The Gerontologist
cite as: *Gerontologist*, 2015, Vol. 55, No. 6, 1079–1099
doi:10.1093/geront/gnv100

THE GERONTOLOGICAL SOCIETY OF AMERICA®

OXFORD

Literature Review

Delirium Screening: A Systematic Review of Delirium Screening Tools in Hospitalized Patients

Jayita De, BSc, MBBS, FRACP*[,1] and Anne P. F. Wand, BSc, MBBS, MPsychiatry, FRANZCP[2,3]

[1]Department of Aged Care, St George Hospital, Kogarah, New South Wales, Australia. [2]Older Adults Mental Health Service, St George Hospital, Kogarah, New South Wales, Australia. [3]Faculty of Medicine, University of New South Wales, Australia.

*Address correspondence to Jayita De, BSc, MBBS, FRACP, Department of Aged Care, St George Hospital, Chapel Street, Kogarah, NSW 2217, Australia. E-mail: Jayita.De@sesiahs.health.nsw.gov.au

Received January 31, 2015; Accepted June 4, 2015

Decision Editor: Rachel Pruchno, PhD

Abstract

Background: Delirium occurs commonly in hospitalized older patients but is poorly recognized. Although there are a plethora of validated delirium screening tools, it is unclear which tool best suits particular populations.

Purpose: To evaluate validation studies of delirium screening tools in non–critically ill hospital inpatients and provide guidance on the choice of screening tool.

Methods: The MEDLINE, CINAHL, and PsychInfo databases were searched for studies comparing delirium bedside screening tools with either the Diagnostic and Statistical Manual or International Classification of Diseases defined diagnosis of delirium in hospital inpatients. Information was also drawn from conference proceedings and discussion with delirium researchers.

Results: Thirty-one studies describing 21 delirium screening tools were included in the systematic review. The majority of studies were conducted across a broad range of inpatient settings internationally in elderly inpatients, including patients with dementia but most excluded nonnative language speakers.

Implications: The Confusion Assessment Method was the most widely used instrument to identify delirium, however, specific training is required to ensure optimum performance. The Delirium Rating Scale and its revised version performed best in the psychogeriatric population but requires an operator with psychiatric training. The Nurses' Delirium Screening Checklist appears best suited to the surgical and recovery room setting. The Single Question in Delirium shows promise in oncology patients. The Memorial Delirium Assessment Scale, while demonstrating good measures of validity in the surgical and palliative care setting, may be better used a measure of delirium severity. The 4As Test performed well when delirium was superimposed on dementia, but it requires further study.

Key Words: Delirium, Delirium Screening Tools, Diagnosis

Delirium is an acute disorder of attention, cognition, and psychomotor activity that commonly affects elderly people. The reported incidence of delirium during admission in the hospitalized adult population is 3%–29% (Siddiqi, House, & Holmes, 2006). The co-occurrence of delirium in patients with dementia is particularly high in hospitalized older adults (22%–89%, Fick, Agostini, & Inouye, 2002).

The overall risk of adverse outcome as a result of hospitalization in the elderly people, particularly those from residential care facilities, is already high for functional decline and falls (Friedman, Mendelson, Bingham, & McCann, 2008). The adverse outcomes for delirious patients may be even graver. They include multiple medical complications, greater lengths of stay, the possibility of not returning to independent living, and death (Cole et al., 2002; Elie et al., 2000; Inouye, 2006). Delirium superimposed upon dementia has been shown to prolong hospital stay and be associated with both cognitive and functional decline (Fick, 2013). Patients with dementia are particularly difficult to evaluate for delirium (Powers et al., 2013). However, there are clear advantages to early detection and targeted treatment (Chong, Chan, Tay, & Ding, 2014; Lundstrom et al., 2005; Mudge, Maussen, Duncan, & Denaro, 2013). It is thus imperative that delirium be correctly identified and managed to reduce the significant morbidity and mortality, particularly in the elderly people.

Delirium was described more than 2000 years ago and is a prevalent condition in hospitalized elderly population. It still remains underrecognized (Inouye, Westendorp, & Saczynski, 2014) and is often misdiagnosed (Inouye, 2006; Voyer, Cole, McCusker, St-Jacques, & Laplante, 2008; Wand et al., 2013). For example, a recent Australian study demonstrated that detection of delirium by staff was poor, with staff correctly identifying only 23% of cases with delirium despite a targeted multimodal educational intervention (Wand et al., 2014). Delirium remains understudied in relation to the proportion of its disease burden (MacLullich et al., 2013).

Delirium was only formally categorized by standardized diagnostic criteria, the Diagnostic and Statistical Manual (DSM) III, in 1980 (American Psychiatric Association, 1980) and in the International Classification for Diseases 10th Edition (ICD-10) in 1992 (World Health Organization, 1992). Prior to its inclusion in the DSM, delirium was described in the literature under various eponyms, such as acute confusional state, toxic encephalopathy, and toxic psychosis. The variable terms used led to much confusion in the detection of delirium and made it difficult to interpret the published literature (Hall, Meagher, & MacLullich, 2012). Over the last three decades, there have been significant improvements in the understanding of delirium and with it, revisions in the delirium diagnostic criteria (American Psychiatric Association 1987, 1994, 2013).

There are now numerous screening tools validated for the assessment of delirium. A recent review by Grover and Kate (2012) comprehensively identified and evaluated 38 separate instruments in use for screening, diagnosis, assessing cognitive function, assessing motor symptoms, risk factors, and grading severity of and quantifying the distress associated with delirium. In response to the limitations of existing screening measures, newer tools have emerged and are being validated (Lin et al., 2015; MacLullich, Ryan, & Cash, 2011; Sands et al., 2010). The plethora of tools available can make it difficult for the clinicians to decide which tool to use and in what context.

Therefore, the primary aim of this review was to identify, compare, and evaluate validation studies of delirium screening tools used in hospital inpatients. A secondary aim was to provide guidance regarding the clinical applicability of the reviewed screening tools to particular patient populations.

Methods

Search Strategy

The MEDLINE, CINAHL (1996 to July 2014), and PsychInfo (1987 to July 2014) databases were searched using the following combinations of keywords, with searches limited to articles with human subjects published in English: "delirium AND screening NOT intensive care unit (ICU),""delirium AND rating scale NOT ICU," "delirium AND tool NOT ICU," "delirium AND ident* NOT ICU," and "delirium AND validation NOT ICU". The citations obtained were examined to identify validation studies of delirium screening tools used in hospital inpatients. Information was also drawn from conference proceedings and discussion with experts involved in delirium research.

Inclusion and Exclusion Criteria

Articles were included if they reported original research comparing a bedside cognitive delirium screening tool against a standardized diagnosis of delirium using DSM or ICD criteria. The study population was limited to hospitalized adult inpatients, including those with dementia or terminal illness.

Publications were excluded if they studied delirium in the ICU population. Delirium in the ICU is usually considered a different entity to delirium in other settings, including the geriatric population, due to the severity of illness, invasive management strategies, and frequent need for intravenous and other sedation. Although there may be overlap in etiology, the methods of delirium screening in the ICU population are significantly different to warrant a review in its own right.

Articles solely concerned with rating delirium severity (rather than delirium identification), those which did not apply DSM or ICD criteria as a reference standard for delirium diagnosis or which validated a non-English version of a delirium screening tool, or looked at delirium screening in community settings were excluded. Review articles and

purely descriptive studies of delirium assessment tools were also not included.

Each abstract was reviewed by the authors to identify publications meeting the inclusion criteria. The full text of the article was retrieved when the information in the abstract was insufficient. If there was a discrepancy in articles sourced or uncertainty about whether a publication should be included, it was discussed between the authors until a consensus was reached. Data were collected where available on the age, sample size and English (or native language) speaking status of participants, number of patients with dementia, sensitivity, specificity, interrater reliability, and the time interval between the delirium and the reference standard assessments.

Validation studies are evaluated using sensitivity and specificity. Sensitivity describes the accuracy of a diagnostic test to identify positive cases of the condition when it is present. As such, a highly sensitive test will have few false negatives and will rule "out" conditions (SnOUT) when negative. Specificity describes the accuracy of a diagnostic test to correctly identify cases where the condition is not present. Thus, a highly specific test has few false positives and is effective in ruling conditions "in" (SpIN) when positive. Both of these tests report features of a given diagnostic test that are independent of the actual prevalence of the disease (retrieved from www.med.emory.edu/EMAC/curriculum/ diagnosis/sensand.html). The following ratings were used to rank performance of tools in this review with regard to sensitivity and specificity: Excellent >95%, Good 80%–94%, Moderate 70%–84%, and Poor <70%.

Assessment of Quality

The quality of data reporting for each individual study was assessed using the Standards for the Reporting of Diagnostic Accuracy (STARD) criteria. The STARD criteria were developed to allow detection of potential biases and judge the generalizability of studies of diagnostic accuracy. It is a 25-item checklist used to verify that all essential elements are included in the reporting of the study. The STARD is scored from 1–25, with a score greater than 20 considered high quality (Bossuyt et al., 2003a, 2003b; Morandi et al., 2012).

Results

There were 3,541 citations identified from the database searches and other sources. Overall 31 citations met the inclusion criteria (Figure 1).The studies were tabulated according to their inpatient population, that is, mixed hospitalized patients (general medical, geriatric, psychogeriatric, surgical, and rehabilitation; 20 studies); exclusively surgical patients (3 surgical ward, 2 recovery room), emergency department (2 studies), oncology (2 studies), and palliative care (2 studies). An overview of the operating characteristics of the most frequently studied tools is provided in Table 1.

