluded that objective tests are useful for prediction of non-delirium but lack specificity for a delirium diagnosis and that global attentional deficits were more indicative of delirium than deficits of specific domains of attention. Furthermore, they concluded that global clinical subjective rating of attention in conjunction with MBT had the best discriminatory ability to identify CAM-defined delirium. O’Regan et al (2014a) examined the months of the year backwards (MBT) test, SSF4 (spatial span forwards test with a cut-off of 4) and a question about evidence of confusion in a sample of 265 general hospital inpatients. The most precise screening method overall was achieved by simultaneously performing MBT and assessing for subjective/objective confusion (sensitivity 93.8%, specificity 84.7%. In older (age 69 and over) patients, MBT alone was most accurate whereas in younger patients (age less than 69) a simultaneous combination of SSF4 and either MBT or assessment of subjective/objective confusion was best. A potential clinical implication from this finding is that brief stand-alone tests of attention may be sufficient for detection of delirium in older patients whereas more challenging tests in combination with assessment of subjective/objective confusion may be needed in younger populations.
In another ‘head to head’ comparison, Fick et al (2015) found that the best single test for delirium was the months of the year backwards test, with sensitivity of 83% and specificity of 69%.

Younger populations

Apart from two studies (Rajlakshmi et al., 2013, Trzepacz et al., 1986), all studies focussed on older populations (i.e. patients over age 65). Performance norms differ considerably for common tests of attention (Meagher et al., 2015), depending on age, and adjustments must be made accordingly for younger populations. In younger populations, more demanding tests may be required to demonstrate significant inattention.

Acceptability to patients and assessors

Acceptability to patients and assessors, along with their preferences for tests used, was not considered in any of the studies in this review. The skill-sets and levels of expertise of treating clinicians should also inform the choice of screening tools used in different clinical contexts. Moreover, studies did not describe the frequency of patient non co-operation with cognitive testing even though in real-world practice it is well recognised that a significant number of patients are unable to engage with cognitive testing. A recent survey (Bellelli et al., 2014) of doctors, nurses, psychologists and physiotherapists found that possible causes of delirium were under-assessed by half of doctors and by the majority of other professionals and while nurses, psychologists and physiotherapists did not answer the case vignettes, doctors identified the correct answer in most cases. Screening tools used by professionals with lower levels of expertise and training should be more explicit and consistent in presentation and rating.

In view of the varied presentations of delirium in diverse clinical contexts, future research should aim to identify optimal assessment tools tailored to these different clinical environments. Development of such tools should also take into account patient and clinical
characteristics such as comorbid neuropsychiatric disorders (especially dementia), sensory impairment, mobility problems and problems with language and expression.
Computerised tests

There is a growing trend towards using computerised-assisted technologies in the detection of delirium (Lowery et al., 2008, Brown et al., 2011d, Tieges et al., 2015b). Computerised approaches have a number of advantages over traditional pen and paper tests. These include clarity and consistency of presentation, acceptability to patients and assessors and objectivity of testing with reduced need for subjective interpretation of findings. Computerised approaches are also likely to reduce the rater error associated with pen and paper tests. Furthermore, computerised tests are more likely to be useful in medically ill populations with marked frailty, where the assessment process may be hampered by problems with patient mobility, sensory impairment and difficulties with language and expression. Smartphone applications in particular have considerable potential in view of increasing use of such devices by healthcare staff for other purposes and the ease of use when delirium tests are administered using them. The Edinburgh Delirium Test Box, (Brown et al., 2011a) progressed to the subsequent development of the DelApp (Tieges et al., 2015a). The DelApp demonstrated high sensitivity and specificity, as well as an ability to determine delirium severity. It is a portable, simple and effective way to evaluate delirium in the hospital setting and does not heavily rely on clinical judgement and specific training.

Study limitations

The search was restricted to publications written in the English language and did not consider unpublished data. In view of the lack of large blinded studies and studies with power calculation, the minimum requirements for inclusion in this review were studies in the English language that involved the use of brief tests of attention in the assessment of delirium in inpatient populations. Due to the heterogeneity of the tests of attention, no meta-analysis was performed. Nevertheless, the findings allow for some conclusions regarding the usefulness of attention testing to assist in delirium diagnosis as well as the relative merits of different tests.

The Downs and Black tool, with its emphasis on clinical trials and intervention studies, is not suited in the sections relating to interventions, adverse events, blinding of study subjects
and assessors and power calculations for the types of study included in this review, which are primarily phenomenological studies of small populations.
Conclusion

Delirium is a common neuropsychiatric condition that is frequently under detected in the hospital setting (O'Regan et al., 2014a). Therefore, more consistent and accurate detection of delirium is a key challenge for healthcare services as we provide for increasingly aged populations. This review identified a variety of objective bedside tests that are used in the detection of delirium. Based on the available evidence, future delirium screening and detection programmes will ideally involve a two phase process. The first phase will involve screening with objective tests of attention that can be used by staff from all backgrounds and levels of clinical expertise. The second phase will involve more detailed assessment by trained experts. The screening tools used in the first phase should be tailored to the clinical characteristics of the target populations and take into account factors such as age, comorbidity with other mental disorders such as dementia and factors that may impact on level of cooperation and completion of tests, such as motor and verbal abilities (O'Connell et al., 2014).

Future studies should focus on the use of effective and reliable objective tests of attention that can be easily administered by health care providers. These studies must also take into account the effect that age, comorbid dementia, depression and post-operative states have on performance in objective tests of attention. In the development of diagnostic tools for delirium, there is a need to recognise that there are multiple causes of cognitive impairment of varying levels of severity, reversibility and chronicity and with numerous aetiologies. In many cases, delirium may not be present, but other cognitive problems may impact upon test performance in delirium diagnosis. For reasons of patient and user acceptability, objectivity, reproducibility and accurate recording, electronic and smartphone devices are likely to offer advantages over traditional pen and paper tests.
Table 10: Summaries of studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Design</th>
<th>Total pop</th>
<th>No. with delirium</th>
<th>No. with comorbid dementia and delirium</th>
<th>Clinical question</th>
<th>Population</th>
<th>Subject Groups</th>
<th>Age</th>
<th>Sex (%M)</th>
<th>Reliability/validity</th>
<th>Test of attention</th>
<th>Scores</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Trzeacz et al., 1986)</td>
<td>Cross-sectional</td>
<td>40</td>
<td>12</td>
<td>0</td>
<td>Can neuropsychological tests be used in the early detection of hepatic encephalopathy?</td>
<td>Delirium (N=12)</td>
<td>18-58</td>
<td>37%</td>
<td>?</td>
<td>Trail Making Test Part A</td>
<td>For delirium: 1 had normal score; 2 had mildly impaired score; 9 had impaired score. For delirium: 1 had normal score; 3 had mildly impaired score; 8 had impaired score. Trail Making Test scores were significantly associated with delirium (Trail Making Test Part A: p=0.001 Trail Making Test Part B: p=0.002) Trail Making Test Part A: sensitivity=75%, specificity=82.1% Trail Making Test Part B: sensitivity=91.7%, specificity=73.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Trzeacz et al., 1988a)</td>
<td>Cross-sectional</td>
<td>47</td>
<td>20</td>
<td>0</td>
<td>Can the Trail Making Test be used in the detection of delirium?</td>
<td>Delirium (N=20)</td>
<td>58.8 ± 16.5</td>
<td>35%</td>
<td>?</td>
<td>Trail Making Test Part A</td>
<td>Delirium; 125 ± 61, Schizophrenia; 89 ± 42 Delirium; 330 ± 112, Schizophrenia; 240 ± 123 Delirium and schizophrenia groups severely impaired on Trailmaking Tests. 9 participants with delirium could not attempt tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Pompel et al., 1998)</td>
<td>Cross-sectional</td>
<td>42</td>
<td>8</td>
<td>0</td>
<td>Can the Digit Span Test and Vigilance A be used in the detection of delirium?</td>
<td>Delirium (N=61)</td>
<td>74.8 ± 7.3</td>
<td>44%</td>
<td>?</td>
<td>Vigilance A test</td>
<td>Worst, median and best scores reported for Vig A and Digit Vigilance ‘A’ Test; 61% sensitive and 77% specific. +LR (95% CI)= 2.7 (2.0-3.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Study Design</td>
<td>N1</td>
<td>N2</td>
<td>N3</td>
<td>N4</td>
<td>Test Used by Clinicians to Identify Patients with Delirium?</td>
<td>Controls (N=367)</td>
<td>Span Tests</td>
<td>LR (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-------------------------------------------------------------</td>
<td>-----------------</td>
<td>------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>Prospective cohort</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td></td>
<td>Test be used by clinicians to identify patients with delirium? (Aged &gt; 65 years)</td>
<td>74.2 ± 6.7</td>
<td>44%</td>
<td>-LR (95% CI)= 0.5 (0.4-0.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospective cohort</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td></td>
<td>Can digit span forward and digit span backward be used to monitor delirium severity?</td>
<td>Delirium (N=10)</td>
<td>65.0 ± 7.2</td>
<td>100%</td>
<td>Digit span forward</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospective cohort</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td></td>
<td>Can Visual Attention Span Backwards and Forwards adequately screen for delirium?</td>
<td>Delirium (N=19)</td>
<td>61.3±1.3</td>
<td>52.6%</td>
<td>Visual Attention span Backward</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospective cohort</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td></td>
<td>What is the relationship between global assessment of attentiveness and tests of attention?</td>
<td>Acute geriatric in-patients</td>
<td>81.9 ± 5.0</td>
<td>Not specified</td>
<td>Digit span forwards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997b</td>
<td>Cross-sectional</td>
<td>87</td>
<td>18</td>
<td>4</td>
<td></td>
<td>Early assessment Del. scores&lt; Con. on DSB,DCT1 and DCT2</td>
<td>DSF: Del.; 4.7 (1.0), Dem.; 5.2 (0.8), Con.; 6.4 (0.7)</td>
<td>DSB: Del.; 2.4 (0.9), Dem.; 3.6 (0.8), Con.; 5.1 (0.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>Cross-sectional</td>
<td>37</td>
<td>22</td>
<td>0</td>
<td></td>
<td>Is the digit</td>
<td>Delirium (N=22)</td>
<td>72.4</td>
<td>100%</td>
<td>Digit Span Forward</td>
<td>Early assessment Del. scores&lt; Con.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Type</td>
<td>Patient/Control Groups</td>
<td>Data Points</td>
<td>Findings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>et al., 1998</td>
<td>prospective cohort</td>
<td>ward in-patients with delirium and controls vs. controls (N=15)</td>
<td>CAM-ICU: Vigilance A task; picture recognition.</td>
<td>Del. : 17.5 ±3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(McNicol et al., 2005)</td>
<td>cross-sectional</td>
<td>ICU patients. (Aged&gt;65 years)</td>
<td>Delirium (N=8); Dementia (N=3); Del. and Dem. (N=7); Controls (N=4)</td>
<td>CAM/MMSE recall and WORLD backwards; Digit Span.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Bosio et al., 2006)</td>
<td>cross-sectional</td>
<td>Cancer patients from medical oncology and palliative care units</td>
<td>Delirium (N=66); Non-Delirium (N=40)</td>
<td>none reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Meagher et al., 2007)</td>
<td>prospective cohort</td>
<td>Palliative care in-patients</td>
<td>Delirium (N=83); Delirium-Dementia (N=17)</td>
<td>CTD Spatial span: 2.1 (1.8) CTD Vigilance: 2.4 (2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Lower et al., 2008)</td>
<td>prospective cohort</td>
<td>Elective hip and knee replacement patients. (Aged &gt;70 years)</td>
<td>Delirium (N=14); Controls (N=80)</td>
<td>Correlation between DRS-R-98 attention and CTD attention: r= - 0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Leong et al., 2008)</td>
<td>prospective cohort</td>
<td>100 consecutive palliative care admissions</td>
<td>Syndromal delirium (N= 29)</td>
<td>impaired in those with delirium in comparison to those with depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total group: mean 49%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 2009

