Aging & Mental Health
Vol. 14, No. 5, July 2010, 543–555

Routledge
Taylor & Francis Group

Delirium scales: A review of current evidence

Dimitrios Adamis[a]*, Naveen Sharma[b], Paul J.P. Whelan[c] and Alastair J.D. Macdonald[d]

[a]*Research and Academic Institute of Athens, 27 Themistokleous Street & Akadimias, Athens, 106 77 Greece;* [b]*South London and Maudsley Foundation NHS Trust, General Adult and Old Age Psychiatry, Ladywell Unit, University Hospital Lewisham, London, SE13 6TD, UK;* [c]*Central and North West London NHS Foundation Trust, 40 Hanson St. Latimer House, London, W1W 6UL, UK;* [d]*Institute of Psychiatry, Old Age Psychiatry, David Goldberg Building, de Crespigny Park, London, SE5 8AP, UK*

(*Received 14 May 2009; final version received 14 September 2009*)

Objectives: Delirium is a common neuropsychiatric condition with many adverse outcomes in elderly populations including death. Despite this, it is often misdiagnosed and mistreated. A number of scales can be used to detect delirium. We review scales that have been used in delirium studies and report their psychometric properties.
Method: An extensive MEDLINE database search and subsequent examination of reference lists was conducted to identify the various delirium scales that have been designed, primarily for use in the elderly.
Results: Twenty-four scales were identified. Delirium instruments differed according to the classification system they were based on, length of time to administer, the rater and whether they were screening scales or measured symptom severity. The psychometric properties of each scale is reported.
Conclusion: A large number of scales exist, but not all are properly evaluated in terms of psychometric properties, and there is not unanimity about which scale is the best. However, a small number of scales may be considered already to be robust and useable: the CAM, the DRS, the MDAS and the NEECHAM.

Keywords: delirium; scales; measurement; instruments; neuropsychological tests

Introduction

Delirium is a severe psychiatric syndrome characterized by an acute change in consciousness and cognitive function caused by an underlying medical condition. It affects up to 30% (Adamis, Treloar, Martin, & Macdonald, 2006a) of elderly general hospital patients. Despite its prevalence the diagnosis of delirium is often missed.

Rating scales can be helpful aids in detecting delirium. Since the DSM and ICD criteria were widely accepted for clinical purposes, the emphasis has been shifted to more sophisticated diagnostic tools for research or clinical purposes, in order to maximize diagnosis and to measure the severity of delirium. However, different definitions and concepts of delirium lead to a different selection of populations under investigation and, inevitably, to different results.

In addition, the question of whether to use a categorical or dimensional approach to diagnosis is pertinent. According to Schmidt, Kotov, and Joiner (2004) the dichotomous nature of diagnosis is arbitrary especially when certain forms of psychopathology are best viewed as dimensional phenomena. Such dimensional symptoms in delirium might include psychomotor activity, consciousness, attention, speech disturbances and, in medical inpatients, the sleep-wake cycle. Indeed, investigators have commented on the weakness of a categorical approach to the diagnosis of delirium. Levkoff et al. (1996) in their seminal work suggested that delirium represents a spectrum of neurobehavioral impairment and they used the term 'Subsyndromal Delirium' (SDD) to describe those who did not meet the full criteria of delirium but who had some delirium symptoms. Patients with SSD fall on a continuum between those with DSM-defined delirium and those with no symptoms of delirium.

There are a number of specific scales for delirium and they represent a variety of methods; some operationalize either DSM-III or DSM-IIIR criteria whereas others are independent of a particular DSM version or have not employed any psychiatric diagnostic criteria (Trzepacz, 1994). Some scales are dichotomous, e.g. the Confusion Assessment Method (CAM), while some allow a dimensional approach which measures what is often called 'severity', e.g. the Delirium Rating Scale (DRS). Severity scales can also be used as categorical ones if a cut-off point is applied. Similarly, categorical scales such as the CAM have been used as dimensional scales after rating each item.

This article reviews recent developments with regard to psychometric qualities, measurement goal, content and rating procedures of delirium scales in clinical practice and research. The aim was to identify such scales and to review their original validation studies and subsequent replication studies. As our main target audience is primarily old age psychiatrists and geriatricians, the review was restricted to scales used specifically for detecting delirium in the elderly

*Corresponding author. Email: dimaadamis@yahoo.com

ISSN 1360–7863 print/ISSN 1364–6915 online
© 2010 Taylor & Francis
DOI: 10.1080/13607860903421011
http://www.informaworld.com

settings they commonly work in. Scales that were developed for use in different populations, such as those used in Intensive Care Unit (ICU) or AIDS-care settings, were not included unless they were later validated for geriatric-specific use.