Delirium Screening Tools

Twenty-one different tools were validated across the 31 studies reviewed (Table 2). The Confusion Assessment Method (CAM), brief CAM, and CAM-ICU were the most commonly studied tool (13/31 studies) followed by the Delirium Rating Scale and its revised version (DRS/DRS-R-98) (6/31 studies).

Patient Characteristics

The age range was 18–97 years, the majority being older patients. The studies were conducted in a broad range of settings (Tables 3–9). Eleven studies were conducted in the United States; four in Canada; three each in the United Kingdom and Australia; two each in Germany and Holland; and one each in Finland, Hong Kong, Italy, Poland, and Spain. Sixty-six percent (20/31) of the studies included patients with dementia. Four studies specifically included nonnative language speakers from the country the study was conducted in.

Reporting Quality

Most studies (25/31, 83%) had a high quality data reporting rating, that is, STARD Score greater than 20. Almost half the studies (13/31, 43%) did not report the time period between the reference standard delirium assessment and index (delirium screening tool) assessment. In the studies that did report the intervals between the assessments, the time periods ranged from 30 min (Bellilli et al., 2014; González et al., 2004, Monette et al., 2001) to more than 24 hr (Andrew et al., 2009; Breitbart et al., 1997; Pompei, Foreman, Cassel, & Cox, 2003; Ryan et al., 2009; Whittamore et al., 2014).

The DSM/ICD raters for delirium included specialist registrars, geriatricians, psychiatrists, or neuropsychologists in all the studies apart from one (Erkinjuntti, Sulkava, Wikstrom, & Autio, 1987), where the reference standard assessment was carried out by research assistants, indicating an advanced level of clinical experience and training overall.

Populations Studied

Mixed Hospital Inpatients

This was a heterogeneous group of studies that evaluated 13 different tools either in isolation or in combination with another tools against the DSM reference standard (Tables 3–5) The DRS (including the DRS-R-98) was the most commonly evaluated tool followed by the CAM (including the CAM-ICU).

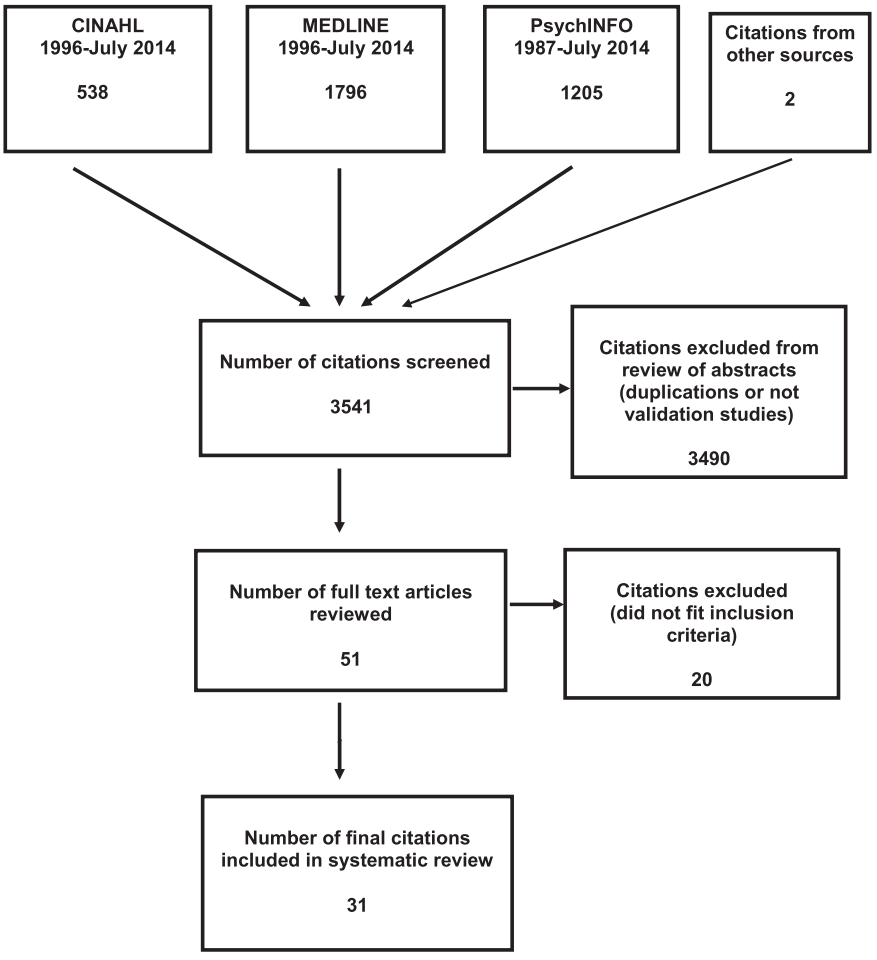

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram: Results of the literature search.

DRS and DRS-R-98

The DRS is a 10-item scale, rated by a clinician with psychiatry training, based on the patient's behavior over a 24-hr period (Table 3; Trzepacz, Baker, & Greenhouse, 1988). Rockwood, Goodman, Flynn, and Stolee (1996) reported a cutoff score greater than 10 on the DRS to achieve good sensitivity/specificity of 82% and 94% respectively, in contrast to the greater sensitivity and specificity obtained by Rosen and colleagues (1994) for the same cutoff scores. However, their differing populations and proportions of subjects with dementia may account for this variability. The two studies that evaluated the DRS-R-98 converged on the cutoff score of greater than 17.75 but demonstrated poor–moderate sensitivity/specificity with values less than 75% in the older adult population. The DRS-R-98 consists of two parts: 13 severity items and 3 diagnostic items. Of note, in both studies, the reference standard and index assessments were conducted more than 24 hr apart (Andrew et al., 2009; Whittamore et al., 2014).

Confusion Assessment Method

The index CAM study reported an excellent performance of the tool with sensitivities and specificities greater than 95% in a small number of elderly participants (Table 4; Inouye et al., 1990). Similar results were obtained in the study where the CAM and Delirium Symptom Index (DSI) were evaluated together, also in an elderly population (Yates et al., 2009). The CAM is comprised of nine criteria derived from the DSM III-R. The interrater reliability was high in both the studies with the best results obtained using trained operators.

Other Tools

The Delirium Diagnostic Tool-provisional (DDT-Pro) study, while reporting excellent sensitivity/specificity with values of 100% and 94% respectively, was tested in a small number of patients with acquired brain injury which limits its generalizability (Table 5). The interrater reliability was high and the tool correlated well with the DRS (Kean, Trzepacz, Murray, Abell, & Trexler, 2010). The DDT-Pro is comprised of the comprehension and vigilance sections from the Cognitive Test for Delirium and the sleep–wake cycle disturbance item from the DRS-R-98 (Kean et al., 2010). The Memorial Delirium Assessment Scale (MDAS) is a 10-item clinician rated scale, which assesses disturbances in arousal and level of consciousness, cognitive function, and

Table 1. Overview of the Most Frequently Studied Screening Tools.

Delirium screening tool	No. of studies	Criteria on which based	No. of items	Rating by	Time taken (min)	Population(s) tested in	Assessment characteristics	Sensitivity (%)	Specificity (%)
CAM (and variants)	13	DSM III-R	4/9	Trained clinician	5–7	ED, mixed, oncology, palliative, surgical	Operator training required to maximize accuracy Limited assessment of drowsy patient	46–100 (depends on operator experience)	63–100 (depends on operator experience)
DRS/DRS-R-98	6	DSM III	16	Clinician with psychiatry training	20–30	Mixed	Psychiatry specific training required Consensus unclear on cutoff score accepted for positive result Longest time to administer	91–100 (depends on cutoff score)	84–92 (depends on cutoff score)
MDAS	4	DSM IV	10	Physician	10–15	Surgical, oncology, palliative	Consensus unclear on cutoff score accepted for positive result	64.1	100
Nu-DESC	3	Confusion Rating Scale	5	Nursing staff	1–3	Surgical	Takes into account psychomotor changes Scale is rated with each shift change, total scores averaged over a 24-hr cycle to give final score	85.7	86.8
SQiD	2	Carers viewpoint	1	Clinician	5	Mixed, oncology	Single question Validated in a small oncology population	80	71
4AT	1	DSM IV	4	Anyone	2–3	Mixed	Allows assessment of the drowsy patient and delirium superimposed on dementia	89.7	84.1

Note: 4AT = The 4As Test; CAM = Confusion Assessment Method; DRS/DRS-R-98 = Delirium Rating Scale, Revised Version; DSM = Diagnostic and Statistical Manual; ED = emergency department; NuDESC = Nursing Delirium Screening Scale; SQiD = Single Question in Delirium.

Table 2. The Frequency of the Various Delirium Screening Tools.

Delirium screening tool	No. of studies
Confusion Assessment Method (CAM)	9
Memorial Delirium Assessment Scale (MDAS)	4
Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)	3
Delirium Rating Scale (DRS)	3
Delirium Rating Scale, Revised (DRS-R-98)	3
Nursing Delirium Screening Checklist (NuDESC)	3
Delirium Detection Score (DDS)	2
Delirium Observation Screening Scale (DOSS)	2
Digit Span Test (DST)	2
Single Question in Delirium (SQiD)	2
Delirium Symptom Interview (DSI)	1
Brief CAM (bCAM)	1
Clinical Assessment of Confusion (CAC)	1
Delirium Diagnostic Tool-provisional (DDT-Pro)	1
Delirium triage screen (DTS)	1
Intensive Care Delirium Screening Checklist (ICDSC)	1
Inter-RAI Acute Care Assessment System (four items pertaining to delirium)	1
Modified Richmond Agitation Sedation Scale (mRASS)	1
Simple Question for Easy Evaluation of Consciousness (SQUEEC)	1
Short Portable Mental Status Questionnaire (SPMSQ)	1
The 4As Test (4AT)	1
Vigilance A Test	1

Note: Some studies validated a combination of tools.

psychomotor activity. It is designed to quantify severity of delirium. The MDAS index study was tested in a modest number of patients with cancer and AIDS (Breitbart et al., 1997), also limiting its generalizability.