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Comparator</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meagher et al., 2010</td>
<td>Prospective cross-sectional</td>
<td></td>
<td>Delirium (N=40)</td>
<td>Dementia (N=20)</td>
<td>Comorbid delirium-dementia (N=40)</td>
<td>Controls (N=40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>68.7 ± 12</td>
<td>78.6 ± 7.8</td>
<td>74.9 ± 8.5</td>
<td>66.3 ± 47%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57%</td>
<td>55%</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spatial Span Forwards</td>
<td>Spatial Span Backwards</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Del.; 6.6 ± 1.9</td>
<td>Dem.; 4.1 ± 2.1</td>
<td>Del.-Dem.; 2.8 ± 2.3</td>
<td>Con.; 5.7 ± 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SSF: Del.; 2.6 ± 1.9</td>
<td>Dem.; 4.1 ± 2.1</td>
<td>Del.-Dem.; 2.8 ± 2.3</td>
<td>Con.; 5.7 ± 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spatial Span Forwards</td>
<td>Spatial Span Backwards</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Del.; 1.3 ± 1.6</td>
<td>Dem.; 2.1 ± 1.7</td>
<td>Del.-Dem.; 1.3 ± 1.7</td>
<td>Con.; 4.0 ± 1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CTD attention scores: Del. and Del. - Dem &lt; Dem. and Con. p&lt;0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CTD total score: Del.; 11.9 ± 6.4</td>
<td>Del. - Dem.; 8.5 ± 5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Attention scores of CTD and DRS-R98 compared:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Score of 4-6 on CTD = 1.4 ± 0.9 on DRS-R-98,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Score of 2 on CTD = 2.4 ± 0.7 on DRS-R-98,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Score of 0 on CTD = 2.7 ± 0.6 p&lt;0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2011

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Comparator</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jabbar et al., 2011</td>
<td>Prospective cross-sectional</td>
<td></td>
<td>Delirium (N=40)</td>
<td>Comorbid delirium and dementia (N=40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>77.2 ± 7.7</td>
<td>81.3 ± 7.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50%</td>
<td>48%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Digit span</td>
<td>Digit span (post op.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Del.; 75.7 ± 56%</td>
<td>Del.; 75.7 ± 56%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pre to post-operative change in Del &gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Ref.</td>
<td>Design</td>
<td>Questions</td>
<td>Cognitive Test</td>
<td>Delirium</td>
<td>No Delirium</td>
<td>Controls</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Brown et al., 2011c</td>
<td>Cross-sectional</td>
<td>Would patients with delirium show impairments of sustained visual attention using the Edinburgh Delirium Test Box?</td>
<td>Stroop</td>
<td>Delirium: Median (IQR)</td>
<td>81.7 (8.8)</td>
<td>79.9 (8.0)</td>
</tr>
<tr>
<td>Beugin et al., 2011</td>
<td>Prospective</td>
<td>Will bone marrow transplant patients with delirium perform worse than bone marrow transplant patients without delirium and controls on tests of attention?</td>
<td>Trail Making Test parts A and B</td>
<td>Delirium (N=19)</td>
<td>No delirium (N=33)</td>
<td>Controls (N=10)</td>
</tr>
<tr>
<td>Meagher et al., 2012</td>
<td>Prospective</td>
<td>Can the CTD be used to accurately distinguish Palliative care in-patients</td>
<td>CTD</td>
<td>Delirium (N=100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study (Author et al., Year)</td>
<td>Study Type</td>
<td>Sample Size</td>
<td>No. Delirium</td>
<td>No. No Delirium</td>
<td>Cognitive Symptoms</td>
<td>Non-Cognitive Symptoms</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
<td>-------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
</tbody>
</table>
| (Rajlakshmi et al., 2013)   | Cross-sectional | 84          | 0            | 84             | Delirium (N=84)   | Medical and surgical inpatients referred to liaison psychiatry service. | CTD attention mean score: 2.35 ±1.58 (0-4)  
Span: 2.4/6(1.6)  
Vigilance: 1.7/6(1.6) | No scores reported | Correlations between DRS-R-98 attention item with spatial span and vigilance task: r = - 0.10 and r = - 0.15 respectively, NS |
| (O’Regan et al., 2014b)     | Cross-sectional | 48          | 19           | 39             | Delirium (N=48)   | Hospital inpatients | MBT 83.3% sensitive (95% CI 69.8 to 92.5) and 90.8% specific (95% CI 86.1 to 94.3)  
Simultaneous use of MBT and SSF 4 had a sensitivity of 87.5% and specificity of 86.3% |
| (Tieges et al., 2015b)      | Case-control  | 50          | 85           | 65             | Delirium (N=50)   | In-patients from Medicine of the Elderly and Orthopaedic wards (age >60 years) | DelApp scores:  
Del. vs Dem.: U=174.5, p<0.001, 95% CI= (50.58, 74.38)  
Dem. vs Con.: U=985.5, p<0.001, 95% CI= (69.25, 92.84) | Is the DelApp an acceptable tool to use in the assessment of attention in older hospitalised patients? |
| (Adami et al., 2015)        | Prospective   | 201         | 34 (17%)     | 167 (83%)      | Consecutive general hospital admissions | Digit span Vigilance ‘A’ test Serial 7s MBT | Clinical subjective rating combined with MBT had best discriminatory ability |
| (Fick et al., 2015)         | Prospective   | 42          | 20 (47%)     | 22 (53%)       | Consecutive general hospital | Total = 201  
42 (21%) with delirium | MBT the single best test, with sens. of 83% and specificity of 69% | What is the best single test? |
<p>| 2015) | delirium test within the 3D-CAM, and best combination of tests? | inpatient admissions | 84 years | Days of week backwards 3 digits backwards |
|------------|------------|--------------------------|----------------|-------------------------------|----------------|--------------------------------|----------------|---------------------------------|-------------------------------|---------------------------------|-------------------------------|---------------------------------|--------------------------------|-----------|
| (Trzepacz et al., 1986) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N/A | Y |
| (Trzepacz et al., 1988a) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N/A | Y |
| (Pompei et al., 1995) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 0 | 1 | 1 | 1 | Y |
| (Christensen et al., 1996) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | 1 | 1 | 0 | 0 | 1 | Y |
| (Hart et al., 1997) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 0 | 0 | 1 | Y |
| (O’Keeffe and Gosney, 1997a) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | Y |
| (Bettin et al., 1998) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 0 | 0 | 1 | Y |
| (McNicol et al., 2005) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | Y |
| (Bosissio et al., 2006) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | Y |
| (Meagher et al., 2007) | 0 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N |
| (Lowery et al., 2008) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N |
| (Leonard et al., 2009) | 0 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N |
| (Meagher et al., 2010) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N |
| (Jabbar et al., 2011) | 1 | 1 | 1 | N/A | 1 | 1 | 1 | N/A | N/A | 1 | 1 | 1 | 1 | N |</p>
<table>
<thead>
<tr>
<th>Study Reference</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>N/A</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>N/A</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>N/A</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>N/A</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>N/A</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown et al., 2011b</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Brown et al., 2011c</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Beglinger et al., 2011</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Meagher et al., 2012</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Rajakshmi et al., 2013</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>O’Regan et al., 2014b</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tieges et al., 2015b</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Adams et al., 2015</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fick et al., 2015</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
### Table 12: Downs and Black Tool parts 14-26

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(Trzepacz et al., 1986)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>Y</td>
<td>13</td>
</tr>
<tr>
<td>(Trzepacz et al., 1988a)</td>
<td>0</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>Y</td>
<td>14</td>
</tr>
<tr>
<td>(Pompei et al., 1995)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>Y</td>
<td>14</td>
</tr>
<tr>
<td>(Christensen et al., 1996)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>11</td>
</tr>
<tr>
<td>(Hart et al., 1997)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>11</td>
</tr>
<tr>
<td>(O'Keeffe and Gosney, 1997b)</td>
<td>0</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>15</td>
</tr>
<tr>
<td>(Bettin et al., 1998)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>11</td>
</tr>
<tr>
<td>(McNicoll et al., 2005)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>14</td>
</tr>
<tr>
<td>(Bosio et al., 2006)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Meagher et al., 2007)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Lowery et al., 2008)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Leonard et al., 2009)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Meagher et al., 2009)</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>Study Reference</td>
<td>Date (Year)</td>
<td>Cluster</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>(Jabbar et al., 2010)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Brown et al., 2011b)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Brown et al., 2011d)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Beglinger et al., 2011)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Meagher et al., 2012)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>(Rajakshmi et al., 2013)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(O'Regan et al., 2014b)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Tieges et al.,</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Adamis et al.,</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
<tr>
<td>(Fick et al., 2015)</td>
<td></td>
<td>N/A</td>
<td>1</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>14</td>
</tr>
</tbody>
</table>
### Table 13: Recommendations for future research

- Tests delivered via computerised and smartphone application platforms can enhance reliability of testing and interpretation
- More studies considering populations with comorbid dementia and delirium
- Increased focus on dementia severity, Mild Neurocognitive Disorder and dementia subtypes
- Increased focus on younger populations, in whom more demanding tests of attention may be required
- ‘Head-to-head’ studies comparing different tests of attention in identification of delirium
Chapter 3: Methods

Subjects and Design

We conducted a prospective cross-sectional study of neuropsychiatric symptoms and cognitive performance in referrals to a psychiatry consultation-liaison service of patients with delirium, dementia, comorbid delirium-dementia, as well as comparison subjects with no neurocognitive disorder (NNCD). Cases with altered mental state were identified on daily rounds by the medical team and referred for assessment and diagnosis by the research team. A purposive sample was collected, with an emphasis on obtaining comparable numbers for the four clinical groups (i.e. those with delirium, dementia, comorbid delirium-dementia and NNCD). Formal power calculation was not performed, but a sample size of 190-200 was estimated as being sufficient for this study, based on previous similar research conducted by this research team (O’Regan et al, 2014a). Assessments were conducted by raters (ML, FA, HO’C, WE, OW, DM) 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.

Delirium was diagnosed according to a cutoff score of ≥ 15 on the severity scale of the DRS-R98 (Trzepacz et al., 2001) and / or presence of DSM 5 criteria (APA, 2013) based upon a full clinical assessment. Dementia was defined as a clear history of documented DSM-5 Major Neurocognitive Disorder (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 (Jorm, 1994). Comorbid delirium-dementia was defined as the presence of both disorders.

The raters administered a battery of delirium screening and diagnostic tools and performed either the initial ‘Rater A’ assessment or the follow-up ‘Rater B’ assessment. In as far as was practically possible, Raters A and B were blinded to the outcomes of their respective findings. Therefore, all patients had a range of assessments performed by two different assessors. Rater A assessments lasted approximately 60 minutes and Rater B assessments were performed within 4-6 hours, lasting 30-45 minutes.
Patients were referred by medical and surgical teams and were assessed for the presence of DSM-5 criteria for delirium. Participants were assessed by Rater A using the revised Delirium Rating Scale (DRS-R98) (Trzepacz et al., 2001) and Cognitive Test for Delirium (CTD) (Hart et al., 1996), Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988), Short-IQCODE (Jorm, 1994) and Delirium Motor Subtype Scale (Meagher et al., 2014). These instruments are described in further detail below. Rater B assessments involved Intersecting Pentagons, Clock Drawing Test, Letter Shape Drawing Test and the Lighthouse app test. Again, these tests are outlined in more detail below. All assessments were conducted by experienced clinicians and Rater B assessments were briefer and more focussed than Rater A assessments: these factors helped minimise participant fatigue.

**Ethical approval**

Ethical approval was granted by the Midwestern Regional Ethics Committee/University Hospital Limerick Regional Ethics Committee. 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-interventional nature of the study, the University Hospital Limerick Regional Ethics Committee approved 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 accordance with the Helsinki Guidelines for Medical Research involving human subjects (World Medical Association, 2004). Information sheets (see appendix) were made available for participants, their families and ward staff.