Methods

A comprehensive search using the MEDLINE database and subsequent examination of reference lists was conducted to identify the various delirium scales in use. The keywords 'delirium', 'acute confusion', 'acute confusion state' and 'acute confusional state' were combined with the terms 'questionnaires', 'measurement', 'scales', 'psychological tests', 'psychiatric status rating scales', 'sensitivity and specificity', 'psychometrics' and 'mass screening' to identify studies that used delirium scales.

Only primary studies, validation studies and studies assessing psychometric properties were included. Exclusion criteria included studies that: (a) were not written in English (although translations of scales were included where relevant); (b) involved primarily ICU, AIDS or palliative care populations (unless the scale was later validated specifically for use with older people); (c) involved children; (d) assessed scales for alcohol- or drug-related delirium; or (e) did not evaluate scale psychometrics. In addition, the studies that evaluated only the DSM or ICD criteria, neuropsychological batteries and the Mini-Mental State Examination (MMSE) were omitted. The MMSE was not included as it is a measurement of cognitive function rather than delirium, although data from the studies of the various delirium scales that report comparisons with the MMSE were reported.

The abstracts of the 858 articles found were screened and 464 were deemed to be of relevance for the purposes of this review. From these articles, 24 delirium scales were identified. These included scales that were diagnostic, measured severity of delirium, or both. The title of each scale was then searched in MEDLINE in an attempt to find additional articles that further evaluated the scale or added extra information relevant to the purposes of this review.

Results

Twenty-four scales have been developed to identify or measure delirium were identified. Each shall be discussed in turn in alphabetical order. In the first part of Table 1, the studies which are not reported in the text are summarized while in the second part studies which compared two or more scales are presented.

Bedside Confusion Scale

This Bedside Confusion Scale (BCS) has been designed and has been evaluated in patients with cancer. It has two items and it takes about two minutes to administer (Stillman & Rybicki, 2000). A total score of 0 is considered as normal, one as borderline and a total of two and above as delirium. Using a cut-off point of two, the sensitivity was one and specificity was 0.85 (compared against the CAM). A later study conducted by Sarhill, Walsh, Nelson, LeGrand, and Davis (2001) claimed to be a validation study for BCS but no actual psychometrics were reported. Although this scale is fast and easy to use as a screening test for delirium, it has not been fully validated.

Clinical Assessment of Confusion-A

The Clinical Assessment of Confusion-A (CAC-A) is based on patients' observed behaviour and was developed to assess a nursing concept of 'confusion' (rather than criterion-defined delirium) (Vermeersch, 1990). It has 25 items, but none of the symptoms rated are specific to delirium (Robertsson, 2002). However, the CAC-A's utility is only partially determined and it cannot distinguish dementia from delirium (Schuurmans, Deschamps, Markham, Shortridge-Baggett, & Duursma, 2003a; Trzepacz, 1994). A revised version has been developed by Rateau (2000).

Clinical Assessment of Confusion-B

The Clinical Assessment of Confusion-B (CAC-B) has seven subscales: cognition; behaviour; motor activity/speech/sensory acuity; orientation; behaviours that threaten the safety of other patients; psychotic-neurotic behaviour and activities of daily living (Vermeersch, 1992). In total, there are 58 items. A score between 66 and 80 indicates possible acute confusion, 81–100 mild confusion, 101–120 moderate confusion and greater than 120 severe confusion. Its internal consistency is 0.95 and interrater reliability is 0.69 in acute care settings (Vermeersch, 1992). However, the aforementioned problem with the CAC-A pertains to the CAC-B too: both measure a broader concept of confusion rather than delirium; thus limiting utility.

Clock Drawing Test

In the Clock Drawing Test (CDT) the patient is required to draw the numbers on a clock face and then insert the hands correctly so that the clock shows a specific time (e.g. 'ten-to-eleven'). The more deranged the clock face the more it indicates the presence of delirium. However, rather conflicting results exist in the literature. Fisher and Flowerdew (1995) found that abnormal CDT scores, but not abnormal MMSE scores, were predictive of the risk of developing postoperative delirium. However, Adamis, Morrison, Treloar, Macdonald, and Martin (2005b) reported that the CDT detects cognitive impairment but not delirium in elderly medical inpatients.