One of the newer tools, the Simple Query for Easy Evaluation of Consciousness (SQEEC) was notable as the only tool that invites a narrative (Lin et al., 2015). The patient is asked to describe a journey and the details of how they would undertake the journey. The SQUEEC's performance was good with a sensitivity of 83% and specificity of 81%. However, the inherent requirement for a narrative excludes assessment of those patients with reduced level of consciousness.

The 4A's Test (4AT) consists of four items, two brief cognitive tests, assessment of level of consciousness, and an acute change in mental state (MacLullich et al., 2011). A single validation study confirmed that the 4AT had a good sensitivity of 90% and specificity of 84% in screening for delirium in elderly hospital inpatients (Bellilli et al., 2014). The reported advantages of the tool were ease of use, rapid screening by an operator without requiring specific training, and the ability to assess patients with fluctuating level of consciousness and hypoactive delirium. A large scale, multicentre study is in progress, led by MacLullich and colleagues from the University of Edinburgh (see http://www.nets.nihr.ac.uk/projects/hta/1114301).

The Delirium Observation Screening Scale (DOSS) is a 25-item scale rated by nurses looking at typical behavior patterns in relation to delirium. Assessments are completed every nursing shift over 24 hr and the final score is the mean of all the shift scores. Although the DOSS had a favorable predictive validity against a DSM IV diagnosis of delirium, the sensitivity/specificity of the tool and time between the index and reference assessments were not reported. It was also validated in relatively small sample size (Schuurmans, Shortridge-Bagett, & Duursma, 2003).

Four items of the InterRAI Acute Care Assessment Systems referring directly to delirium (acute change in mental state, variable mental function through the day, disorganized speech, and distractibility) were extracted to produce a screening tool that performed well both in the general medical population and in those with dementia (Salih, Paul, Klein, Lakhan, & Gray, 2012). However, this study comprised of a subanalyses of data from a larger study looking at geriatric syndromes and outcomes (Lakhan et al., 2011). The tool requires further refinement to be adapted into common clinical practice.

The Clinical Assessment of Confusion (CAC), Vigilance A Test, and Digit Span Test (DST) demonstrated poor sensitivities of less than 70% in the mixed populations (Leung et al., 2011; Pompei et al., 1995). The CAC is a checklist of 25 items looking at psychomotor behaviors associated with confusion, with more behaviors indicating severe confusion (Vermeersch, 1990). The Vigilance A Test requires the operator to read out a series of 60 letters, among which "A" appears with greater than random frequency. The patient is asked to indicate each time an "A" is heard. In the DST, the patient is asked to listen to and repeat a series of numbers (Strub & Black, 2000). The modified Richmond Agitation Sedation Scale (mRASS) is an objective scale of determining level of consciousness, a modification of the RASS to include assessment of inattention (Chester, Beth Harrington, & Rudolph, 2012). It was studied in a population that lacked diversity and excluded those with cognitive impairment, a group particularly vulnerable to delirium. The DSI is a diagnostic interview for delirium, based on DSM III criteria (Albert et al., 1992). It showed good sensitivity and specificity of 91% and 80%, respectively (Yates et al., 2009), but has not been widely studied (Grover & Kate, 2012).

The Short Portable Mental Status Questionnaire (SPMSQ) was validated in an elderly population, including patients with dementia (Erkinjuntti et al., 1987). The SPMSQ is a 10-item tool that measures cognitive function in the elderly people (Pfeiffer, 1975). The quality of the data reporting was the lowest of all the studies. In particular, there was no mention of blinding between the assessors performing the index and reference tests, the time between assessments and quantification of interrater reliability.

Table 3. Mixed Population: Delirium Rating Scale and Revised Version.

Author (year), country	Study population	*N* (% male)	Mean age (*SD*)	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of linterrater reliability	Notes	STARD Score
Andrew et al. (2009), Canada DRS-R-98	Geriatric inpatients	145 (34)	81 .2	DSM IV	Not specified	Yes *n* = 58	Whole group: 56% Dementia: 59%	Whole group: 82% Dementia: 67%	>24 hr Single assessment	High intraclass correlation coefficient 0.92		23
Rockwood et al. (1996), Canada DRS	Geriatric and psychogeriatric inpatients	104	80	DSM III-R	Not specified	Yes *n* = 16	DRS cutoff 10: 82% DRS cutoff 8: 90%	DRS cutoff 10: 94% DRS cutoff 8: 82%	Range between "minutes to hours" Single assessment	High intraclass correlation coefficient 0.91		22
Rosen et al. (1994), USA DRS	Psychogeriatric inpatients	791 (33)	72.6 (9.4)	DSM III-R	Not specified	Yes *n* = 196	DRS threshold score >10: 94%	DRS threshold score >10: 82%	Simultaneous Single assessment	Not reported	DRS high sensitivity, thus good screening test Lower specificity limits the usefulness of DRS as a diagnostic tool	20
Trzepacz et al. (1988), USA DRS	General medical and surgical inpatients (referred for assessment by consultation—Liaison Psychiatry Service)	47 (43)	Delirium 58.8 (16.5) Dementia 78.4 (4.4) Schizophrenia 33.3 (7.3) Controls 42.1 (13.2)	DSM III	Not specified	Yes—used as a control group	Not reported	Not reported	Not reported Single assessment	High intraclass correlation coefficient 0.97	Box plots of DRS Scores in each of the groups clearly differentiated the delirious group (DRS 12–30) from the various control groups	19
Trzepacz et al. (2001), USA DRS-R-98	Medical, surgical, psychiatric, rehabilitation, and nursing home care inpatients	68 (75)	Delirium 64.0 (16.8) Dementia 75.5 (8.6) Schizophrenia 41.2 (15.2) Depression 57.9 (18.6) Other psych 49.6 (19.8)	DSM IV	Yes	Yes *n* = 13	DRS-R-98 total score cutoff 15.25: 92% DRS-R-98 severity scale best cutoff score 15.25: 92%	DRS-R-98 total score cutoff 15.25: 93% DRS-R-98 severity scale best cutoff score 15.25: 93%	Not reported Single assessment	High intraclass correlation coefficients 0.98–0.99	Box plots of the DRS-R-98 total score in each of the groups clearly differentiated the delirious group from the other groups	21
Whittamore et al. (2014), UK DRS-R-98	Geriatric trauma orthopedic general medical inpatients	249 (34)	84	DSM IV	Not specified	Yes *n* = 106	DRS-R-98 cutoff >17.75: 75%	DRS-R-98 cutoff >17.75: 71%	Conducted during same admission Single episode of assessment	Not reported	Difficult population to recruit and study due to the procedures of acute medical admissions in UK	22

Note: CIND = cognitively impaired no dementia; DRS = Delirium Rating Scale; DRS-98 = Delirium Rating Scale, Revised; DSM = Diagnostic and Statistical Manual; STARD = Standards for the Reporting of Diagnostic Accuracy.

Table 4. Mixed Population: Confusion Assessment Method.

Author (year), country	Study population	*N* (% male)	Mean age (*SD*)	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
González et al. (2004), Spain	General hospital inpatients, excluding psychiatry ward patients	123 (38)	77.7 (8.2)	DSM IV	Yes	Yes *n* = 62	Whole group: 90% Dementia: 87% No Dementia: 93%	Whole group: 100% Dementia: 100% No Dementia: 100%	30 min Single assessment	High kappa rating 0.89	CAM adapted for Spanish speaking population	22
Inouye et al. (1990), USA	General medical inpatients	Site 1: 30 (86) Site 2: 26 (86)	Site 1: 80.3 (7.9) Site 2: 77.1 (8.7)	DSM III	Not specified	Yes *n* = 12 Site 1: 9 Site 2: 3	Site 1: 100% Site 2: 94%	Site 1: 95% Site 2: 90%	6 hr Single assessment	High kappa rating 1.0	Small numbers Training required for optimal use of the CAM	25
Yates et al. (2009), UK CAM/DSI	General medical inpatients	62	81	DSM IV	Not specified	Not specified	CAM vs. DSM-IV: 90% DSI vs. DSM-IV: 91%	CAM vs. DSM IV: 96% DSI vs. DSM-IV: 80%	Not reported Single assessment	Not reported	DSI and CAM showed similar high internal consistency (Chronbach alpha of 0.857 and 0.841) CAM was a good screening instrument but requires specific training for use Lowest quality of data reporting	13
Powers et al. (2013), USA CAM-ICU/RASS	Geriatric rehabilitation inpatients	61 (97)	73	DSM IV	Not specified	Yes *n* = 7	Not reported	Not reported	Up to 4 hr Multiple assessments	Not reported	97% (59) of participants were men RASS used to assess motor subtype of delirium Poor agreement between the CAM-ICU and reference standard assessment in identifying delirium in this population ALL of the dementia patients demonstrated hypoactive delirium by at least one of the assessment methods	22

Note: CAM = Confusion Assessment Method; CAM-ICU = Confusion Assessment Method for the Intensive Care Unit; DSI = Delirium Symptom Interview; DSM = Diagnostic and Statistical Manual; RASS: Richmond Agitation Sedation Scale; STARD = Standards for the Reporting of Diagnostic Accuracy.