**Definitions for delirium and dementia**

A DRS-R98 severity score of ≥ 15 was taken to indicate syndromal delirium. A mean Short IQCODE score of ≥ 3.5 was considered indicative of dementia. Participants were thus categorized into four groups: delirium in the context of background dementia, delirium alone, dementia alone and those who were cognitively intact or normal (i.e. not having either delirium or dementia, i.e. having no neurocognitive disorder, NNCD).
**Statistical analyses**

Statistical analysis was conducted using SPSS-19. Demographic and rating scale data are expressed as means plus standard deviation. Continuous variables (e.g. age, total DRS-R98 and SIQ Code scores) were compared by one way ANOVA with independent t-tests used for post hoc comparisons. Non-normal data (e.g. cognitive test scores) were compared with Wilcoxon signed ranks and Mann-Whitney U tests for between group comparisons. Box’s Test reveals that variance-covariance matrices are equal (Box’s M=17.8, F=2.9, p<0.001) and thus the assumptions for discriminant analysis were not held. In order to evaluate attention tests (and their combinations) in terms of distinguishing delirium, a logistic regression analysis was performed with the binary classification (DSM-IV delirium or not) as the dependent variable and performance on the attention, vigilance and visuospatial tests as independent variables. Finally, for the binary tests of cognition (and their combinations), sensitivity and specificity as well as positive and negative likelihood ratio, Positive Predictive Value (PPV), and Negative Predictive Value (NPV) were calculated, with confidence intervals testing significance at 95%. In comparing the test performance of the novel Lighthouse app with ‘gold standard’ tests, *a priori* cut-offs were not defined. However, optimal test sensitivity was considered the priority (as opposed to test specificity) in view of the clinical importance of delirium, the significant morbidity and mortality associated with missed diagnosis and the fact that the Lighthouse test is envisaged as an initial screening phase in delirium detection and diagnosis.
**Rater A clinical assessments**

*Principal Rater A assessments*

Delirium Etiology Checklist (Trzepacz and Meagher, 2008, Trzepacz et al., 2011)

Mini-Mental State Examination (MMSE) (Folstein et al., 1975)

Delirium Rating Scale revised version (DRS-R98) (Trzepacz et al., 2001)

Cognitive test for Delirium (CTD) (Hart et al., 1996)

Short-IQCODE (Jorm, 1994)

Neuropsychiatric Inventory (NPI-Q) (Kaufer et al., 2000)

Clinical Dementia Rating Scale (CDR) (Hughes et al., 1982)

Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988)

Delirium Motor Subtype Scale 4 (DMSS-4) (Meagher et al., 2014)

*Additional cognitive testing included a number of common bedside tests:*

Spatial Span Forwards (SSF) (Hart et al., 1996)

Spatial Span Backwards (SSB) (Hart et al., 1996)

Vigilance A test (Hart et al., 1996)

Vigilance B test (Hart et al., 1996)

Days of week/months of year backwards (Katzman et al., 1983)

WORLD backwards (Folstein et al., 1975)

Global assessment of visuospatial abilities (Trzepacz et al., 2001, Trzepacz et al., 2010)
Tests and rating scales

*The Delirium Etiology Checklist (DEC) (Trzepacz and Meagher, 2008)*

This was used to document etiological underpinnings of delirium. This standardised checklist captures delirium etiology according to twelve categories. The presence and suspected role of multiple potential causes were documented for each case of delirium, rated on a 5-point scale for degree of attribution to the delirium episode, ranging from ‘ruled out/not present/not relevant’ (0) to ‘definite cause’ (4). The DEC allows for multiple concomitant causes as contributing etiologies and thus is sensitive to the multifactorial nature of delirium etiology.

*Mini-Mental State Examination (MMSE) (Folstein et al., 1975)*

This is a universally recognised brief 30 point cognitive screening test covering the domains of orientation, attention and calculation, registration, short-term memory, language and visuospatial skills. A cut-off of 23 or less is generally taken to imply cognitive impairment and suggests further cognitive assessment is indicated. The test can easily be administered at the bedside and takes 5-10 minutes to complete.

*Delirium Rating Scale-Revised-98 (DRS-R98) (Trzepacz et al., 2001)*

This was 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. For determination of
item frequencies in this study, any item score ≥ 1 was considered as being “present”. It has high interrater reliability, sensitivity and specificity for detecting delirium in mixed neuropsychiatric and other hospital populations including dementia, depression and schizophrenia.

_Cognitive Test for Delirium (CTD) (Hart et al., 1996)_

This is specifically designed to allow for wide coverage in the assessment of a range of neuropsychological domains in hospitalized delirium patients, including those with sensory and other impairments that might impact upon their ability to undergo cognitive testing (e.g. those who are intubated or unable to speak or write). It assesses 5 domains: orientation, attention, memory, comprehension and vigilance. Each domain is assessed using an operationalised approach to testing based upon suitable cognitive tests. Scores range between 0-30 with higher scores indicating better cognitive function. The CTD reliably differentiates delirium from other neuropsychiatric conditions including dementia, schizophrenia and depression (Hart et al., 1997).

_The IQCODE - short version (Jorm, 1994)_

This 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 short-IQCODE takes approximately 10 minutes to administer. The total score divided by the number of questions provides a mean item score where ratings of 3.5 or more are considered indicative of longstanding cognitive difficulties and dementia.
**The Neuropsychiatric Inventory Questionnaire (NPI-Q) (Kaufer et al., 2000)**

This was developed for assessing neuropsychiatric symptoms in patients with Alzheimer’s disease and other neurodegenerative disorders. The NPI-Q is a self-administered short questionnaire version of the NPI intended for use in everyday clinical practice. Neuropsychiatric symptom severity is assessed in the same way as the original NPI. The NPI-Q includes ten behavioural and two neurovegetative items that are assessed by an informed caregiver who is knowledgeable about the patient’s daytime and night-time behaviours. Symptoms are rated over the past four weeks. Each of the 12 symptom domains is assessed by a screening question derived from the NPI-Q that covers symptom manifestations with anchor points for symptom severity rated on a three point scale and caregiver distress ratings rated on a five point scale. The total NPI-Q severity score represents the sum of individual symptom scores and ranges from 0 to 36. The NPI-Q provides a brief, reliable, informant-based assessment of neuropsychiatric symptoms and associated caregiver distress that may be suitable for use in general clinical practice. It is used as a self-administered questionnaire with written instructions; all domains for severity (0-3) and caregiver distress (0-5) are well anchored. The wording is taken directly from NPI screening questions. It has interscale correlations of 0.91 for NPI-total and NPI-Q. The total severity NPI-Q represents the sum of the scores 0-36 and the total caregiver distress represents the sum of the scores 0-60.

**The Clinical Dementia Rating (CDR) scale (Hughes et al., 1982)**

This was developed for a prospective study of mild senile dementia—Alzheimer type (SDAT). It is a global rating device which was found to distinguish unambiguously among older subjects with a wide range of cognitive function, from healthy to severely impaired. Global CDR scores equate to severity of cognitive impairment and dementia as follows: 0 (no cognitive impairment); 0.5 (Mild Cognitive Impairment); 1 (mild dementia); 2 (moderate dementia); 3 (severe dementia).
The Cornell Scale for Depression in Dementia (CSDD) (Alexopoulos et al., 1988)

This measures depression severity in older adults with and without dementia; a cut-off point of 6 or more indicates clinically significant symptoms. The scale consists of 19 items covering the domains of mood-related signs, behavioural disturbance, physical signs, cyclic functions and ideational disturbance. Ratings should be based on symptoms and signs occurring during the week before interview. No score should be given if symptoms result from physical disability or illness.

The Delirium Motor Subtype Scale 4 (DMSS-4) (Meagher et al., 2014)

This scale allows for rapid assessment of clinical subtypes in delirium and has high concordance with the longer and well-validated Delirium Motor Subtype Scale (DMSS). The DMSS-4 was derived from the DMSS following latent class analysis that identified four classes based upon four items, two hyperactive (increased amount of activity and loss of control of actions) and two hypoactive (decreased speed of actions and reduced verbal output).
Cognitive testing

Spatial span forwards (SSF)

This was conducted according to the description in the Cognitive Test for Delirium (CTD) (Hart et al., 1996). The SSF 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 x 1cm 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).

Spatial Span Backwards (SSB)

Similarly, the SSB 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 description in the CTD (Hart et al., 1996).

Vigilance A test

The vigilance “A” test was also derived from the CTD scale (Hart et al., 1996). 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) cut-off, 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 (Hart et al., 1996).

**Months Backward 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 subjects best performance was noted. Scoring was according to that proposed by Meagher et al (2015) where patients were rated as unable to engage (0), able to engage but unable to reach July without more than one error (1), able to reach July with less than two errors (2), able to reach January without error (3). In subjects over age 60, failure to reach July without more than one error of omission is considered to equate with clinically significant inattention (and thus a failed test).

**WORLD backwards**

The WORLD backwards test was applied according to the format of the MMSE (Folstein et al., 1975) Each participant was asked to spell WORLD backwards. A point was awarded for each letter correctly identified and the total number of points was recorded. Patients who self-corrected their own mistakes without prompting when spelling WORLD backwards were given the point for each letter they were able to correct. Additionally if a patient was unable to recite a particular letter, they were told that letter in order for them to make an accurate attempt at the next letter to follow. Failure to correctly recite all five letters is considered to equate with clinically significant inattention (and thus a failed test).
Global assessment of visuospatial abilities

Visuospatial ability was rated according to a four point scale based upon the general guidance for DRS-R98 item 13 (Trzepacz et al., 2010) 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 were rated as: no evidence of impairment (3), mild impairment as evidenced by minor difficulty with probes and/or navigating surroundings (2), moderate impairment in terms of responses to probes and/or reported need for redirection to keep from getting lost in the environment or trouble locating familiar objects in immediate environment (1), and severe impairment evidenced by inability correctly respond to any probes and navigate environment (0). In keeping with DRS-R98 guidance, a score of <2 suggests pathological impairment of visuospatial abilities.
Rater B assessments

*Rater B assessments and scales (conducted by a different Rater and within 4-6 hours of assessment by Rater A)*

**Lighthouse application tests:**

- Lighthouse Identification (LH-ID) test
- Focussed Attention (FA) test
- Sustained Attention (SA) test

Intersecting pentagons (Folstein et al., 1975, Bender, 1938)

Clock Drawing test (Freedman et al., 1994, Brodaty et al., 2002)

Letter Shape Drawing (LSD) test
Rater B assessments involved smartphone tests developed by the Cognitive Impairment Research Group (CIRG) at the University of Limerick. The background to the development of the Lighthouse application, along with detail on its operation, is included in Background.

**Lighthouse test**

This test involves presentation of a flashing lighthouse on a standard smartphone screen (3” x 5”). The Lighthouse Test is a novel test that has been developed for presentation as a Smartphone app that aims to provide an alternative to conventional verbal bedside tests of attention. It has three stages that are designed to assess comprehension, focused attention and sustained attention respectively.

The stages are as follows: (i) subjects are asked to name the lighthouse object (which is designed to test basic orientation to the task and simple comprehension), (ii) three separate series of lighthouse flashes (3 flashes, 5 flashes, 4 flashes) and after each the subject is asked to report how many flashes occurred, which is designed to assess ability to focus attention, and (iii) three series that are more complex as they involve sequences of flashes that are interrupted by pauses (e.g. 2-4-3) with each series lasting a total of 15-20 seconds. The subject is asked to report the total number of flashes over the sequence. This is designed to assess sustained attention.