Table 1. Summary of studies witch examine psychometric properties for delirium scales.

Scale	Author	Type of study	Setting/Population	Raters	Items and time	Sensitivity	Specificity	Validity	Reliability
CAM	Inouye et al. (1990)	Primary	General medicine wards (two sites) and in an outpatient geriatric assessment unit. $n=$ 56, 26 delirious 30 non-delirious.	Researchers	9 items <5 min	Site 1: 100% Site 2: 94%	Site 1: 95% Site 2: 90%	With MMSE $k=0.64$, with story recall $k=0.59$, VAS-AC $k=0.82$, with digit span test $k=0.66$; PPV (site 1: 91%, site 2: 94%); NPV (site 1: 100%, site 2: 90%). Validated against psychiatrists' diagnoses	IRR: $\kappa=0.81$–1
CAM	Zou et al. (1998)	Validation	Emergency department $n=87$	Nurses		89%	100%	κ coefficient 0.86 between CAM and consensus diagnosis	
CAM	Laurila et al. (2002)	Validation	Two acute geriatric wards $n=81$	Geriatrician		Rates: 81–86%	Rates: 63–84%	DSM-III, DSM-III-R, DSM-IV, and ICD-10 criteria of delirium used as reference standards. PPV rates of 24–76%, NPV rates of 87–96%, PLR rates of 2.18–4.98, NLR: rates of 0.21–0.32	
CSE	Robertsson (1997)	Primary	Delirious elderly +/− dementia: nursing home, orthopaedic and psychogeratric ward	Nurses/doctors/ psychologists	22 items			Construct validity – CSE against global rating by psychiatrists correlation co-efficient $r=0.79$; Internal validity – mean $r=0.69$	weighted κ 0.38–0.93
DI	McCusker et al. (1998)	Primary	Medical inpatients >65 years with known diagnosis of delirium (diagnosed with CAM)	Psychiatrists, Research assistants	7			Criterion validity correlation with DRS (spearman correlation coefficient, r) – 0.84; Concstruct validity – convergent validity $r=-0.60, -0.70$ and discriminant validity $r=0.26, -0.42$	IRR: 0.78 between research assistants IRR: 0.88 between researchers and psychiatrists

(*continued*)

Table 1. Continued.

Scale	Author	Type of study	Setting/Population	Raters	Items and time	Sensitivity	Specificity	Validity	Reliability
DI	McCusker et al. (2004)	Validation	Medical admissions ≥ 65 years – assessments repeated at 8 weeks after discharge and 6 and 12 months after admission	Research assistants, Research nurses				External validity – correlation coefficient r for correlation with BI, APS and CSI – 0.43, 0.17 and 0.36, respectively	Intraclass correlation coefficient of inter-rater reliability – 0.98; Cronbach's a – 0.74
DOS	Schuurmans et al. (2003a,b)	Primary	Geriatric medicine inpatients = 82 and elderly hip fracture inpatients = 92	Nurses	25			Concurrent validity as compared with CAM for hip fracture group – 0.63 ($p \leq 0.001$)	Internal consistency a – 0.93; a – 0.96;
DRS	Trzepacz et al. (1988)	Primary	Liaison service patient with medical or surgical problems. delirium, 20; dementia, 9; schizophrenia, 9; other medically ill, 9; Total = 47	Two psychiatrists	10 items			Items correlated with each other (r: 0.40–0.66) $r = -0.43$ with MMSE, $r = 0.66$ with Trailmaking B	IRR: 0.97 (ICC)
DRS	Rosen et al. (1994)	Validation	Inpatient psychogeriatric unit ($n = 791$, 70 had delirium)	Research clinicians		Cut-off of 10 gave sensitivity of 94.3%	82%	False negatives in the delirium group were low (6%) but false positives in the non-delirium but with dementia group were high (32%)	IRR 0.69–0.99
DRS	Rockwood et al. (1996)	Validation	Geriatric medicine $n = 70$ and geriatric psychiatry consultation services $n = 34$	Physicians		Cut-off point of 10: 82% Cut-off point of 8: 90%	Cut-off point of 10: 94% Cut-off point of 8: 82%	Construct validity comparison with MMSE $r = -0.78$	Cronbach's $a =$ 90%, IRR: 91% (κ)
DRS-R-98	Andrew et al. (2009)	Validation	Geriatric medicine $n = 147$	Physicians		0.56	0.82		Pearson's $r = 0.93$, ICC = 0.92
DSS	Bettin (1998)	Primary	Older general hospital inpatients with delirium and without dementia	Physicans and trained non-physicians. Study: raters were trained researchers	Max. score 59 10 min			r between DSS and expert ratings – −0.497, −0.436 and 0.523 at 1–12 h, 24–42 h and 2–14 days post-enrolment – at $p = 0.016$	Correlation coefficient for research assistant inter-rater reliability – 0.99