Table 5. Mixed Population: Other Screening Tools.

Author (Year), country	Study population	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Bellilli et al. (2014), Italy	Geriatric inpatients	234 (36)	84 (5.9)	4AT	DSM IV	No	Yes *n* = 74	Whole group: 90% No dementia: 83% Dementia: 94%	Whole group: 84% No dementia: 91% Dementia: 91%	15–30 min Single assessment	Not assessed	Chronbach's alpha for internal consistency 0.8 (good)	23
Breitbart et al. (1997), USA	General hospital inpatients, cancer and AIDS inpatients	Pilot study: 33 (45) Follow-up study: 51 (57)	Pilot study: 56 Follow-up study: 58	Pilot study: MDAS Follow-up study: MDAS vs. DRS/ MMSE/ Clinicians' global rating of delirium severity	Pilot study: DSM III/IV Follow-up study: DSM IV	Not specified	Not specified	Pilot study: MDAS cutoff score >13: 65%	Pilot study: MDAS cutoff score >13: 94%	24 hr Single assessment	Not reported	Some study participants had received treatment for their delirium prior to the study assessments	21
Chester et al. (2012), USA	General medical inpatients	95 (94)	81(7.3)	Modified RASS	DSM IV	Not specified	No	Single day assessment: 64% Longitudinal assessment: 74%	Single day assessment: 93% Longitudinal assessment: 92%	Not reported Multiple assessment	High 98% agreement with kappa 0.48 ($p < .001$)	Lack of diversity within the study population (94% men, 89% Caucasian)	20
Erkinjuntti et al. (1987), Finland	General medical inpatients	282 (39)	75.5 (7.2)	SPMSQ	DSM III (based on publication date, not specified in text)	Not specified	Yes *n* = 34	Cutoff point of 4 errors on SPMSQ 17% Sensitivity rose to 73% with cutoff of 2 errors	Cutoff point of 4 errors on SPMSQ 100% Specificity fell to 89% with cutoff of 2 errors	Not reported Single assessment	Not reported	SPMSQ is more sensitive screening tool for dementia than for delirium 52 subjects (20%) of the delirious patients were untestable with the SPMSQ Poor quality of data reporting	11
Kean et al. (2010), USA	Rehabilitation inpatients	36 (64)	44.2 (15.7)	DDT-Pro	DSM IV TR & DRS-98	No	Not specified	DDT-pro cutoff >7: 100%	DDT-Pro cutoff >7: 94%	"Few hours" Multiple assessments	High intraclass correlation coefficient 0.98	Younger population, acquired brain injury	20
Lin et al. (2015), Australia	General medical inpatients	100	≥75	SQEEC/ SQiD	DSM IV	Not specified	Yes *n* = 30	SQEEC whole group: 83% SQEEC dementia: 83% SQiD 77%	SQEEC whole group 81% SQEEEC dementia: 59% SQiD 51%	Not reported Single assessment	Data not available	Excluded patients with reduced level of consciousness (i.e., excluded possible hypoactive delirium)	N/A[a]

Table 5. *Continued*

Author (Year), country	Study population	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Leung et al. (2011), Hong Kong	Geriatric and medical inpatients	144 (54)	77.9 (6.5)	DST	DSM IV	Yes	Yes *n* = 44	Delirium vs. no delirium (inc delirium superimposed on dementia) Digit span forward: 58% Digit span backward: 81%	Delirium vs. no delirium (inc delirium superimposed on dementia) Digit span forward: 72% Digit span backward: 63%	4 hr Single assessment	Not reported		21
Pompei et al. (1995), USA	Medical and surgical inpatients from four units	432	74.2 (6.7)	DST/ Vigilance A Test/ CAC/CAM	DSM III	No	Yes (numbers not specified)	DST: 34% Vigilance A Test: 61% CAC: 36% CAM: 46%	DST: 90% Vigilance A Test: 77% CAC: 95% CAM: 92%	24 hr Multiple assessment	Not reported	All tests performed less well in the subgroup with cognitive impairment	21
Salih et al. (2012), Australia	General medical inpatients at two sites	239 (49)	81.6 (6.4)	Four items within the inter-RAI acute care system	DSM IV	Yes	Yes *n* = 60	Whole population: 82% Pre-existing cognitive impairment (as determined by CPS ≥2): 90%	Whole population: 91% Pre-existing cognitive Impairment (as determined by CPS ≥2): 69%	4 hr Single assessment	Not reported	Sensitivity not affected by the presence of cognitive impairment.	20
Schuurmans et al. (2003), Holland	Study 1: geriatric inpatients	82 (29)	Study 1:83 (6.17)	Study 1: DOSS	Studies 1 & 2: DSM IV	Not specified	Yes	Not reported	Not reported	Not reported Multiple assessment in both studies	Not reported	Patients excluded if delirious on admission	21
	Study 2: Surgical ward postoperative hip fracture patients	92 (13)	Study 2: 82 (6.65)	Study 2: DOSS & CAM									

4AT = The 4As Test; CAC = Clinical Assessment of Confusion; CAM = Confusion Assessment Method; CPS = Cognitive Performance Scale; DDT-Pro = Delirium Diagnostic Tool-provisional; DOSS = Delirium Observation Screening Scale; DRS = Delirium Rating Scale; DSM = Diagnostic and Statistical Manual; DST = Digit Span Test; MDAS = Memorial Delirium Assessment Scale; MMSE = Mini-Mental State Examination; mRASS = Modified Richmond Agitation Sedation Scale; SQiD = Single Question in Delirium; SQEEC = Simple Question for Easy Evaluation of Consciousness; SPMSQ = Short Portable Mental Status Questionnaire; STARD: Standards for the Reporting of Diagnostic Accuracy.

[a]Information drawn from conference presentation, manuscript in press.

Table 6. Surgical Population.

Author (year), country Setting	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speakers inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Kazmierski et al. (2008), Poland Postcardiothoracic surgery	260 (76)	62 (9.0)	MDAS	DSM IV/ ICD-10	Not specified	Not specified	MDAS cutoff >9 vs. ICD-10: 96% MDAS cutoff > 9 vs. DSM-IV: 97%	MDAS cutoff > 9 vs. ICD-10: 93% MDAS cutoff > 9 vs. DSM-IV: 96%	Not reported Single assessment	Not reported	ICD-10 and DSM IV agreement kappa statistic 0.86 (very good) MDAS cutoff > 9 produced the optimal results for positive result In contrast to MDAS cutoff of 13 by Breitbart et al.)	21
Koster et al. (2009), Holland Postcardiothoracic surgery	112 (63)	70 (7.3)	DOSS	DSM IV	Not specified	No	100%	97%	Not reported Multiple assessments	Not reported	Specificity was estimated assuming that there was no delirium in patients with DOSS<2.	19
Radtke et al. (2010), Germany Postoperative elective surgery	88 (65)	66	CAM/DDS/ NuDESC	DSM IV	No	No	CAM: 75% DDS: 71% NuDESC: 98%	CAM: 100% DDS:87% NuDESC: 92%	Not reported Multiple assessments	CAM: high kappa rating 1.0 DDS: good kappa rating 0.77 NuDESC: good kappa rating 0.83	NuDESC most sensitive for delirium screening on the surgical ward	20
Radtke et al. (2008), Germany Recovery room: postoperative general anesthetic	154 (60)	55	CAM/DDS/ NuDESC	DSM IV	No	No	CAM: 43% DDS: 14% NuDESC: 95%	CAM: 98% DDS: 99% NuDESC: 87%	Not reported Single assessment	Not reported	DDS criteria do not correspond directly to the DSM-IV criteria DDS also looks for agitation, not psychomotor retardation, noting that hypoactive delirium is more frequent form of delirium	20

Table 6. *Continued*

Author (year), country Setting	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speakers inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Neufeld et al. (2013), USA Recovery room: postoperative general anesthetic	91 (42)	79 (6)	NuDESC/ CAM-ICU	DSM IV	No	No	CAM-ICU in recovery room: 28% CAM-ICU repeated on ward 28% NuDESC in recovery room 32% NuDESC (scoring threshold >2) repeated on ward 29% NuDESC (scoring threshold >1) In recovery room: 80% NuDESC (scoring threshold >1) repeated on ward: 72%	CAM-ICU in recovery room: 98% CAM-ICU repeated on ward 100% NuDESC in recovery room 92% NuDESC (scoring threshold >2) repeated on ward 96% NuDESC (scoring threshold >1) In recovery room: 69% NuDESC (scoring threshold >1) repeated on ward: 80%	60 min Multiple assessments	CAM-ICU: high kappa rating 1.0 NuDESC: high kappa rating 1.0 DSM iV: high kappa rating 0.93	Operators were trained in the use of the CAM	24

Note: CAM = Confusion Assessment Method; CAM-ICU = Confusion Assessment Method for the Intensive Care Unit; DDS = Delirium Detection Score; DOSS = Delirium Observation Screening Scale; DSM = Diagnostic and Statistical Manual; ICD = International Classification of Diseases; MDAS = Memorial Delirium Assessment Scale; NuDESC = Nursing Delirium Screening Checklist; STARD = Standards for the Reporting of Diagnostic Accuracy.

Table 7. Emergency Department.