**Intersecting Pentagons Test (IPT)**

This geometric copying test is derived from the original Bender Gestalt test (Bender, 1938). 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 developed by Bourke et al (1995) where 6 represents a perfect reproduction and 1 represents the poorest effort. A score of < 4 is considered to reflect a failed performance.
**Clock Drawing Test (CDT)**

The Clock Drawing Test 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 (Freedman et al., 1994). 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 (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.

**Letter and Shape Drawing (LSD) Test**

The Letter and Shape Drawing test (LSD) is a novel test developed by the Cognitive Impairment Research Group (CIRG). It aims to provide a test of visuospatial ability that is simpler than existing bedside tests (e.g. clock drawing) and that can be readily presented via computer-assisted devices to allow for convenient and reliable testing.

An initial design of the LSD was piloted for acceptability and coverage in a series of trials in a nursing home population with varying degrees of cognitive impairment. The test includes a series of 15 designs that link 1cm spheres arranged in increasingly complex 3x3, 4x4, 5x5 and 6x6 grids that the subject copies to an adjacent blank grid (for an example see figure X). Each grid occupies one half of an A5 page. A correct performance requires that all relevant spheres are connected to complete the required shape. Omissions (but not commissions) are rated as errors. As such, performance can be precisely determined as a correct reproduction of the letter or shape or not. Subjects are permitted a single trial of each of the 15 items. The test typically takes 5 minutes to complete. The current version is presented in a pen and paper format but is designed to allow for ease of transfer to a touch sensitive grid on an android tablet platform. Performance was measured by totalling the number of correctly completed grids (scoring range 0-15).
Chapter 4: Results

Description of study subjects

200 participants were assessed between January 2014 and January 2015. Issues around consent in a population such as this, with high levels of neuropsychiatric morbidity, are covered in the section on ethical approval on page 98.

Seven of the 200 participants were unable to take part, in view of their clinical status (e.g. reduced levels of consciousness or extreme agitation/distress). A total of 193 patients (i.e. those over age 60) were included for this analysis [mean age 79.9 ± 7.3; 97 male]. The frequencies of the neurocognitive diagnoses were delirium (n=32), dementia (n=42), comorbid delirium-dementia (n=53) and no neurocognitive disorder (NNCD) (n=66). Demographic, medication and general clinical data for patients from these four groups are shown in Table 8. There were no significant differences between the four groups in respect of age, gender distribution or number of medications received, while the number of psychotropic medications was higher in those with any neurocognitive diagnosis.

The principal underlying etiologies for delirium (n=85) as captured on the DEC (Trzepacz et al., 2011) were; systemic infection (48), metabolic/endocrine disturbance (27), cerebrovascular (13), organ insufficiency (12), seizure-related (7), drug intoxication or withdrawal (4), neoplasm (4), CNS infection (1), traumatic brain injury (1).

Table 14 compares mean scores for the four groups for the DRS-R98 and IQ-CODE. Both delirium groups were more impaired than the dementia group on total scores for the DRS-R98. The mean short IQCODE scores distinguished the groups, with both dementia groups scoring well in excess of the suggested cut off score.
Table 14: Demographic and medication data for the four neurocognitive groups (mean ± SD)

<table>
<thead>
<tr>
<th></th>
<th>Total group (n=193)</th>
<th>Delirium (n=32)</th>
<th>Comorbid delirium-dementia (n=53)</th>
<th>Dementia (n=42)</th>
<th>No Neurocognitive Disorder (n=66)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (%)</td>
<td>97 (50%)</td>
<td>16 (50%)</td>
<td>32 (60%)</td>
<td>16 (38%)</td>
<td>33 (50%)</td>
</tr>
<tr>
<td>Age</td>
<td>79.9 ± 7.3</td>
<td>77.8 ± 8.6</td>
<td>80.2 ± 7.2</td>
<td>81.3 ± 4.9</td>
<td>79.7 ± 7.8</td>
</tr>
<tr>
<td>Total number of medicines</td>
<td>10.0 ± 4.7</td>
<td>9.4 ± 4.1</td>
<td>10.2 ± 4.8</td>
<td>9.9 ± 4.9</td>
<td>10.3 ± 4.8</td>
</tr>
<tr>
<td>Number of Psychotropic medications*</td>
<td>1.6 ± 1.6</td>
<td>2.1 ± 1.7</td>
<td>2.2 ± 1.7</td>
<td>1.7 ± 1.5</td>
<td>0.9 ± 1.1</td>
</tr>
<tr>
<td>MMSE*</td>
<td>18.9 ± 7.1</td>
<td>17.1 ± 6.6</td>
<td>12.4 ± 6.6</td>
<td>20.7 ± 3.9</td>
<td>23.8 ± 4.7</td>
</tr>
<tr>
<td>DRS-R98 total†*</td>
<td>16.7 ± 9.1</td>
<td>25.9 ± 4.7</td>
<td>25.1 ± 5.6</td>
<td>12.4± 4.8</td>
<td>9.0± 5.2</td>
</tr>
<tr>
<td>Short IQCODE**</td>
<td>3.7 ± 0.7</td>
<td>3.1 ± 0.1</td>
<td>4.5 ± 0.5</td>
<td>4.1 ± 0.5</td>
<td>3.0 ± 0.1</td>
</tr>
<tr>
<td>CDR global score</td>
<td>0.83 ± 0.91</td>
<td>0.26 ± 0.29</td>
<td>1.83 ± 0.92</td>
<td>1.10 ± 0.70</td>
<td>0.22 ± 0.27</td>
</tr>
</tbody>
</table>

†delirium and comorbid delirium-dementia > dementia at p < 0.001,
*all three neurocognitive groups > NNCD at p < 0.001.

CDR: no significant difference between NNCD and Delirium alone. NNCD score less than both Dementia and comorbid delirium-dementia groups (p < 0.001). Delirium score less than both Dementia and comorbid delirium-dementia groups (p < 0.001). Comorbid delirium-dementia group > Dementia (p < 0.001)

*Due to non-normal distribution, Kruskal-Wallis H Test was used to compare median MMSE across clinical groups and post hoc tests using Mann-Whitney U Tests were used to compare between groups. MMSE for NNCD > than for all other groups; MMSE for Dementia alone group was > than for
both dementia and delirium group and delirium groups; MMSE for delirium alone group was > MMSE for dementia and delirium group, all at P < 0.01.

**Due to non-normal distribution, Kruskal-Wallis H Test was used to compare median Short IQCODE across clinical groups and post hoc tests using Mann-Whitney U Tests were used to compare between groups. Short IQCODE scores were as follows: lower for the NNCD group in comparison to all other groups (P < 0.001); Dementia and delirium group > dementia alone group; dementia alone group > delirium alone group; dementia and delirium group > delirium alone group (all at P < 0.001)
Table 15: CDR scale scores for neurocognitive groups

<table>
<thead>
<tr>
<th>CDR</th>
<th>NNCD</th>
<th>Dementia</th>
<th>Comorbid delirium dementia</th>
<th>Delirium</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>0.5</td>
<td>26</td>
<td>13</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>19</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>38</td>
<td>47</td>
<td>29</td>
</tr>
</tbody>
</table>

The numbers and percentages for each neurocognitive group completing the three different components of the Lighthouse test are summarised in Table 16.

Table 16: Numbers from each neurocognitive group completing different components of the Lighthouse test

<table>
<thead>
<tr>
<th>Group</th>
<th>LHID</th>
<th>FA</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNCD</td>
<td>65/66 (98.5%)</td>
<td>66/66 (100%)</td>
<td>66/66 (100%)</td>
</tr>
<tr>
<td>Dementia</td>
<td>41/42 (97.6%)</td>
<td>41/42 (97.6%)</td>
<td>41/42 (97.6%)</td>
</tr>
<tr>
<td>Comorbid delirium-dementia</td>
<td>47/53 (88.7%)</td>
<td>48/53 (90.6%)</td>
<td>48/53 (90.6%)</td>
</tr>
<tr>
<td>Delirium</td>
<td>30/32 (93.8%)</td>
<td>32/32 (93.8%)</td>
<td>32/32 (93.8%)</td>
</tr>
</tbody>
</table>
**Performance on ‘gold standard’ tests**

Performance on the bedside assessments of cognitive function is described in Table 17. The ability to meaningfully engage with each of the tests varied from 84% (DSF) to 57% (Vig B test). In particular, the frequency of non-engagement was a considerable problem for the comorbid delirium-dementia group for all tests, ranging from 31 (DSF) to 77% (Vigilance B test).

In terms of performance scores, applying a correction of $p<0.01$ for multiple comparisons, the NNCD group scored significantly higher than the comorbid delirium-dementia group for all tests and higher than the delirium group for all tests except the DSB. The dementia group scored better than the comorbid delirium-dementia group for all tests and better than the delirium group for the MBT, Vigilance A and B tests, Global assessment of visuospatial ability and the IPT. The comorbid delirium-dementia and delirium without dementia groups differed in terms of performance on the WORLD backwards and DSB. The dementia group was distinguished from the NNCD group on the Vigilance B test and the CDT.

Using pass / fail cutoff scores, the NNCD pass rate was greater than those with comorbid delirium and dementia for all tests at $p<0.001$; the NNCD pass rate was greater than the delirium group at $p<0.001$ for MBT, Vigilance A, Vigilance B, Global VSP and IPT and at $p<0.01$ for WORLD, DSF, CDT, and at $p<0.05$ for DSB; the NNCD pass rate was greater than for the dementia group at $p<0.01$ for WORLD, $p<0.05$ for GVSP, CDT and IPT; Dem > comorbid del-demart $p<0.001$ for all tests except WORLD ($p<0.05$) and CDT ($p<0.01$); dementia > del at $p<0.01$ for MBT, Vigilance A, Vigilance B and GVSP; the frequency of test passing was higher for the delirium group compared with comorbid delirium and dementia at $p<0.001$ for the DSB and $p<0.05$ for DSF and IPT.

The comparisons of test performance amongst different motor subtypes in patients with delirium (see Table 18) indicated higher scores on the majority of tests for the no subtype group but otherwise there were few differences across the groups with hyperactive, hypoactive and mixed subtypes. Of note, the no subtype group had less severe delirium as evidenced by significantly lower DRS-R98 severity scale scores. In terms of percentage pass rate for the different tests there were minimal differences for the spatial span tests that were
not evident after correction for multiple comparisons. There were no significant differences across subtypes in terms of patient numbers that were able to engage meaningfully with each of the tests but this was notably high with approximately 50% of delirious subjects unable to engage with these tests.

Table 19 shows the accuracy of each of the tests (sensitivity, specificity, PPV, NPV) for presence of delirium. Table 20 shows a logistic regression where delirium / no delirium is the outcome variable and each of the cognitive tests were included as predictor variables. Three tests emerged as predictors of delirium; the MBT, Vigilance A test and the Global assessment of visuospatial ability. Table 21 shows the accuracy of various combinations of these three tests whereby combining any two of these three tests allowed for correct prediction of 77% of cases.
Table 17: Cognitive Test scores for total group and neurocognitive disorder groups (mean ± SD; % engagement, and pass/fail frequency).