NEECHAM	Neelon et al. (1996)	Primary	Two samples of elderly medical inpatients ($n = 168$ and 258)	Nurses and researchers	< 10 min			High correlation with MMSE ($r = 0.87$) and moderate with the sum of DSM-III-R diagnosis ($r = 0.54$–0.7)	High internal consistancy (Cronbach's $\alpha = 0.9$) and IRR ($r = 0.91$)
NEECHAM	Jagmin (1998)	Validation	Hip surgery patients, $n = 70$	Nurses				NEECHAM vs. MMSE $r = 0.65$. Morning: NEECHAM vs. MMSE $r = 0.83$	
Nu-DESC	Gaudreau et al. (2005)	Primary	Hemato-oncology/ internal medicine $n = 146$	Nurses	5 items 1 min	85.70%	0.86.8%	Nu-DESC were significantly correlated with the DSM-IV and the MDAS DSM-IV $r = 0.71$, MDAS $r = 0.67$	
OBS	Gustafson et al. (1988)	Primary	Swedish elderly population with confusion or dementia ($n = 55$)	Geriatrician and clinical psychologist	16 interview items; 39 observer items; 40 min				IRR = 0.90–0.95
Comparison studies									
Nu-DESC, CAM	Leung, Leung, Leung, and Pan (2008)	Validation/ comparison	Geriatric medicine $n = 100$	Physician, psychiatrist, nurses	Chinese version	CAM: 0.76 Nu-DESC: 0.96	CAM: 1, Nu-DESC: 0.79	$\kappa = 0.521$ between CAM and Nu-DESC	IRR for Nu-DESC: a coefficient of 0.94.
CAM, DRS	Adamis et al. (2005a,b)	Validation/ comparison	Acute admissions to a geriatric unit ($n = 94$)	Psychiatrist				CAM and DRS ($k = 0.77$ for DRS cut-point 12 and 0.92 for cutpoint 10)	
CAM, DRS, MDAS	Grassi et al. (2001)	Primary	Neurological or psychiatric consultation, cancer patients $n = 105$, mean age = 67.7 (SD13.18) years.	Research psychologist	Italian versions of scales	DRS cut-off 12 = 80% DRS cut-off 10 = 95% MDAS = 68%	DRS cut-off 12 = 76, DRS cut-off 10 = 61%, MDAS = 94%	DRS cut-off 12, PPV = 85%, NPV = 69% DRS cut-off 10: PPV = 80%, NPV = 89% MDAS:PPV = 95% NPV = 63%	MDAS and DRS $r = 0.76$; DRS: Cronbach's $\alpha =$ 0.70 MDAS Cronbach's $\alpha =$ 0.89
CAM, MDAS	Marcantonio, Ta, Duthie, and Resnick (2002)	Validation	Acute hip fracture surgery, $n = 122$	Trained research assistant		Cut-off point of 5 in MDAS: 87%. cut-off poit of 9: 88%	Cut-off point of 5 in MDAS: 86%. Cut-off point of 9: 91%		

(*continued*)

Table 1. Continued.