Author (year), country Setting	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speakers inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of Interrater reliability	Notes	STARD Score
Han et al. (2013), USA	406 (50)	73.5	2 step: delirium triage screen/ brief CAM	DSM IV	No	No	DTS: Physician: 98% RA: 98% bCAM: Physician 84% RA: 78% **Combined:** RA DTS+ Physician bCAM 84% RA DTS + bCAM 78% Physician DTS+bCAM 82%	DTS: Physician 55% RA 56% bCAM: Physician 96% RA 97% **Combined:** RA DTS+ Physician bCAM 96% RA DTS+ bCAM 97% Physician DTS+bCAM96%	3 hr Single assessment	High kappa rating 0.88	Enrolment of a convenience sample of the elderly patient ED presentations Single-centre urban setting limits generalizability Screening tests performed by research assistants given a 6- to 8-hr training session prior to study initiation	24
Monette et al. (2001), Canada	906	>65 (mean age not reported)	CAM by trained non physician interviewer	DSM III-/ IV	No	Yes Number not reported	CAM total: 86% CAM total in dementia: 96% CAM total without dementia: 95%	CAM total: 100% CAM total in dementia: 84% CAM total without dementia: 95%	Simultaneous Single assessment	High kappa rating 0.91	5-day training provided to clinicians prior to use of CAM	19

Note: CAM = Confusion Assessment Method; bCAM = brief CAM; DSM = Diagnostic and Statistical Manual; DTS = delirium triage screen; ED = emergency department; RA = research assistant; STARD = Standards for the Reporting of Diagnostic Accuracy.

Table 8. Oncology Care Population.

Author (year), country	*N* (%male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Neufeld et al. (2011), USA	139 (45)	57	CAM-ICU/ ICDSC	DSM IV	Not specified	Yes *n* = 2	CAM-ICU first day of observation: 18% CAM-ICU multiple days of observation: 18%	CAM-ICU first day of observation: 99% CAM-ICU multiple days of observation: 99%	1 hr Multiple assessments	CAM-ICU: High kappa rating 1.0	Operators were trained in use of the CAM-ICU and ICDSC	24
							ICDSC first day of observation: 47% ICDSC multiple days of Observation: 66%	ICDSC first day of observation: 98% ICDSC multiple days of observation: 98%		ICDSC: high DSM IV: high kappa rating 0.93	Screening tools all used in rousable subjects, may not have captured some patients with hypoactive delirium CAM-ICU and ICDSC insensitive for delirium screening in this population	
Sands et al. (2010), Australia	19 (58)	53.2	SQiD+MDAS/ MMSE/CAM	DSM IV	Not specified	Not specified	SQiD vs. psychiatric interview: 80% SQiD vs. CAM: 67% SQiD vs. MDAS: not detected SQiD vs. MMSE: 50%	SQiD vs. psychiatric interview: 71% SQiD vs. CAM: 67% SQiD vs. MDAS: 64% SQiD vs. MMSE: 59%	Assessments conducted sequentially by same clinician: SQiD was applied first, followed by the CAM, MDAS and MMSE Single assessment	Not reported—Assessments conducted by same clinician	SQiD tool was applied before any of the other assessments Very small sample size Of note, CAM vs. DSM IV Psychiatry interview sensitivity was low at 40% when administered by staff with minimal training	20

Note: CAM = Confusion Assessment Method; CAM-ICU = Confusion Assessment Method for the Intensive Care Unit; DSM = Diagnostic and Statistical Manual; ICDSC = Intensive Care Delirium Screening Checklist; MDAS = Memorial Delirium Assessment Scale; MMSE = Mini-Mental State Examination; SQiD = Single Question in Delirium.

Table 9. Palliative Care Population.

Author (year), country	*N* (% male)	Mean age (*SD*)	Tool	Reference standard	Nonnative language speaker inclusion	Dementia	Sensitivity	Specificity	Time between index and reference assessment	Rating of interrater reliability	Notes	STARD Score
Lawlor et al. (2000), Canada	104	64.1 (10)	MDAS	DSM IV	Not specified	Yes *n* = 3	MDAS cutoff score of 7: 97%	MDAS cutoff score of 7: 95%	Not reported Multiple assessments	Good Intraclass correlation coefficient 0.89	MDAS cutoff score of 7 gave optimal results The MDAS and DSM assessments were carried out by the same physicians	20
Ryan et al. (2009), Ireland	Pilot phase: 32 (47) Main phase: 52 (46)	Pilot phase: 69.3 Main phase: 69.2	CAM	Psychiatrist interview, delirium diagnosis with DRS-R-98	No	Yes *n* = 7 (main phase)	Pilot phase: 50% Main phase: 88%	Pilot phase 100% Main phase: 100%	24 hr Single assessment	Not reported	Pilot phase: Non consultant hospital doctors had a 1-hr training session with CAM Main phase: enhanced training program for the non consultant Hospital doctors: two 1-hr sessions Improved sensitivity with the training SOMC was used in place of MMSE for cognitive status	20

Note: CAM = Confusion Assessment Method; DRS-R-98 = Delirium Rating Scale, Revised; DSM; Diagnostic and Statistical Manual; MDAS = Memorial Delirium Assessment Scale; MMSE = Mini-Mental State Examination; STARD = Standards for the Reporting of Diagnostic Accuracy; SOMC: Short Orientation Memory Concentration Test.

Exclusively Surgical or Recovery Ward Inpatients

There were five studies of exclusively surgical or recovery ward patients evaluating the CAM, Delirium Detection Score (DDS) and Nurses Delirium Screening Checklist (NuDESC) (Table 6). The three screening tools performed better in surgical ward patients (Radtke et al., 2010) when compared with the recovery room study (Radtke et al., 2008). The population age was older in the surgical ward patients with perhaps a higher incidence of delirium compared with the relatively younger recovery room patients. The NuDESC is a 5-item scale assessing disorientation, inappropriate behavior and communication, hallucinations, and psychomotor retardation over a 24-hr period (Gaudreau, Gagnon, Harel, Tremblay, & Roy, 2005). It had the best sensitivity and specificity of the three tools in the surgical ward population. The DDS is also a 5-item scale rating orientation, hallucinations, agitation, anxiety, and paroxysmal sweating (Radtke et al., 2010). When the CAM, CAM-ICU, DDS, and NuDESC were tested on postoperative patients in the recovery room, the tools demonstrated a low sensitivity (values less than 50%) and high specificities (values greater than 95%, increasing the rate of false-negative screening for delirium in this population (Neufeld et al., 2013; Radtke et al., 2008). Patients with dementia were excluded in both of the studies, and there was poor agreement with the CAM-ICU and NuDESC tools when compared with the reference standard interview (Neufeld et al., 2013). This may be due to the reference standard DSM IV assessment including a neuropsychiatric examination and detecting more subtle cases of delirium. The NuDESC showed improved sensitivity when the scoring threshold was lowered (Neufeld et al., 2013) and a good performance in the other study (Radtke et al., 2008).

Two of the studies examined the DOSS and MDAS in a postcardiac surgery population (Kazmierski et al., 2008; Koster, Hensens, Oosterveld, Wijma, & van der Palen, 2009) and reported excellent sensitivity and specificity with values greater than 95%. The MDAS cutoff score was lower than that reported in the index study (10 vs. 13; Breitbart et al., 1997). The DOSS study had potential for bias as the patients were only referred for the delirium reference standard interview if their DOSS was greater than two (Koster et al., 2009). The specificity was estimated assuming that there was no delirium in patients who scored less than two on the DOSS. Thus there was also the potential for false-negative results and missing subsyndromal delirium or fluctuations.

Emergency Department

The CAM and its abbreviated form, brief CAM (bCAM), were used in both emergency department (ED) studies (Table 7). The study carried out by Monette and colleagues (2001) is one of the few validation studies of the CAM in ED geriatric population and included dementia patients. In both studies, the operators were provided training prior to the use of the screening tools. In the study by Monette and colleagues (2001), a 5-day training period was provided to the researchers prior to use of the CAM, which may account for the reported high sensitivity of 86% and excellent specificity of 100%. In the study by Han and colleagues (2013), a large number of participants were evaluated in a two-step delirium screen. Patients were initially screened with the delirium triage screen (DTS) to assess level of consciousness and attention. Those who screened positive with the DTS were assessed with the bCAM. The operators were provided with a 4- to 6-hr training session prior to the start of the study. The DTS had excellent sensitivity (98%), and the bCAM had excellent specificity (96%). Both tools can be used by any health care clinician. The DTS consists of measuring level of consciousness using the RASS and attention with spelling "lunch" backwards (Han et al., 2013).

Oncology Population

CAM-ICU/Intensive Care Delirium Screening Checklist (ICDSC) and Single Question in Delirium (SQiD) were evaluated in oncology patients (Table 8). The study by Neufeld and colleagues (2011) only assessed rousable patients, potentially missing those with hypoactive delirium. The CAM-ICU and ICDSC had poor sensitivities with values of well below 75%. The ICDSC consists of eight items each corresponding to a characteristic behavior in delirium, scored by the patient's primary nurse over 24 hr (Bergeron, Dubois, Dumont, Dial, & Skrobik, 2001). The SQiD consists of a single question directed to carers asking whether the patient was more confused than usual. The sample size was very small in the study by Sands and colleagues (2010). The sensitivity/specificity of the SQiD was moderate, with values of 80% and 71% respectively, when compared with the DSM reference standard interview. Notably, the CAM's sensitivity for detecting delirium in this study was poor when used by staff with minimal training (Sands et al., 2010).

Palliative Care Population

The CAM and MDAS were evaluated in the palliative care setting across a wide range of ages (Table 9; Lawlor et al., 2000; Ryan et al., 2009). The CAM was validated in a pilot and subsequent main phase. The operators participated in a comprehensive training session for using the CAM between the two phases with consequent improvement in the sensitivity of the tool. The MDAS performance was excellent with sensitivity and specificity values greater than 95% at a lower cutoff score of 7 in contrast to the MDAS index study which used a higher cutoff score of 13 (Breitbart et al., 1997).