<table>
<thead>
<tr>
<th></th>
<th>Total group (n=193)</th>
<th>No Neurocognitive Disorder (NNCD) (n = 66)</th>
<th>Delirium (n = 32)</th>
<th>Comorbid delirium-dementia (n = 53)</th>
<th>Dementia (n = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>World Backwards (n=192)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>2.2 ± 1.9</td>
<td>3.3 ± 1.9</td>
<td>2.1 ± 1.8</td>
<td>0.9 ± 1.4</td>
<td>2.2 ± 1.7</td>
</tr>
<tr>
<td>Unable to engage (%)</td>
<td>62 (32%)</td>
<td>11 (17%)</td>
<td>10 (31%)</td>
<td>33 (62%)</td>
<td>8 (19%)</td>
</tr>
<tr>
<td>Pass (%)</td>
<td>45 (23%)</td>
<td>31 (47%)</td>
<td>5 (16%)</td>
<td>2 (4%)</td>
<td>7 (17%)</td>
</tr>
<tr>
<td><strong>Months Backward Test (n=193)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>1.2 ± 1.1</td>
<td>1.9 ± 1.1</td>
<td>0.8 ± 0.8</td>
<td>0.5 ± 0.7</td>
<td>1.4 ± 1.0</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>63 (33%)</td>
<td>10 (15%)</td>
<td>14 (44%)</td>
<td>30 (57%)</td>
<td>9 (21%)</td>
</tr>
<tr>
<td>% pass</td>
<td>78 (40%)</td>
<td>44 (67%)</td>
<td>8 (25%)</td>
<td>5 (9%)</td>
<td>21 (50%)</td>
</tr>
<tr>
<td><strong>Digit Span Forwards (n=187)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>3.7 ± 2.4</td>
<td>4.8 ± 2.1</td>
<td>3.2 ± 2.4</td>
<td>2.1 ± 1.9</td>
<td>4.6 ± 2.1</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>29 (16%)</td>
<td>4 (6%)</td>
<td>7 (22%)</td>
<td>17 (31%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>% pass</td>
<td>72 (39%)</td>
<td>38 (58%)</td>
<td>10 (31%)</td>
<td>6 (11%)</td>
<td>18 (42%)</td>
</tr>
<tr>
<td><strong>Digit Span Backwards (n=187)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>2.1 ± 2.0</td>
<td>3.0 ± 1.9</td>
<td>2.1 ± 2.3</td>
<td>0.8 ± 1.2</td>
<td>2.4 ± 2.0</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>61 (33%)</td>
<td>10 (15%)</td>
<td>12 (38%)</td>
<td>31 (59%)</td>
<td>8 (19%)</td>
</tr>
<tr>
<td>% pass</td>
<td>71 (38%)</td>
<td>37 (56%)</td>
<td>12 (38%)</td>
<td>3 (6%)</td>
<td>19 (45%)</td>
</tr>
<tr>
<td><strong>Vigilance A Test (n=188)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>2.3 ± 1.8</td>
<td>3.3 ± 1.4</td>
<td>1.7 ± 1.7</td>
<td>1.0 ± 1.5</td>
<td>3.0 ± 1.6</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>54 (29%)</td>
<td>7 (11%)</td>
<td>12 (38%)</td>
<td>31 (58%)</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>% pass</td>
<td>93 (49%)</td>
<td>50 (75%)</td>
<td>9 (28%)</td>
<td>8 (15%)</td>
<td>26 (62%)</td>
</tr>
<tr>
<td><strong>Vigilance B Test (n=188)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>1.3 ± 1.4</td>
<td>2.3 ± 1.5</td>
<td>0.7 ± 0.9</td>
<td>0.7 ± 0.3</td>
<td>1.7 ± 1.4</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>80 (43%)</td>
<td>13 (20%)</td>
<td>16 (50%)</td>
<td>41 (77%)</td>
<td>10 (24%)</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>% pass</td>
<td>51 (27%)</td>
<td>34 (52%)</td>
<td>2 (6%)</td>
<td>2 (4%)</td>
<td>13 (31%)</td>
</tr>
</tbody>
</table>

**Global visuospatial function (n=189)**

<table>
<thead>
<tr>
<th>Mean ± SD</th>
<th>1.2 ± 1.2</th>
<th>2.0 ± 1.1</th>
<th>0.6 ± 0.7</th>
<th>0.5 ± 0.8</th>
<th>1.5 ± 1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>% unable to engage</td>
<td>69 (37%)</td>
<td>8 (12%)</td>
<td>16 (50%)</td>
<td>36 (68%)</td>
<td>9 (21%)</td>
</tr>
<tr>
<td>% pass</td>
<td>76 (40%)</td>
<td>45 (68%)</td>
<td>4 (13%)</td>
<td>8 (15%)</td>
<td>19 (45%)</td>
</tr>
</tbody>
</table>

**Interlocking Pentagons Test (n=185)**

<table>
<thead>
<tr>
<th>Mean ± SD</th>
<th>3.3 ± 2.0</th>
<th>4.4 ± 1.8</th>
<th>2.7 ± 1.7</th>
<th>1.9 ± 1.5</th>
<th>3.9 ± 1.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>% unable to engage</td>
<td>54 (29%)</td>
<td>6 (9%)</td>
<td>11 (34%)</td>
<td>33 (62%)</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>% pass</td>
<td>78 (42%)</td>
<td>43 (65%)</td>
<td>10 (31%)</td>
<td>9 (17%)</td>
<td>16 (38%)</td>
</tr>
</tbody>
</table>

**Clock Drawing Test (n=185)**

<table>
<thead>
<tr>
<th>Mean ± SD</th>
<th>4.5 ± 2.9</th>
<th>6.3 ± 2.4</th>
<th>3.7 ± 2.7</th>
<th>2.5 ± 2.4</th>
<th>5.0 ± 2.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>% unable to engage</td>
<td>55 (30%)</td>
<td>9 (14%)</td>
<td>11 (34%)</td>
<td>32 (60%)</td>
<td>3 (7%)</td>
</tr>
<tr>
<td>% pass</td>
<td>87 (47%)</td>
<td>47 (71%)</td>
<td>11 (34%)</td>
<td>8 (15%)</td>
<td>21 (50%)</td>
</tr>
</tbody>
</table>

Scores on tests compared with MWU: NNCD > Dd at p<0.001 for all tests; NNCD > Del at p<0.05 for DSB, p<0.01 for WORLD and DSF, and p<0.001 for all other tests; NNCD > Dem at p<0.05 for WORLD, MBT, Global VSP and p<0.01 for Vig B and CDT.

Dem > Dd at p<0.001 for all tests; Dem > Del at p<0.01 for MBT, Vig A, Vig B, Global Vsp and IPT; <0.05 for DSF. Dd< Del at p<0.01 for WORLD and DSB; p<0.05 for DSF, Vig B, CDT, IPT.

Using the pass / fail cut off scores, NNCD pass rate was > comorbid Del-dem for all tests at p<0.001; NNCD pass rate > delirium group at p<0.001 for MBT, Vig A, Vig B, Global VSP and IPT and p<0.01 for WORLD, DSF, CDT, and at p<0.05 for DSB; NNCD > dementia group at p<0.01 for WORLD, p<0.05 for GVSP, CDT and IPT; Dem > comorbid del-dem at p<0.001 for all tests except WORLD (p<0.05) and CDT (p<0.01); dementia > del at p<0.01 for MBT, Vig A, Vig B and GVSP; the frequency of test passing was higher for del vs comorbid at p<0.001 for the DSB and p<0.05 for DSf and IPT.
Table 18: DRS-R98 Severity and cognitive test scores for patients with delirium (n=85) according to clinical subtype as per the Delirium Motor Subtype Scale (mean ± SD; % engagement, and pass/fail frequency).

<table>
<thead>
<tr>
<th></th>
<th>Hyperactive Subtype (n=30)</th>
<th>Mixed Subtype (n=15)</th>
<th>Hypoactive Subtype (n=23)</th>
<th>No Motor Subtype (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRS-R98 Severity Scale Score</td>
<td>22.6 ± 4.5</td>
<td>22.7 ± 4.4</td>
<td>21.1 ± 4.6</td>
<td>16.7 ± 3.1</td>
</tr>
<tr>
<td>World Backwards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>1.7 ± 1.9</td>
<td>1.6 ± 1.8</td>
<td>1.0 ± 1.4</td>
<td>1.1 ± 1.2</td>
</tr>
<tr>
<td>Unable to engage (%)</td>
<td>15 (50%)</td>
<td>7 (47%)</td>
<td>13 (57%)</td>
<td>8 (47%)</td>
</tr>
<tr>
<td>Pass (%)</td>
<td>4 (13%)</td>
<td>2 (13%)</td>
<td>1 (4%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Months Backward Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>0.7 ± 0.8</td>
<td>0.6±0.6</td>
<td>0.7 ± 0.7</td>
<td>0.5 ± 0.7</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>17 (57%)</td>
<td>7 (47%)</td>
<td>9 (39%)</td>
<td>11 (65%)</td>
</tr>
<tr>
<td>% pass</td>
<td>7 (23%)</td>
<td>1 (7%)</td>
<td>3 (13%)</td>
<td>2 (12%)</td>
</tr>
<tr>
<td>Digit Span Forwards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>3.2 ± 2.3</td>
<td>1.6 ± 1.9</td>
<td>2.1± 2.2</td>
<td>2.8± 1.9</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>7 (23%)</td>
<td>6 (40%)</td>
<td>8 (35%)</td>
<td>3 (18%)</td>
</tr>
<tr>
<td>% pass</td>
<td>10 (33%)</td>
<td>1 (7%)</td>
<td>3 (13%)</td>
<td>2 (12%)</td>
</tr>
<tr>
<td>Digit Span Backwards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>1.5 ± 1.9</td>
<td>1.0±1.8</td>
<td>1.0±1.6</td>
<td>1.4± 1.6</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>13 (43%)</td>
<td>11 (73%)</td>
<td>11 (48%)</td>
<td>8 (47%)</td>
</tr>
<tr>
<td>% pass</td>
<td>7 (23%)</td>
<td>3 (20%)</td>
<td>1 (4%)</td>
<td>4 (24%)</td>
</tr>
<tr>
<td>Vigilance A Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>1.5 ± 1.7</td>
<td>1.1±1.5</td>
<td>0.9± 1.3</td>
<td>1.6± 1.8</td>
</tr>
<tr>
<td>% unable to engage</td>
<td>14 (46%)</td>
<td>8 (53%)</td>
<td>12 (52%)</td>
<td>9 (53%)</td>
</tr>
<tr>
<td>Test</td>
<td>% Pass</td>
<td>% Unable to Engage</td>
<td>% Pass</td>
<td>% Unable to Engage</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Vigilance B Test</strong></td>
<td>8 (26%)</td>
<td>19 (63%)</td>
<td>2 (7%)</td>
<td>11 (37%)</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>0.6 ± 0.9</td>
<td>0.7± 1.1</td>
<td>0.3± 0.6</td>
<td>0.4± 0.7</td>
</tr>
<tr>
<td><strong>Global visuospatial function</strong></td>
<td>2 (14%)</td>
<td>9 (60%)</td>
<td>2 (13%)</td>
<td>10 (66%)</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>0.7 ± 1.0</td>
<td>0.6± 1.0</td>
<td>0.6± 0.8</td>
<td>0.6± 0.8</td>
</tr>
<tr>
<td><strong>Clock Drawing Test</strong></td>
<td>3 (10%)</td>
<td>14 (46%)</td>
<td>3 (20%)</td>
<td>9 (30%)</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>3.5±2.8</td>
<td>2.2±2.1</td>
<td>2.9±2.4</td>
<td>2.5±2.7</td>
</tr>
<tr>
<td><strong>Interlocking Pentagons Test</strong></td>
<td>3 (13%)</td>
<td>11 (37%)</td>
<td>3 (18%)</td>
<td>4 (26%)</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>2.5±1.5</td>
<td>1.8±1.5</td>
<td>2.1±1.8</td>
<td>2.0±1.8</td>
</tr>
</tbody>
</table>

No subtype < other three groups: p<0.001 for hyperactive and mixed subtype, p=0.01 for hypoactive subtype.

No > Hyper at p<0.001 for MBT, Vig B, Global VSP; p<0.01 for DSB, Vig A, CDT, IPT.No> Mixed at p<0.001 for DSF, Vig A, IPT, CDT; p<0.01 for MBT, DSB, Global VSP; p<0.05 for Vig B.