Scale	Author	Type of study	Setting/Population	Raters	Items and time	Sensitivity	Specificity	Validity	Reliability
CAM, Nu-DESC, DDS	Radtke et al. (2008)	Validation/ comparison	Recovery room $n=154$	Research assistants		CAM: 0.43 DDS: 0.14 Nu-DESC: 0.95	CAM: 0.98 DDS: 0.99, Nu-DESC: 0.87		
CAC-A, CAM, DST & V′A′T	Pompei et al. (1995)	Validation/ comparison	432 patients older than 65 years admitted to medical and surgical wards	Clinician investigators		CAM: 46%. CAC-A: 36%	CAM: 92%. CAC-A: 95%	DSM-III-R criteria as standards CAM: (PLR: 5.4 NLR: 0.6), CAC-A: (PLR: 7.8 NLR: 0.7)	
CAC-A, CAC-B, NEECHAM, VAS-AC	Cacchione (2002)	Validation/ comparison	Frail older adults in long-term care	Nurses		CAC-A: 93.1%, CAC-B: 89.7%, NECAAM: 89.7%, VAS-AC: 96.6%	CAC-A: 37%, CAC-B: 76.1%, NECAAM: 69.6%, VAS-AC: 80.5%	Coefficient *a* s: CAC-A: 0.82; CAC-B: 0.86; and NEECHAM: 0.80.	IRR:CAC-B: 0.90 NEECHAM: 0.87 VAS-AC: 0.80
DRS-R-98, DRS and CTD	Trzepacz et al. (2001)	Validation/ comparison	Adults from variety of settings and diagnosis – delirium, dementia, schizophrenia, depression and other psychiatric patients	Psychiatrists				Correlation between CTD & DRS: $r=-0.41$, $p<0.005$; Correlation between CTD and DRS-R-98: $r=-0.63$m, $p<0.001$	

Notes: CAC-A: Clinical Assessment of Confusion – A; CAC-B: Clinical Assessment of Confusion – B; CAM: Confusion Assessment Method; CTD: Cognitive Test for Delirium; CSE: Confusion State Evaluation; DRS: Delirium Rating Scale; DRS-R-98: Delirium Rating Scale-Revised-98; DAS: Delirium Assessment Scale; DDS: Delirium Detection Score (used in ICU); Delirium Index; DST: Digital Span Test; DOSS: Delirium Observation Severity Scale; DSS: Delirium Severity Scale; DOM: Delirium-O-Meter; MCV-NDRS: MDAS: Memorial Delirium Assessment Scale; NEECHAM CS: NEECHAM Confusion Scale; Nu-DESC: Nursing Delirium Screening Scale; OBS: Organic Brain Syndrome; VAS-AC: Visual Analog Scale for Acute Confusion; V′A′T: Visual Analog Test; APS: Acute Physiology Score; BI: Barthel Index; CSI: Clinical Severity of Illness Scale; IRR: Inter-rater reliability; NPV: Negative Predictive Value; PPV: Positive Predictive Value; PLR: Positive Likelihood Ratio; NLR: Negative Likelihood Ratio.

Cognitive Test for Delirium

The Cognitive Test for Delirium (CTD) was designed to identify delirium in ICUs (Hart et al., 1996). It measures five cognitive domains (orientation, attention, memory, comprehension and vigilance), each of which is scored 0–6 in two-point increments, except for comprehension which is scored in single-point increments. Total scores range between 0 and 30, with higher scores indicating better cognitive function. Because of the population this scale was developed for use in non-verbal modes of communication (visual and auditory) are assessed in determining delirium, but the CTD has also been validated for use in non-ICU settings (e.g. Meagher et al., 2007; Trzepacz et al., 2001).

Confusion Assessment Method

The CAM is a screening instrument that detects delirium for use by non-psychiatrist clinicians (Inouye et al., 1990). It was originally validated against the diagnosis made by a psychiatrist using DSM-IIIR criteria (Inouye et al., 1990). It consists of nine operationalized DSM-IIIR criteria. The four 'cardinal' criteria are acute onset and fluctuating course, inattention, disorganized thinking and altered level of consciousness. These contribute to an algorithm for diagnosis of delirium. The presence of both first and second and either the third or fourth criteria are compatible with a diagnosis of delirium. Among older acute medical patients, the CAM used by trained nurses had a sensitivity of 0.68 and specificity of 0.97 (Rockwood et al., 1994) while Pompei, Foreman, Cassel, Alessi, and Cox (1995) found an overall specificity of 92% and sensitivity of 46%. Agreement of CAM was closer with the DSM-IV than with the DSM-IIIR, from which it had been originally devised (Laurila, Pitkala, Strandberg, & Tilvis, 2002). The CAM is a rapidly and easily administered instrument and has been used in various settings for clinical and research purposes. However, appropriate training and use of other cognitive tests, e.g. the MMSE, or thorough patient assessments is necessary to enhance its sensitivity and specificity. The CAM has been translated and validated for use in various non-English speaking populations.

Confusion Rating Scale

The Confusion Rating Scale (CRS) is designed for nurses to detect 'confusion' (Williams, 1991). In clinical settings, it needs to be accompanied by a psychiatric examination and, in research, with a more comprehensive scale for assessing delirium (Trzepacz, 1994). The CRS was subsequently modified in the development of the Nursing Delirium Screening Scale (see below).