Discussion

This systematic review highlights the eclectic range of delirium screening tools validated for use in the inpatient setting. Older inpatients were well represented reflecting

current international hospital populations, as were patients with dementia. The studies were conducted in numerous countries, although there were few from Asia. There were 21 different tools evaluated, 11 tools were represented by single studies only.

Particular tools performed better in specific patient populations. In summary, the CAM performed well across the ED, postoperative and mixed inpatient settings, however, with several limitations, which are discussed later. The DRS/DRS-R-98 performed best in the psychogeriatric population and was tested on a very small number of surgical patients. The 4AT's sensitivity and specificity were good in the mixed inpatient setting, specifically in geriatric and dementia populations. The NuDESC and MDAS performed well in the postoperative settings. The SQiD is a simple tool that shows promise in the oncology setting. The MDAS had excellent validity in the palliative care population.

This review confirms that the CAM is a widely used instrument to identify delirium in research (Wei, Fearing, Sternberg, & Inouye, 2008; Wong et al., 2010). The original validation study for the CAM was based on observations made during a brief, structured interview that included the Mini-Mental State Examination (MMSE; Inouye et al., 1990). Specific training is recommended to ensure optimum performance (Inouye, 2003). This review highlighted poor sensitivity when the CAM is used by untrained operators (Ryan et al., 2009; Monette et al., 2001; Sands et al., 2010). However, the CAM was well-suited to (and the only tool tested in) the ED, where there is a need for a quick, simple tool. The two-step screening employed by Han and colleagues (2013) using the DTS/bCAM in trained operators particularly warrants further research. The MDAS has a lack of consensus on the cutoff scores for a positive screen, and it does not include key features of delirium, such as acuity of onset and variability of symptoms, limiting its use as a delirium screening tool (Trzepacz, 1994). A systematic review by Adamis, Sharma, Whelan, and Macdonald (2010) concluded that the MDAS is best used to rate delirium severity after diagnosis.

In mixed hospital patients, the use of the DRS and DRS-R-98 in routine clinical practice may be limited by the training required to use the tool, time taken to administer, and essential baseline knowledge of psychiatry. The DRS does not have an item for inattention or the ability to distinguish between the motor subtypes of delirium, limitations that were rectified in the DRS-R-98 (Rosen et al., 1994). The Vigilance A Test and DST have limited value in screening for delirium when used in isolation (Wong, Holroyd-Leduc, Simel, & Straus, 2010). The CAC was developed to assess "global confusion" rather than delirium (Vermeersch, 1990). It has 25 items, but none of the rated symptoms are specific to delirium (Adamis et al., 2010). The DOSS performed well in the surgical population but has also not been widely studied. It was difficult to objectively compare the remaining tools as they were evaluated in single studies with varying methodologies and results.

The NuDESC shows potential in the surgical and recovery room population. (Radtke et al., 2008, 2010). However, the total score is based on a 24-hr cycle of observation across nursing shifts, which while addressing the issue of fluctuation, is time consuming and may be difficult to achieve in patients with short lengths of stay. In oncology patients, the SQiD appears to be a reasonable choice, However, the psychometric properties were evaluated in a very small number of subjects. Further research is underway to assess its utility in a larger inpatient population.

Given the high prevalence of dementia in older inpatients and high rates of superimposed delirium, it is important to examine the validity of delirium screening tools in this population. Two thirds of the studies in this systematic review specifically included patients with dementia. This is a group that is poorly recognized (Fick et al., 2002), as it is often a challenge to differentiate signs of delirium from the cognitive impairment of dementia. A recent systematic review found that the CAM and CAM-ICU had the most support for the diagnosis of delirium superimposed on dementia, however, the evidence base was small (Morandi et al., 2012). There are also significant issues when interpreting the results of the CAM when administered by untrained operators in this setting. It may be difficult to accurately establish an acute change or fluctuation in the subject's mental state without a reliable collateral history. Impairment in attention and disorganized thinking may exist prior to the onset of delirium (Bellilli & Trabucci, 2010). The DRS-R-98 can differentiate delirium from dementia, depression, and schizophrenia, however, it requires training and psychiatric expertise. In contrast, the 4AT with its brevity and ease of use may be utilized when delirium is superimposed on dementia.

It has been highlighted that non-English (or nonnative language speaking) patients are often excluded from delirium research studies (Siddiqi et al., 2006; Wand et al., 2013). There were only 4 of 31 (13%) studies identified by this review which included such patients, and each evaluated a different tool (González et al., 2004; Leung et al., 2011; Salih et al., 2012; Trzepacz et al., 2001). It is therefore still unclear which tool is preferred in culturally and linguistically diverse patients.

Methodological Problems of Studies

There were some common methodological limitations of the validation studies. Almost half of the studies did not clearly report the time between the index and reference assessments. Given the hallmark of delirium is fluctuation, assessments markedly separated in time creates the potential for a missed diagnosis and bias (Morandi et al., 2012). Similarly, only 10 of the 31 studies included multiple, repeated index and reference standard delirium assessments over the course of admission. Given the fluctuation

in symptoms of delirium, a single assessment may be insufficient to detect delirium.

The index assessments in the studies were mostly carried out by experienced researchers or those who had been specifically trained in the use of a particular delirium screening tool. This limits the generalizability to the "real world" clinical setting, where it may not be feasible to offer specific training to all staff who would encounter delirious patients.

Limitations

The limitations of this systematic review include the restriction of the search to articles published in English, inpatient populations, the specific use of "delirium" as a search term (vs. acute confusional state, organic brain syndrome etc.), and publication bias. The intention was to ensure all included studies were referring to a standardized diagnosis of delirium by a validated method. Another limitation is the exclusion of studies that validated a non-English language version of a delirium screening tool.

Not all validated screening tools were identified. For example, the search strategy did not locate the primary validation study of the Neelon and Champagne (NEECHAM) Confusion Scale (Neelon, Champagne, Carlson, & Funk, 1996). However, it has been suggested that the NEECHAM scale measures "acute confusion" rather than delirium as defined by standardized criteria (Rapp et al., 2000), which may account for it not appearing in the database search.

Conclusion

The wide variety of delirium screening tools available adds to the complexity of delirium assessment (MacLullich et al., 2013). The tools themselves vary in number of items, time required to administer, level of knowledge assumed, and necessary training prior to use. The setting must also be considered when choosing the most appropriate tool. For example, brevity and ease of use are key factors in ED patients, demonstrated by the use of the DTS/bCAM (Han et al., 2013). Discrimination of delirium from psychiatric illness is critical in a psychogeriatric population, where despite its limitations, the DRS-R-98 is most useful. An additional challenge is the lack of consensus on cutoff scores for some tools (DRS and MDAS).

It is also important to have delirium screening tools that are relevant in real-world patients, such as those with comorbid dementia, those with various motor subtypes of delirium, and patients who are non-English speaking. Studies of delirium phenomenology have shown that hypoactive delirium is more prevalent than the hyperactive variant (Meagher et al., 2011, 2012; Spiller & Keen, 2006), yet some of the newer and other validated tools either cannot be used or do not perform well in hypoactive patients (Lin et al., 2015; Radtke et al., 2008). The 4AT is notable in this regard for its ability to assess patients with psychomotor changes and markedly reduced level of consciousness.

This systematic review identified a paucity of literature on delirium screening tools used in patients from cultural and linguistically diverse populations. With the availability of the major screening tools in multiple languages and the wide use of health care service interpreters, further research should be conducted in this field.

There is an ongoing need to develop and validate a delirium screening tool that is brief, can be used by operators without specific training, and can be used to assess motor fluctuations and delirium in dementia. It is also imperative that the tool be easily incorporated into busy clinical practice. From this review, the tools meeting these criteria are the 4AT, DTS-bCAM, and NuDESC. The approach to each acutely unwell patient needs to be "is there a delirium?" with the use of an appropriate screening tool for early detection and better management to improve outcomes for patients.

References

Adamis, D., Sharma, N., Whelan, P. J. P., & Macdonald, A. J. D. (2010). Delirium Scales: A review of current evidence. *Aging and Mental Health*, **14**, 543–555. doi: 10.1080/13607860903421011

Albert, M. S., Levkoff, S. E., Reilly, C., Liptzin, B., Pilgrim, D., Cleary, P. D., Evans, D., & Rowe, J. W. (1992). The delirium symptom interview: An interview for the detection of delirium symptoms in hospitalized patients. *Journal of Geriatric Psychiatry and Neurology*, **5**, 14–21. doi: 10.1177/002383099200500103

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). (DSM-III). Washington, DC: American Psychiatric Association. doi: 10.1176/dsm 10.1176/appi.books.9780521315289

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Revised (DSM-III-R). Washington, DC: American Psychiatric Association. doi: 10.1176/dsm 10.1176/appi.books.9780890420188

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Text Revision (DSM-IV). Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890423349.11547

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Text Revision (DSM-V).Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596

Andrew, M. K., Bhat, R., Clarke, B., Freter, S. H., Rockwood, M. R. H., & Rockwood, K. (2009). Inter-rater reliability of the DRS-R-98 in detecting delirium in frail elderly patients. *Age and Ageing*, **38**, 241–244. doi: 10.1093/ageing/afn298

Bellilli, G., & Trabucci, M. (2010). Reply to editor. *Journal of the American Medical Association*, **304**, 2124–2125. doi:10.1001/jama.2010.1615