No > Hypo at p<0.001 for all tests except p<0.01 for WORLD and Global VSP

Hypo and Hyper NS for all tests; hypo and mixed NS all tests; Hyper > mixed at p<0.01 for DSF and p<0.05 for IPT

For pass/fail; hyper > mix for DSF at p=0.04; per >po for DSB at p=0.04; rest NS

Coverage: mix < other 3 subtype categories at p=0.08.
Table 19: Accuracy of individual tests of attention, vigilance and visuospatial function for the detection of DSM-IV delirium

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity (95% C.I.)</th>
<th>Specificity (95% C.I.)</th>
<th>PPV (95% C.I.)</th>
<th>NPV (95% C.I.)</th>
<th>Overall Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>WORLD</td>
<td>91.7 (83.8-96.6)</td>
<td>35.2 (26.2-45.0)</td>
<td>52.7 (44.3-60.9)</td>
<td>84.4 (70.5-93.5)</td>
<td>60%</td>
</tr>
<tr>
<td>MBT</td>
<td>85.9 (76.6-92.5)</td>
<td>61.1 (51.3-70.3)</td>
<td>63.5 (54.0-72.3)</td>
<td>84.6 (74.7-91.2)</td>
<td>72%</td>
</tr>
<tr>
<td>DSF</td>
<td>80.9 (70.9-88.7)</td>
<td>53.9 (43.8-63.8)</td>
<td>59.1 (49.6-68.2)</td>
<td>77.5 (66.0-86.5)</td>
<td>64%</td>
</tr>
<tr>
<td>DSB</td>
<td>82.1 (72.3-89.7)</td>
<td>54.4 (44.3-64.2)</td>
<td>59.5 (49.9-68.5)</td>
<td>78.9 (67.6-87.7)</td>
<td>67%</td>
</tr>
<tr>
<td>Vigilance A</td>
<td>79.8 (69.6-87.8)</td>
<td>73.1 (63.5-81.3)</td>
<td>70.5 (60.3-79.4)</td>
<td>81.7 (72.6-88.9)</td>
<td>76%</td>
</tr>
<tr>
<td>Vigilance B</td>
<td>95.2 (88.3-98.7)</td>
<td>45.2 (35.4-55.3)</td>
<td>58.4 (49.7-66.8)</td>
<td>92.2 (81.1-97.8)</td>
<td>68%</td>
</tr>
<tr>
<td>Global VSP</td>
<td>85.5 (76.1-92.3)</td>
<td>60.4 (50.4-69.8)</td>
<td>62.8 (53.2-71.7)</td>
<td>84.2 (74.0-91.6)</td>
<td>71%</td>
</tr>
<tr>
<td>CDT</td>
<td>67.4 (56.7-77.0)</td>
<td>56.9 (46.7-66.6)</td>
<td>57.7 (47.6-67.3)</td>
<td>66.7 (55.8-76.4)</td>
<td>65%</td>
</tr>
<tr>
<td>IPT</td>
<td>75.3 (64.5-84.2)</td>
<td>65.7 (55.6-74.8)</td>
<td>63.5 (53.1-73.1)</td>
<td>77.0 (66.8-85.4)</td>
<td>69%</td>
</tr>
</tbody>
</table>
Table 20: Logistic Regression for predictors of DSM-IV delirium amongst tests of attention, vigilance and visuospatial ability.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Wald</th>
<th>$R^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigilance A</td>
<td>-1.49</td>
<td>15.72</td>
<td>0.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Global VSP</td>
<td>-1.62</td>
<td>15.34</td>
<td>0.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MBT</td>
<td>-1.72</td>
<td>11.76</td>
<td>0.51</td>
<td>0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>1.65</td>
<td>29.14</td>
<td></td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 21: Accuracy of combined tests of attention, vigilance and visuospatial function for the detection of DSM-IV delirium

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (95% C.I.)</th>
<th>Specificity (95% C.I.)</th>
<th>PPV (95% C.I.)</th>
<th>NPV (95% C.I.)</th>
<th>Overall Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBT and Vigilance A</td>
<td>94.1 (86.8-98.1)</td>
<td>50.0 (40.1-59.9)</td>
<td>60.2 (51.3-68.5)</td>
<td>91.4 (81.0-97.1)</td>
<td>70%</td>
</tr>
<tr>
<td>Vigilance A and Global VSP</td>
<td>95.2 (88.3-98.7)</td>
<td>47.1 (37.3-57.2)</td>
<td>59.3 (50.5-67.6)</td>
<td>92.5 (81.8-97.9)</td>
<td>69%</td>
</tr>
<tr>
<td>MBT and Global VSP</td>
<td>96.5 (90.0-99.3)</td>
<td>39.8 (30.5-49.7)</td>
<td>55.8 (47.4-64.0)</td>
<td>93.5 (82.1-98.6)</td>
<td>65%</td>
</tr>
<tr>
<td>MBT plus Vigilance A plus Global VSP</td>
<td>98.8 (93.6-99.9)</td>
<td>34.0 (25.0-43.8)</td>
<td>54.6 (46.3-62.6)</td>
<td>97.3 (85.8-99.9)</td>
<td>63%</td>
</tr>
<tr>
<td>Any Two of MBT, Vigilence A or Global VSP</td>
<td>88.3 (79.4-94.2)</td>
<td>68.2 (58.5-76.9)</td>
<td>68.8 (59.2-77.34)</td>
<td>88.0 (79.0-94.1)</td>
<td>77%</td>
</tr>
</tbody>
</table>
Performance on the Lighthouse app tests

Lighthouse identification (LH-ID)

Overall, 60/183 (33%) correctly identified the Lighthouse.

The specific responses to identifying the Lighthouse were assessed in a random sample of 100 patients all of whom were assessed by the same assessor (H.O’C).

Of these 100 participants, 30 did not give a response, 35 gave the correct response and 35 gave a response that was incorrect. Of the 35 correct responses, 3 participants gave a description of the lighthouse but were unable to give the actual name (e.g. ‘signals out for ships at sea’). For the incorrect responses, answers varied widely and included ‘a cone’, ‘chimney’, ‘traffic light’, ‘a village pump’ and ‘a tower’.

Table 22: LH-ID cross-tabulation for the 4 clinical groups

<table>
<thead>
<tr>
<th>Neurocognitive groups</th>
<th>NNCD</th>
<th>Dementia alone</th>
<th>Dementia + Delirium</th>
<th>Delirium alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number passing the LH-ID test/total in that neurocognitive group (%)</td>
<td>28/65 (43%)</td>
<td>12/41 (29.3%)</td>
<td>7/47 (14.9%)</td>
<td>13/30 (43.3%)</td>
</tr>
</tbody>
</table>

The four neurocognitive groups significantly differed in terms of correct response rate ($p = 0.008$). A comparison of those who identified the Lighthouse with those who did not found that age was greater in those who failed ($p = 0.01$) and MMSE was greater in those who passed ($p = 0.001$).
Table 23: Accuracy of LH-ID test for identification of delirium

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>74.0%</td>
<td>62.8-83.4</td>
</tr>
<tr>
<td>Specificity</td>
<td>37.7%</td>
<td>28.5-47.7</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>46.3%</td>
<td>37.3-55.6</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>66.7%</td>
<td>53.3-78.3</td>
</tr>
</tbody>
</table>

The LH-ID test when used alone had sensitivity of 74% (95% CI 62.8-83.4%) and specificity of 37.7% (95% CI 28.5-47.7%) for detection of delirium. Positive predictive value of the test was 46.3% (95% CI 37.3-55.6%) and negative predictive value was 66.7% (53.3-78.3%).
Focussed Attention (FA) component of Lighthouse test

The second part of the Lighthouse test consists of a series of three tests examining the subject’s capacity to focus attention by describing the number of times the lighthouse flashes (3 sequences of 4 flashes, 3 flashes and 5 flashes). Performance was rated from 0-3 depending on the number of correct answers. Focussed Attention Pass/Fail analysis was conducted, whereby subjects correctly completing all 3 FA tests were judged to have passed while subjects scoring any less than 3 (i.e. 0, 1 or 2) were judged to have failed.

Overall, 85/187 (45%) correctly completed the FA component of the Lighthouse test. Table 24 shows the relationship between neurocognitive diagnosis and performance on the FA test.

Table 24: FA test performance by neurocognitive group

<table>
<thead>
<tr>
<th></th>
<th>NNCD</th>
<th>Dementia alone</th>
<th>Dementia + Delirium</th>
<th>Delirium alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed/total in group</td>
<td>39/66 (59.1%)</td>
<td>22/41 (53.7%)</td>
<td>9/48 (18.8%)</td>
<td>15/32 (46.9%)</td>
</tr>
</tbody>
</table>

For those with delirium, 24/80 (30%) passed the FA test. For those without delirium, 61/105 (58.1%) passed the FA test.

Those who passed the FA test had higher mean MMSE scores than those who failed (p < 0.001)
Table 25: FA test performance for identification for delirium

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>70.0%</td>
<td>58.7-79.7</td>
</tr>
<tr>
<td>Specificity</td>
<td>57.0%</td>
<td>47.1-66.5</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>54.9%</td>
<td>44.7-64.8</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>71.8%</td>
<td>70.0-81.0</td>
</tr>
</tbody>
</table>

The FA test when used alone had sensitivity of 70% (CI 58.7-79.7) and specificity of 57% (CI 47.1-66.5%). Positive predictive value of the FA test was 54.9% (44.7-64.8%) and negative predictive value was 71.8% (70-81%).
Sustained Attention (SA) component of Lighthouse test

A SA Pass/Fail analysis was conducted, whereby subjects scoring 3/3 on the SA test were judged to have passed and subjects scoring any less than 3 (i.e. 0, 1 or 2) were judged to have failed.

Overall, 66/187 (35.3%) completed the 3 parts of the SA test without error.

Table 26: SA test performance by neurocognitive group

<table>
<thead>
<tr>
<th></th>
<th>NNCD</th>
<th>Dementia alone</th>
<th>Dementia + Delirium</th>
<th>Delirium alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Passed/total group</td>
<td>36/66 (54.5%)</td>
<td>20/41 (48.8%)</td>
<td>4/48 (8.3%)</td>
<td>6/32 (18.8%)</td>
</tr>
</tbody>
</table>

A significantly higher number of those with delirium, 10/80 (12.5%) passed the SA test, compared with those without delirium, 56/107 (52.3%) (p < 0.001)

MMSE scores for those who passed were higher than for those who failed (p < 0.001)
Table 27: SA test performance for identification of delirium

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>87.5%</td>
<td>78.2-93.8</td>
</tr>
<tr>
<td>Specificity</td>
<td>52.3%</td>
<td>42.5-62.1</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>57.9%</td>
<td>48.5-66.8</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>84.9%</td>
<td>74.0-92.5</td>
</tr>
</tbody>
</table>

The FA test when used alone had sensitivity of 87.5% (CI 78.2-93.8%) and specificity of 52.3% (CI 42.5-62.1%). Positive predictive value of the FA test was 57.9% (48.5-66.8%) and negative predictive value was 84.9% (74-92.5%).
The three components of the Lighthouse test (LH-ID, FA and SA) were then analysed in a range of potential combinations.

**Table 28: Accuracy of different combinations of the individual Lighthouse tests**

<table>
<thead>
<tr>
<th>Test</th>
<th>Controls</th>
<th>Dementia alone</th>
<th>Dementia and Delirium</th>
<th>Delirium alone</th>
<th>Sensitivity for delirium</th>
<th>Specificity for delirium</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID + FA</td>
<td>21/66 (31.8%)</td>
<td>9/41 (22%)</td>
<td>2/49 (4.1%)</td>
<td>4/31 (12.9%)</td>
<td>86.3%</td>
<td>28%</td>
</tr>
<tr>
<td>Passed/total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID + SA</td>
<td>18/65 (27.7%)</td>
<td>10/41 (24.4%)</td>
<td>2/49 (4.1%)</td>
<td>4/31 (12.9%)</td>
<td>92.5%</td>
<td>26.4%</td>
</tr>
<tr>
<td>Passed/total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA + SA</td>
<td>32/66 (48.5%)</td>
<td>16/41 (39%)</td>
<td>2/48 (4.2%)</td>
<td>5/32 (15.6%)</td>
<td>91.3%</td>
<td>44.9%</td>
</tr>
<tr>
<td>Passed/total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID + FA + SA</td>
<td>17/66 (25.8%)</td>
<td>8/41 (19.5%)</td>
<td>1/49 (2%)</td>
<td>3/31 (9.7%)</td>
<td>95%</td>
<td>23.4%</td>
</tr>
</tbody>
</table>

Therefore, when all 3 components of the Lighthouse test were combined, 4/80 subjects with delirium passed. These subjects can be categorised as ‘false negatives’ (FN) on the ID, FA and SA tests combined tests and termed ‘ID-FA-SA-FN’ (N = 4). There were no significant differences in age, gender or MMSE score between the IDFASAFN group and others.
Table 29: Overall test accuracy for individual and combined Lighthouse tests

<table>
<thead>
<tr>
<th>Test</th>
<th>True positives</th>
<th>True negatives</th>
<th>Total</th>
<th>OTA %</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH-ID alone</td>
<td>57</td>
<td>40</td>
<td>183</td>
<td>53</td>
</tr>
<tr>
<td>FA test alone</td>
<td>56</td>
<td>61</td>
<td>187</td>
<td>63</td>
</tr>
<tr>
<td>SA test alone</td>
<td>70</td>
<td>56</td>
<td>187</td>
<td>67</td>
</tr>
<tr>
<td>ID + FA</td>
<td>74</td>
<td>30</td>
<td>187</td>
<td>56</td>
</tr>
<tr>
<td>ID + SA</td>
<td>74</td>
<td>28</td>
<td>186</td>
<td>55</td>
</tr>
<tr>
<td>FA + SA</td>
<td>73</td>
<td>48</td>
<td>187</td>
<td>65</td>
</tr>
<tr>
<td>ID + FA + SA</td>
<td>76</td>
<td>25</td>
<td>187</td>
<td>54</td>
</tr>
</tbody>
</table>

Overall test accuracy (calculated as Lighthouse test true positives added to true negatives and divided by total tested and given as percentage) was highest for the SA test alone (67%) and lowest for the LH-ID test alone (53%). All results are summarised above in Table 29.
Those who failed all three components of the Lighthouse test were subsequently identified. These were subjects who failed LH-ID, FA and SA tests and are thus referred to as LHIDFASA Failers. The number of LHDFASA Failers in each neurocognitive group are summarised below.

The percentages of each neurocognitive group who failed all three components of the Lighthouse test were as follows: 25.8% of controls; 34.1% of the dementia alone group; 38.7% of the delirium alone group and 67.4% of the comorbid dementia and delirium group.
Table 30: LHIDFASA Failers compared to others (i.e. those who passed at least one component of the Lighthouse test – LHID, FA or SA)

<table>
<thead>
<tr>
<th></th>
<th>LHIDFASA Failers</th>
<th>Others</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (%)</td>
<td>51</td>
<td>49</td>
<td>NS</td>
</tr>
<tr>
<td>Age</td>
<td>80.2</td>
<td>79.7</td>
<td>NS</td>
</tr>
<tr>
<td>Total number of medications</td>
<td>9.8</td>
<td>10.1</td>
<td>NS</td>
</tr>
<tr>
<td>Number of psychotropics</td>
<td>1.8</td>
<td>1.5</td>
<td>NS</td>
</tr>
<tr>
<td>MMSE</td>
<td>15.0</td>
<td>21.8</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>DRS-R98 total</td>
<td>21.0</td>
<td>13.3</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Short IQCODE</td>
<td>3.7</td>
<td>3.3</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>CDR global score</td>
<td>1.2</td>
<td>0.6</td>
<td>p &lt; 0.001</td>
</tr>
</tbody>
</table>

LHIDFASA Failers: those participants who failed all three components of the Lighthouse app test

When those who failed all three components of the Lighthouse test (LHIDFASA Failers) were compared to others (i.e. those who passed at least one component of the Lighthouse test), there were no significant differences in gender, age, total number of medications or total number of psychotropic medications. However, the LHIDFASA Failers had lower MMSE scores, higher DRS-R98 Total scores, higher Short IQCODE scores and higher global CDR scores.
Concordance measures

In order to assess the relationship between ‘gold standard tests’ (i.e. DRS-R98 and CTD) and the components of the Lighthouse test, DRS-R98 scores on items 9-13 inclusive were divided into ‘normal’ (</=2) and ‘pathological’ (>/=3). Likewise CTD scores for Orientation, Attention, Memory and Vigilance items were divided into ‘normal’ (</=2) and ‘pathological’ (>/=4). The CTD Comprehension item was divided into ‘normal’ (</=3) and ‘pathological’ (>/=4).

Kappa values were then calculated, relating the three components of the Lighthouse test to ‘normal’ scores on DRSR98 items 9-13 inclusive and all components of the CTD test.

Table 31: Kappa values for DRS, CTD and Lighthouse test items

<table>
<thead>
<tr>
<th></th>
<th>LHID pass/fail</th>
<th>FA test pass/fail</th>
<th>SA test pass/fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRSR98 item 9 (Orientation)</td>
<td>-0.072</td>
<td>-0.036</td>
<td>-0.103</td>
</tr>
<tr>
<td>DRSR98 item 10 (Attention)</td>
<td>-0.058</td>
<td>-0.141</td>
<td>-0.252</td>
</tr>
<tr>
<td>DRSR98 item 11 (Short term memory)</td>
<td>-0.236</td>
<td>-0.176</td>
<td>-0.288</td>
</tr>
<tr>
<td>DRSR98 item 12 (Long term memory)</td>
<td>-0.170</td>
<td>-0.137</td>
<td>-0.252</td>
</tr>
<tr>
<td>DRSR98 item 13 (Visuospatial)</td>
<td>-0.148</td>
<td>-0.283</td>
<td>-0.363</td>
</tr>
<tr>
<td>CTD Orientation</td>
<td>0.197</td>
<td>0.256</td>
<td>0.234</td>
</tr>
<tr>
<td>CTD Attention</td>
<td>0.248</td>
<td>0.372</td>
<td>0.361</td>
</tr>
<tr>
<td>CTD Memory</td>
<td>0.220</td>
<td>0.365</td>
<td>0.339</td>
</tr>
<tr>
<td>CTD Comp</td>
<td>0.125</td>
<td>0.233</td>
<td>0.231</td>
</tr>
<tr>
<td>CTD Vigilance</td>
<td>0.298</td>
<td>0.330</td>
<td>0.406</td>
</tr>
</tbody>
</table>
Table 32: Concordance for Lighthouse tests and other tests of cognition (kappa values)

<table>
<thead>
<tr>
<th></th>
<th>LHID</th>
<th>FA</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBT Pass</td>
<td>0.149</td>
<td>0.304</td>
<td>0.334</td>
</tr>
<tr>
<td>WORLD Pass</td>
<td>0.148</td>
<td>0.169</td>
<td>0.205</td>
</tr>
<tr>
<td>DSF Pass</td>
<td>0.193</td>
<td>0.390</td>
<td>0.359</td>
</tr>
<tr>
<td>DSB Pass</td>
<td>0.278</td>
<td>0.336</td>
<td>0.314</td>
</tr>
<tr>
<td>Vig A Pass</td>
<td>0.315</td>
<td>0.418</td>
<td>0.407</td>
</tr>
<tr>
<td>Vig B Pass</td>
<td>0.230</td>
<td>0.261</td>
<td>0.321</td>
</tr>
<tr>
<td>GVS Pass</td>
<td>0.136</td>
<td>0.260</td>
<td>0.346</td>
</tr>
<tr>
<td>CDT Pass</td>
<td>0.308</td>
<td>0.345</td>
<td>0.351</td>
</tr>
<tr>
<td>IPT (4 or more a Pass)</td>
<td>0.276</td>
<td>0.521</td>
<td>0.438</td>
</tr>
<tr>
<td>IPT (6 or less a Fail)</td>
<td>0.203</td>
<td>0.312</td>
<td>0.339</td>
</tr>
</tbody>
</table>

The findings in Table 33 suggest that the components of the Lighthouse test have substantial overlap with standard tests of attention. The SA test is generally more highly concordant with standard tests, the LHID test appears to be least concordant and the FA test concordance measures generally lie between those of the other two Lighthouse tests. Specifically, the SA test appears more strongly linked to vigilance and more complex/prolonged tests of attention. In comparison with the standard tests, the Lighthouse tests have the advantages of more consistent delivery and interpretation with defined cut-off scores. Furthermore, as outlined in Tables 16 and 17, the Lighthouse test has greater clinical coverage with all clinical groups than standard tests of attention.
Study hypotheses

As outlined earlier in Chapter 1, the study hypotheses were as follows:

1. The clinical utility of the novel Lighthouse tests in terms of accuracy for distinguishing delirium from non-delirium in patients with mixed neuropsychiatric presentations will compare favourably with commonly used ‘gold-standard’ cognitive screening instruments.

2. The coverage, i.e. the proportion of patients with various neurocognitive disorders who are able to engage meaningfully (i.e. are rateable) with the Lighthouse test will be at least equivalent with standard tests.

Regarding the first hypothesis, this study has demonstrated that the utility of the Lighthouse tests, individually and in combination, compare favourably with commonly used ‘gold standard’ cognitive screening instruments. Detailed results are displayed in Table 19 (for ‘gold standard’ tests) and in Tables 28 and 29 (for the Lighthouse tests). Overall Test Accuracy scores for the ‘gold standard’ tests (as outlined in Table 19 above) ranged from 60% for WORLD up to 76% for the Vigilance A test. Of the ‘gold standard’ tests, the Vigilance B test had the highest sensitivity for detection of delirium and the Vigilance A test had highest specificity (73.1%). Overall Test Accuracy for the Lighthouse tests ranged from 53% (LH-ID alone) to 67% (for the SA test alone). When all three components of the Lighthouse test were combined, 95% sensitivity in detection of delirium was demonstrated. Therefore, sensitivity of the Lighthouse test was superior to that of the gold standard tests and Overall Test Accuracy score for the Lighthouse test was similar to gold standard tests. Overall, specificity of the Lighthouse test was poor, with the optimal combination of tests (FA and SA tests combined) having a specificity of 44.9%. The clinical importance of early detection and rapid intervention in delirium means that the trade-off of poor specificity for excellent sensitivity is reasonable, particularly when it is envisaged that use of the Lighthouse app will be the first ‘screening’ phase of a process, with more detailed assessments for those screening ‘positive’.
Regarding the second hypothesis, ability to engage with testing was superior in the Lighthouse in comparison to ‘gold standard’ tests. Table 16 demonstrates engagement levels of 88.7-100% for the different neurocognitive groups with the different components of the Lighthouse test. In contrast (as outlined in Table 17), 7-68% of the different neurocognitive groups were unable to engage with ‘gold standard’ tests.
Chapter 5: Discussion

Delirium is a common, severe and life threatening neuropsychiatric disorder. The focus of this study has been on a general hospital population, where delirium affects at least one in five inpatients (Ryan et al., 2013). There is also a dearth of effective and objective bedside tests of cognition, especially computerised tests, as evidenced by our review of this area.

Research over the past decade by the Cognitive Impairment Research Group (CIRG) has helped establish a number of key findings on the phenomenology of delirium. Attentional deficits are the key cognitive findings in delirium across different patient populations and clinical settings. Furthermore, such attentional deficits persist throughout the course of an episode of delirium. Therefore, the development of objective, brief and user friendly bedside tests of attention will be vital in improving the levels of delirium detection and diagnosis.