Bellilli, G., Morandi, A., Davis, D. H. J., Mazzola, P., Turco, R., Gentile, S., … MacLullich A. M. J. (2014). Validation of the 4AT, a new instrument for rapid delirium screening: A study in 234 hospitalised older people. *Age and Ageing*, **43**, 496–502. doi: 10.1093/ageing/afu021

Bergeron, N., Dubois, M. J., Dumont, M., Dial, S., & Skrobik, Y. (2001). Intensive Care Delirium Screening Checklist: Evaluation of a new screening tool. *Intensive Care Medicine*, **27**, 859–864. doi: 10.1007/s001340100909

Breitbart, W., Rosenfeld, B., Roth, A., Smith, M. J., Cohen, K., & Passik, S. (1997). The Memorial Delirium Assessment Scale. *Journal of Pain and Symptom Management*, **13**, 128–137. doi: 10.1016/S0885-3924(96)00316-8

Bossuyt, P. M., Reitsma, J. R., Bruns, D. E., Gatsonis, C. A., Gasziou, P. P., Irwig, L. M., … Lijmer, J. G. (2003a).The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clinical Chemistry*, **49**, 7–18. doi: 10.1373/49.1.7

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., … de Vet, H. C. (2003b). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *British Medical Journal*, **326**, 41–44. doi: 10.1136/bmj.326.7379.41

Chester, J. G., Beth Harrington, M., & Rudolph, J. L. (2012). Serial administration of a modified Richmond Agitation and Sedation Scale for delirium screening. *Journal of Hospital Medicine*, **7**, 450–453. doi: 10.1002/jhm.1003

Chong, M. S., Chan, M., Tay, L., & Ding, Y. Y. (2014). Outcomes of an innovative model of acute delirium care: The Geriatric Monitoring Unit (GMU). *Clinical Interventions in Aging*, **9**, 603–612. doi: 10.2147/CIA.S60259

Cole, M. G., McCusker, J., Bellavance, F., Primeau, F. J., Bailey, R. F., Bonnycastle, M. J., & Laplante, J. (2002). Systematic detection and multidisciplinary care of delirium in older medical inpatients: A randomized trial. *Canadian Medical Association Journal*, **167**, 753–759. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/12389836

Elie, M., Rousseau, F., Cole, M., Primeau, F., McCusker, J., & Bellavance, F. (2000). Prevalence and detection of delirium in elderly emergency department patients. *Canadian Medical Association Journal*, **163**, 977–981. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/11068569

Erkinjuntti, T., Sulkava, R., Wikstrom, J., & Autio, L. (1987). Short portable Mental Status Questionnaire as a screening test for dementia and delirium among the elderly. *Journal of the American Geriatrics Society*, **35**, 412–416. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/3571790

Fick, D. M., Agostini, J. V., & Inouye, S. K. (2002). Delirium superimposed on dementia: A systematic review. *Journal of the American Geriatrics Society*, **50**, 1723–1732. doi: 10.1046/j.1532-5415.2002.50468.x

Fick, D. M. (2013). Delirium superimposed on dementia is associated with prolonged length of stay and poor outcomes in hospitalized older adults. *Journal of Hospital Medicine*, **8**, 500–505. doi: 10.1002/jhm.2077

Friedman, S. M., Mendelson, D. A., Bingham, K. W., & McCann, R. M. (2008). Hazards of hospitalization: Residence prior to admission predicts outcomes. *The Gerontologist*, **48**, 537–541. doi: 10.1093/geront/48.4.537

Gaudreau, J. D., Gagnon P., Harel, F., Tremblay, A., & Roy, M. A. (2005). Fast, systematic, and continuous delirium assessment in hospitalized patients: The nursing delirium screening scale. *Journal of Pain and Symptom Management*, **29**, 368–375. doi: 10.1016/j.jpainsymman.2004.07.009

González, M., De Pablo, J., Fuente, E., Valdes, M., Peri, J. M., Nomdedeu, M., & Matrai, S. (2004). Instrument for detection of delirium in general hospitals: Adapation of the Confusion Assessment Method. *Psychosomatics*, **45**, 426–431. doi: 10.1176/appi.psy.45.5.426

Grover, S., & Kate, N. (2012). Assessment scales for delirium: A review. *World Journal of Psychiatry*, **2**, 58–70. doi: 10.5498/wjp.v2.i4.58

Hall, R. J., Meagher, D. J., & MacLullich, A. M. J. (2012). Delirium detection and monitoring outside the ICU. *Best Practice and Research Clinical Anaesthesiology*, **26**, 367–383. doi: 10.1016/j.bpa.2012.07.002

Han, J. H., Wilson, A. W., Vasilevkis, E. E., Shintani, A., Schnelle, J. F., Dittus, R. S., … Wesley Ely, E. (2013). Diagnosing delirium in older emergency department patients: Validity and reliability of the delirium triage screen and the brief confusion assessment method. *Annals of Emergency Medicine*, **62**, 458–465. doi: 10.1016/j.annemergmed.2013.05.003

Inouye, S. K., van Dyck, C. H., Alessi, C. A., Balkin, S., Siegal, A. P., & Howitz, R. I. (1990). Clarifying confusion: The confusion assessment method. A new method for detection of delirium. *Annals of Internal Medicine*, **113**, 941–948. doi: 10.7326/0003-4819-113-12-941

Inouye, S. K. (2003). *The Confusion Assessment Method (CAM): Training Manual and Coding Guides 2003*. Boston: Yale University School of Medicine.

Inouye, S. K. (2006). Delirium in older persons. *New England Journal of Medicine*, **354**, 1157–1165. doi: 10.1056/NEJMra052321

Inouye, S. K., Westendorp, R. G. J., & Saczynski, J. S. (2014). Delirium in elderly people. *Lancet*, **383**, 911–922. doi: 10.1016/S0140-6736(14)60994-6

Kazmierski, J., Kowman, M., Banach, M., Fendler, W., Okonski, P., Banys, A., Jaszweski, R., Sobow, T., & Kloszewska, I. (2008). Clinical utility and use of DSM-IV and ICD-10 criteria and the Memorial Delirium Assessment Scale in establishing a diagnosis of delirium after cardiac surgery. *Psychosomatics*, **49**, 73–76. doi: 10.1176/appi.psy.49.1.73

Kean, J., Trzepacz, P. T., Murray, L. L., Abell, M., & Trexler L. (2010). Initial validation of a brief provisional diagnostic scale for delirium. *Brain Injury*, **24**, 1222–1230. doi: 10.3109/02699052.2010.498008

Koster, S., Hensens, A. G., Oosterveld, F. G., Wijma, A., & van der Palen, J. (2009). The delirium observation screening scale recognizes delirium early after cardiac surgery. *European Journal of Cardiovascular Nursing*, **8**, 309–314. doi: 10.1016/j.ejcnurse.2009.02.006

Lakhan, P., Jones, M., Wilson, A., Courtney, M., Hirdes, J., & Gray, I. C. (2011). A prospective cohort study of geriatric syndromes among older medical patients admitted to acute care hospitals. *Journal of the American Geriatrics Society*, **59**, 2001–2008. doi: 10.1111/j.1532-5415.2011.03663.x

Lawlor, P. G., Nekolaichuk, C., Gagnon, B., Mancini, I. L., Pereira, J. L., & Bruera, E. D. (2000). Clinical utility, factor analysis, and further validation of the memorial delirium assessment scale in patients with advanced cancer: Assessing delirium in advanced cancer. *Cancer*, **88**, 2859–2867. doi: 10.1002/1097-0142(20000615)88:12<2859:: AIDNCR29>3.3.CO;2-K http://dx.doi.org/10.1002/1097-0142(20000615)88:12<2859::AID-CNCR29>3.0.CO;2-T

Leung, J. L. M., Lee, G. T. H., Lam, Y. H., Chan, R. C. C., & Wu, J. Y. M. (2011). The use of the digit span test in screening for cognitive impairment in acute medical inpatients. *International Psychogeriatrics*, **23**, 1569–1574. doi: 10.1017/S1041610211000792

Lin, H. -S., Eeles, E., Pandy, S., Pinsker, D., Brasch, C., & Yerkovich, S. (2015). Screening in delirium: A pilot study of two screening tools, the Simple Query for Easy Evaluation of Consciousness (SQeeC) and Single Question in Delirium (SQiD). *Australasian Journal on Ageing*. doi:10.1111/ajag.12216

Lundstrom, M., Edlund, A., Karlsson, S., Brannstrom, B., Bucht, G., & Gustafson, Y. (2005). A multifactorial intervention program reduced the duration of delirium, length of hospitalization and mortality on delirious patients. *Journal of the American Geriatrics Society*, **53**, 622–628. doi: 10.1111/j.1532-5415.2005.53210.x

MacLullich, A. M. J., Anand, A., Davis, D., Jackson, T., Barugh, A., Hall, R. J., … Cunningham, C. (2013). New horizons in the pathogenesis, assessment and management of delirium. *Age and Ageing*, **42**, 667–674. doi: 10.1093/ageing/aft148 PMid:24067500

MacLullich, A., Ryan, T., & Cash, H. (2011). *The 4A Test: Screening instrument for cognitive impairment and delirium*. Edinburgh, Scotland: University of Edinburgh and NHS Lothian. Retrieved from: http://www.the4at.com Accessed 28 August 2014

Meagher, D. J., Leonard, M., Donnelly, S., Conroy, M., Adamis, D., & Trzepacz, P. T. (2011). A longitudinal study of motor subtypes in delirium: Relationship with other phenomenology, aetiology, medication exposure and prognosis. *Journal of Psychosomatics Research*, **71**, 395–403. doi: 10.1016/j.jpsychores.2011.06.001