The systematic review of the literature on bedside tests of attention identified thirteen different tests of attention that were used across a variety of healthcare settings in the detection and diagnosis of delirium. Study sizes and methodologies varied widely and there was a dearth of information on the use of computerised tests of cognition.
‘Gold standard’ tests

We compared performance on bedside tests of cognition in elderly medical inpatients with a variety of neurocognitive presentations, including a cognitively-intact group. Participants were carefully diagnosed using a full neuropsychiatric assessment that included well-validated instruments – the DRS-R98 and Short IQCODE. We found that patients with active delirium (both with and without comorbid dementia) could be distinguished from patients with dementia-alone in respect of performance on a variety of simple tests of attention, vigilance and visuospatial abilities and the combination of these tests allowed for accurate distinction of different neurocognitive diagnostic groups. However, the ability to engage with testing was limited for many patients, especially those with comorbid delirium and dementia.

The findings in respect of differences in cognitive profile between delirium and dementia extend previous work where performance on tests of attention and vigilance distinguished both delirium and comorbid delirium-dementia from dementia alone (Meagher et al., 2010). Attentional disturbances in delirium are in respect of the ability to direct, focus, sustain and shift attention. Vigilance is a term that is equated with the ability to sustain attention to a task and thus is often referred to as ‘vigilant’ attention (Oken et al., 2006). In addition, this study of elderly medical inpatients found that visuospatial ability also distinguished delirium groups from dementia. Similarly, Brown et al (2011) compared performance among patients with delirium, dementia and unimpaired cognition on a series of tests of sustained visual attention and found that delirious patients could be distinguished across a range of tests, while performance among the patients with dementia was relatively preserved and equivalent to the unimpaired controls. These findings highlight how efforts to improve detection of delirium (e.g. developing screening tools) can be enhanced by emphasizing attention, vigilance and visuospatial ability as key cognitive functions for testing.

The choice of optimal testing methods for everyday clinical practice involves careful consideration of overall accuracy, relative importance of sensitivity versus specificity, coverage of the population in terms of being able to engage with the testing procedures, and ease of delivery and interpretation. Our findings highlighted the importance of applying (any) routine and formal testing of cognitive status in improving delirium recognition. The
consistency of test performance was such that all of the tests were potentially useful for assisting in the diagnosis of neurocognitive disorders and all of the tests had a sensitivity for delirium of greater than 75% (except for the clock drawing test). In particular, three tests emerged as most useful for delirium detection – the MBT, Vigilance A and global VSP. The combination of all three tests allowed for very high sensitivity while the comparisons of combinations of any two of these tests indicated pairing the MBT with the Global VSP had the highest sensitivity for delirium.

Ultimately, the choice of test also involves a careful consideration of the skillset of those performing the cognitive testing allied to the needs of individual patients. For patients (e.g. who are intubated) who lack verbal capacity the Vigilance A test is suitable, while for others the MBT is easy, non-threatening to patients (i.e. does not require calculation) and has established norms for delivery and interpretation (Meagher et al., 2015). Moreover, the MBT is also available in a computerised version that can enhance test reliability (Donoghue et al., 2016). In addition, the global assessment of visuospatial function which is based upon simple questions relating to the immediate environment allows for a verbal emphasis which, unlike the CDT and IPT, is thus suited to subjects with fine motor or other impairments.

Our findings highlight the challenge in assessing cognitive function amongst highly morbid populations where there is wide variability in capacity to engage with testing procedures. We found limited difference in test performances according to motor subtype of delirium which is in keeping with previous work highlighting how clinical subtypes differ in terms of motor behaviour (and general neuropsychiatric burden) but not in cognition (Meagher, 2009, Meagher et al., 2012, Leonard et al., 2011). In addition, we found similar coverage across motor subtypes in terms of the ability to engage with tests. Although these findings reflected contrasting difficulties in terms of drowsiness compared to agitation, a common issue was that patients were unable to engage with testing in a way that allowed for rating above a zero score. It is relevant that the tests included herein were largely developed as elements of test batteries designed to assess cognition in subjects at a higher level of functioning and are thus not ideally suited to assessing patients with severe neurocognitive disorders such as delirium. These bottoming-out limitations impact upon the capacity of cognitive tests to allow for detailed assessment of severity and monitoring progress over time in patients with more severe delirium. Guidance around application of DSM-5 criteria
suggests that the inability to engage with tests of attention should be equated with severe inattention (European Delirium Association, 2014) but alternate approaches are needed in order to more adequately capture cognitive performance in this group. As a consequence, new approaches to assessing arousal that are specifically developed for use in delirious patients and that relate to inattention have recently emerged and can allow for more systematic documentation of the ability to engage with testing of cognition (Tieges et al., 2013b, Neerland et al., 2014). This increasing focus upon tools that are specifically designed for the assessment of cognitive function in delirium can also allow for better capture of patients at the more severe end of the performance spectrum.
Lighthouse app tests

Computerised tests offer advantages over traditional pen and paper tests, for reasons of objectivity and reproducibility. Furthermore, smartphone applications are used widely by healthcare professionals in their clinical work on a daily basis (Weir et al., 2014). These factors led to the development of the Lighthouse test application by the CIRG and the Department of CSIS at the University of Limerick, building upon several years of CIRG delirium research, particularly in studies relating to assessment of attention in delirium (Meagher et al., 2007, Jabbar et al., 2011, Leonard et al., 2015, Adamis et al., 2015).

The Lighthouse test application provides a brief (less than three minutes) and portable bedside test that is acceptable and user friendly for assessor and patient alike. Training required for use of the Lighthouse app test is minimal, in view of its simplicity and brevity, and should require no more than five minutes of explanation and practice. Furthermore, healthcare professionals of all backgrounds and levels of experience/skill-sets are likely to be able to use the test.

The Lighthouse test involves three main components, examining recognition, focussed attention and sustained attention. In this study, sensitivity of up to 95% was achieved for detection of delirium in what was a heterogenous real-world clinical population. Specificity was low for different Lighthouse test components, individually and in different combinations. Level of engagement of the different neurocognitive groups with the components of the Lighthouse test were superior to that seen with ‘gold standard’ tests.

While objective and computerised tests of attention are likely to have a key role in the development of delirium detection programmes, it should also be outlined that key research in the area of bedside testing of attention in delirium has highlighted the importance and indeed superiority of subjective assessment by experienced clinicians. O’Keeffe and Gosney (1997a) compared a variety of approaches to assessment of attention in older medical inpatients and found that while objective tests of attention could distinguish delirium from non-delirium, a clinical global assessment of attention was the most distinguishing approach to delirium detection. Likewise, when O’Regan et al (2014) examined the utility of a number of bedside tests of attention in an older inpatient population, they concluded that the most
precise screening method involved a combination of clinical assessment for subjective/objective confusion along with the MBT test.

Adamis et al (2015) also examined a number of bedside tests of attention in an elderly inpatient population and again concluded that clinical subjective rating of attention in conjunction with the MBT had the best discriminatory ability to identify delirium. The authors also reported a high negative predictive value for the bedside tests (i.e. good accuracy in excluding delirium) but low positive predictive value (meaning they are not diagnostic for delirium).

The findings of the latter three studies (O'Keeffe and Gosney, 1997b, O'Regan et al., 2014b, Adamis et al., 2015) are also reflected in the results of this research whereby the Lighthouse app has demonstrated excellent sensitivity but low specificity. These findings are important in informing how we shape delirium screening and diagnostic programmes in the future.

The evidence on objective tests, including the Lighthouse test, suggests that they are useful in excluding non-delirium cases. However, more detailed assessment by trained clinicians is subsequently required in order to confirm delirium status in those who have screened positive initially with objective tests such as the Lighthouse test. This two-phase approach, along with daily screening for all those at risk, is advocated by the National Institute of Health and Care Excellence (NICE) (Young et al., 2010).

Future modifications to the Lighthouse app will also be informed by the findings from this research and will involve improvements in graphics to make the Lighthouse more clearly identifiable as such (see Figures 3 and 4). Adjustments will also be made to allow for scores to be recorded immediately on the device.

Specificity of the Lighthouse test is likely to be improved by adding questions about key contextual clinical factors, especially acuity of onset of cognitive problems, to form a more comprehensive delirium test battery. It will be possible to add such questions to the Lighthouse battery so that they can be applied algorithmically by clinicians of all backgrounds and expertise, thus reducing reliance on highly trained experts.

Ultimately, a brief portable bedside battery of cognitive tests should be available for all clinical staff in the general hospital setting in the detection and diagnosis of delirium. Such a
test battery should be objective, user friendly and have high levels of clinical utility. This test battery should also form a vital part of an overall cognitive friendly hospital programme whereby delirium is prevented, detected and managed promptly and effectively.
Table 33: Future research

- Future research will involve a modified and improved version of the Lighthouse app with an enhanced lighthouse image. It will also be possible to immediately store and retrieve patient scores on the test.

- The modified Lightouse app will be tested for use by healthcare professionals of varying backgrounds and levels of experience and expertise.

- Clinical populations that have been relatively understudied will also be included in future research. These include younger populations, populations with high levels of comorbid delirium and dementia, dementia complicated by behavioural and psychological symptoms (BPSD) and other important comorbid neuropsychiatric conditions such as depression.

- Understudied clinical contexts will also be included in future research, such as community, outpatient, day-care and nursing home populations.
Study limitations

We studied consecutive referrals to a consultation-liaison service for assessment of neuropsychiatric status. Attempts were made throughout recruitment to assess approximately equal numbers of people with no neurocognitive dysfunction (NNCD), dementia alone, dementia with comorbid delirium and delirium alone and controls.

As such, these patients are likely to have a heightened symptom burden and are not representative of elderly inpatients in general. Cross-sectional studies cannot fully capture the profile of conditions such as delirium where symptom fluctuation is prominent.

Rater A and Rater B assessments were conducted on the same day, within 4-6 hours of each other. However, a shorter timescale between Rater A and Rater B assessments would have been preferable. Data was incomplete for some individuals, due to factors such as clinical condition and ability to cooperate with the full battery of assessments.

We included a range of bedside tests but in order to avoid testing fatigue this was not exhaustive. For example, the serial sevens test was not included, although previous work has suggested that this is pitched at a level of complexity that does not allow for favourable comparison with other tests when used in the assessment of patients with delirium (Adamis et al., 2015). We did not specify the stage or primary cause of dementia but evidence indicates that both neurocognitive burden and neuropsychiatric disturbance varies across dementia types (D'Onofrio et al., 2012).

Specifically regarding the Lighthouse application, an android specific application was used and this limits generalizability to other platforms.
Implications

Improved identification of major neurocognitive disorders is a key challenge in our increasingly aged society. In particular, accurate and consistent detection of delirium is a priority issue as evidence indicates that more than half of cases are missed or detected late in everyday practice (O’Hanlon et al, 2014). A fundamental factor in enhancing recognition rates is to identify a simple and brief means of establishing the presence of clinically significant cognitive impairment at the bedside. Guidance regarding differentiation of delirium and dementia is relatively lacking in the definition of delirium in DSM-5 (American Psychiatric Association, 2013) or ICD-10 (World Health Organisation, 1993), suggesting that these diagnostic systems would be advanced by criteria to guide efforts to distinguish these common conditions. Although both delirium and dementia involve generalised disturbance of cognitive function, this work emphasises how delirium can be distinguished by virtue of the disproportionate impairment of attention, vigilance and visuospatial abilities. These cognitive functions can be assessed with simple bedside tests and should be emphasised in efforts to identify delirium.
References


APA 1980. Diagnostic and Statistical Manual for Mental Disorders.


APA 1994. DSM-IV.

APA 2000. DSM-IV-TR.


BOSISIO, M., CARACENI, A. & GRASSI, L. 2006. Phenomenology of delirium in cancer patients, as described by the Memorial Delirium Assessment Scale (MDAS) and the Delirium Rating Scale (DRS). *Psychosomatics,* 47, 471-8.


EUROPEAN DELIRIUM ASSOCIATION, A. D. S. 2014. The DSM-5 criteria, level of arousal and delirium diagnosis: inclusiveness is safer. BMC Med, 12, 141.


RCO., P. 2012. Standardising the Assessment of Acute Illness Severity in the NHS. London: RCP.