Meagher, D. J., Leonard, M., Donnelly, S., Conroy, M., Adamis, D., & Trzepacz, P. T. (2012). A longitudinal study of motor subtypes in delirium: Frequency and stability during episodes. *Journal of Psychosomatic Research*, **72**, 236–241. doi: 10.1016/j.jpsychores.2011.11.013

Monette, J., Galbaud, G., Fung, S. H., Massoud, F., Moride, Y., Arsenault, L., & Afilalo, M. (2001). Evaluation of the Confusion assessment method (CAM) as a tool for delirium in the emergency room. *General Hospital Psychiatry*, **23**, 20–25. doi: 10.1016/S0163-8343(00)00116-X

Morandi, A., McCurley, J., Vasilevskis, E. E., Fick, D. M., Bellilli, G., Lee, P., … Maclullich, A. M. J. (2012). Tools to detect delirium superimposed on dementia: A systematic review. *Journal of the American Geriatrics Society*, **60**, 2005–2013. doi: 10.1111/j.1532-5415.2012.04199.x

Mudge, A. M., Maussen, C., Duncan, J., & Denaro, C. P. (2013). Improving quality of delirium care in a general medical service with established interdisciplinary care: A controlled trial. *Internal Medical Journal*, **43**, 270–277. doi: 10.1111/j.1445-5994.2012.02840.x

Neelon, V. J., Champagne, M. T., Carlson, J. R., & Funk, S. G. (1996). The NEECHAM Confusion Scale: Construction, validation and clinical testing. *Nursing Research*, **45**, 324–330. doi: 10.1097/00006199-199611000-00002

Neufeld, K. J., Hayat, M. J., Coughlin, J. M., Huberman, A. L., Leistikow, N. A., Krumm, S. K., & Needham, D. M. (2011). Evaluation of two intensive care delirium screening tools for non-critically ill hospitalised patients. *Psychosomatics*, **52**, 133–140. doi: 10.1016/j.psym.2010.12.018

Neufeld, K. J., Leoutsakos, J. S., Sieber, F. E., Joshi, D., Wanamaker, B. L., Rios-Robles, J., & Needham, D. M. (2013). Evaluation of two delirium screening tools for detecting post-operative delirium in the elderly. *British Journal of Anaesthesia*, **111**, 612–618. doi: 10.1093/bja/aet167

Pfeiffer, E. (1975). A short portable mental state questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*, **23**; 433–441. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/1159263

Pompei, P., Foreman, M., Cassel, C. K., & Cox, D. (1995). Detecting delirium among hospitalised older patients. *Archives of Internal Medicine*, **155**, 301–307. doi: 10.1001/archinte.155.3.301

Powers, J. S., Doering, T., Gordeon, S., Eden, S., Shintanu, A., & Schnelle, J. (2013). Exploring the utility of ultra-breif delirium assessments on a nonintensive care geriatric population: The GEM study. *The Gerontologist*, **53**, 1051–1055. doi: 10.1093/geront/gns161

Rapp, C. G., Wakefield, B., Kundrat, M., Mentes, J., Tripp-Reimer, T., & Culp, K. (2000). Acute confusion assessment instruments: Clinical versus research usability. *Applied Nursing Research*, **13**, 37–45. doi: 10.1016/S0897-1897(00)80017-8

Radtke, F. M., Franck, M., Scneider, M., Luetz, A., Seeling, M., Heinz, A., … Spies, C. D. (2008). Comparison of three scores to screen for delirium in the recovery room. *British Journal of Anaesthesia*, **101**, 338–343. doi: 10.1093/bja/aen193

Radtke, F. M., Franck, M., Schust, S., Boehme, L., Pascher, A., Bail, H. J., … Spies, C. D. (2010). A comparison of three scores to screen for delirium on the surgical ward. *World Journal of Surgery*, **34**, 487–494. doi: 0.1007/s00268-009-0376-9

Rockwood, K., Goodman, J., Flynn, M., & Stolee, P. (1996). Cross validation of the delirium rating scale in older patients. *Journal of the American Geriatrics Society*, **44**, 839–842. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/8675936

Rosen, J., Sweet, R. A., Mulsant, B. H., Rifai, A. H., Pasternak, R., & Zubenko, G. S. (1994). The delirium rating scale in a psychogeriatric inpatient setting. *The Journal of Neuropsychiatry and Clinical Neurosciences*, **6**, 30–35. doi: 10.1176/jnp.6.1.30

Ryan, K., Leonard, M., Guerin, S., Donnelly, S., Conroy, M., & Meagher, D. J. (2009). Validation of the confusion assessment method in the palliative care setting. *Palliative Medicine*, **23**, 40–45. doi: 10.1177/0269216308099210

Salih, S. A., Paul, S., Klein, K., Lakhan, P., & Gray, L. (2012). Screening for delirium within the interRAI acute care assessment system. *Journal of Nutrition, Health & Aging*, **16**, 695–700. doi: 10.1007/s12603-012-0074-4

Sands, M. B., Danotc, B. P., Hartshom, A., Ryan, C. J., & Lujic, S. (2010). Single Question in Delirium (SQiD): Testing its efficacy against psychiatrist interview, the Confusion Assessment Method and the Memorial Delirium Assessment Scale. *Palliative Medicine*, **24**, 561–565. doi: 10.1177/0269216310371556

Schuurmans, M. J., Shortridge-Bagett, L. M., & Duursma, S. A. (2003). The delirium observation screening scale: A screening instrument for delirium. *Research and Theory for Nursing Practice*, **17**, 31–47. doi: 10.1891/rtnp.17.1.31.53169

Siddiqi, N., House, A., & Holmes, J. (2006). Occurrence and outcome of delirium in medical in-patients: A systematic literature review. *Age and Ageing*, **35**, 350–364. doi: 10.1093/ageing/afl005

Spiller, J. A., & Keen, J. C. (2006). Hypoactive delirium: Assessing the extent of the problem for inpatient specialist palliative care. *Palliative Medicine*, **20**, 17–23. doi: 10.1191/0269216306pm1097oa

Strub, R. L., & Black, F. W. (2000). *The mental state evaluation in neurology* (4th ed.). Philadelphia: FA Davis Company Publishers. ISBN-13: 978-0803604278 10.5498/wjp.v2.i4.58

Trzepacz, P. T. (1994). A review of delirium assessment instruments. *General Hospital Psychiatry*, **16**, 397–405. doi: 10.1016/0163-8343(94)90115-5

Trzepacz, P. T., Baker, R. W., & Greenhouse, J. (1988). A symptom rating scale for delirium. *Psychiatry Research*, **23**, 89–97. doi: 10.1016/0165-1781(88)90037-6

Trzepacz, P. T., Mittal, D., Torres, R., Kanary, K., Norton, J., & Jimerson, N. (2001). Validation of the Delirium Rating Scale-revised-98: Comparison with the delirium rating scale and the cognitive test for delirium. *Journal of Neuropsychiatry & Clinical Neurosciences*, **13**, 229–242. doi: 10.1176/jnp.13.2.229

Vermeersch, P. E. (1990). The clinical assessment of confusion-A. *Applied Nursing Research*, **3**, 128–133. doi: 10.1016/S0897-1897(05)80132–6

Voyer, P., Cole, M. C., McCusker, J., St-Jacques, S., & Laplante, J. (2008). Accuracy of nurse documentation of delirium symptoms in medical charts. *International Journal of Nursing Practice*, **14**, 165–177. doi: 10.1111/j.1440-172X.2008.00681.x

Wand, A. P. F., Thoo, W., Ting, V., Baker, J., Sciuriaga, H., & Hunt, G. E. (2014). A multifaceted educational intervention to prevent delirium in older inpatients: A before and after study. *International Journal of Nursing Studies*, **51**, 974–982. doi: 10.1016/j.ijnurstu.2013.11.005

Wand, A. P. F., Thoo, W., Ting, V., Baker, J., Sciuriaga, H., & Hunt, G. E. (2013). Identification rates of delirium in elderly medical inpatients from diverse language groups. *Geriatric Nursing*, **34**, 355–360. doi: 10.1016/j.gerinurse.2013.05.004

Wei, L. A., Fearing, M. A., Sternberg, E. J., & Inouye, S. K. (2008). The confusion assessment method: A systematic review of current usage. *Journal of the American Geriatrics Society*, **56**, 823–830. doi: 10.1111/j.1532-5415.2008.01674.x

Whittamore, K. H., Goldberg, S. E., Gladman, J. R., Bradshaw, L. E., Jones, R. G., & Harwood, R. H. (2014). The diagnosis, prevalence and outcome of delirium in a cohort of older people with mental health problems on general hospital wards. *International Journal of Geriatric Psychiatry*, **29**, 32–40. doi: 10.1002/gps.3961

Wong, C. L., Holroyd-Leduc, J., Simel, D. L., & Straus, S. E. (2010). Does this patient have delirium? Value of bedside instruments. *Journal of the American Medical Association*, **304**, 779–786. doi: 10.1001/jama.2010.1182

World Health Organization (WHO). (1992). *Mental disorders: Glossary and guide to their classification in accordance with the tenth revision of the international classification of diseases*. Geneva: WHO 1992. Retrieved from: http://www.who.int/iris/handle/10665/62413#sthash.7TYLs50Y.dpuf

Yates, C., Stanley, N., Cerejeira, J. M., Jay, R., & Mukaetova-Ladinska, E. B. (2009). Screening instruments for delirium in older people with an acute medical illness. *Age and Ageing*, **38**, 235–237. doi: 10.1093/ageing/afn285