Confusion State Evaluation

The Confusion State Evaluation (CSE) was designed to evaluate the efficacy of treatment in delirious people, especially the elderly (Robertsson, Karlsson, Styrud, & Gottfries, 1997). It has 22 items, 12 of which were deemed as core symptoms (disorientation to person, time, space and situation, thought, and memory disturbances, inability to concentrate, distractibility, perseveration, impaired contact, paranoid delusions and hallucinations). The sum of these 12 items represents a total 'confusion score'. Despite good psychometrics (Table 1), this scale has not been evaluated in other studies and, perhaps, has not been used again since its initial validation.

Delirium Assessment Scale

The Delirium Assessment Scale (DAS) was designed to measure the severity of delirium using DSM-III operationalised criteria (O'Keeffe, 1994). Its sensitivity is reported as 0.83–0.88 and specificity 0.79–0.88 using three different raters. No clear cut-off point was reported for delirium. In addition, the DAS has not been evaluated using repeated measures and is not able to differentiate between delirium and dementia (Schuurmans et al., 2003a).

Delirium Index

The Delirium Index (DI) was developed to measure changes in severity of symptoms in delirious patients and was designed to be used in conjunction with the MMSE (McCusker, Cole, Bellavance, & Primeau, 1998). It has seven items that are operationalized from the CAM. In the original study, subjects were initially diagnosed with delirium by the CAM. The reliability, validity, and responsiveness of the scale were further evaluated by the scale development team (McCusker, Cole, Dendukuri, & Belzile, 2004) (Table 1). The DI had good to excellent 'responsiveness' (to change over time) at six months (effect sizes −0.49 and −0.71; standardized response mean −0.45 and −0.81) and moderate responsiveness at 12 months (effect sizes −0.39 and −0.51; standardized response mean −0.34 and −0.51) in delirious patients with and without dementia. Although the DI appears to be a reliable, valid and responsive measure for assessing and monitoring the severity of delirium for clinical and research purposes it lacks popularity; from the 15 studies identified that used the DI, all were conducted by the team that developed the scale.

Delirium Observation Screening Scale

The Delirium Observation Screening (DOS) scale is a 25-item scale based on DSM-IV criteria that was developed for screening for delirium by nursing observations (Schuurmans, Shortridge-Baggett, & Duursma, 2003b). It has acceptable predictive validity

against the DSM-IV, good correlation with the MMSE, and high internal consistency. The scale was subsequently shortened to 13 items and can be rated in less than five minutes (Schuurmans, Donders, Shortridge-Bagget, & Duursma, 2002). A score of zero is defined as 'normal behaviour', meaning the absence of behavioural alterations. The highest total score is 13 and the cut-off point is 3. Three or more points indicate the presence of delirium. The sensitivity of the scale was 0.89 and specificity was 0.88.

Delirium Rating Scale

The DRS was developed as a complement to the DSM-III criteria to measure the severity of delirium (Trzepacz, Baker, & Greenhouse, 1988). It is a 10-item rating scale, intended to be completed by a clinician with psychiatric training. Each item can be scored from zero to four points depending on the item. The maximum score for the scale is 32 points. A total scale score of 12 or more is compatible with a diagnosis of delirium (Trzepacz et al., 1988). Sensitivity and specificity are substantial (Trzepacz, 1999) but also depend on the cut-off point used for diagnosis of delirium. Using the cut-point of 10, the sensitivity is 0.82 and specificity is 0.94 (Rosen et al., 1994). With a cut-off point of 8, the sensitivity of the DRS rose to 0.90 with specificity falling to 0.82 (Rockwood, Goodman, Flynn, & Stolee, 1996; Rosen et al., 1994; Trzepacz, 1999). Like the CAM, various non-English language versions of the DRS exist.

In a head-to-head comparison of DRS and CAM, a high level of agreement was found between the two and this was best when a cut-off point of 10 was chosen for the DRS (Adamis et al., 2005a).

Delirium Rating Scale-Revised-98

The DRS was revised to address a number of shortcomings: its inability to distinguish between hypoactive and hyperactive subtypes of delirium, the lack of separate scoring item for attention deficit, and a lack of clarity about 'clouding of consciousness' (Trzepacz et al., 2001). The Delirium Rating Scale-Revised-98 (DRS-R-98) rectifies these problems by incorporating 16 clinician-rated items (13 for severity and three for diagnosis), which have greater phenomenological preciseness than its predecessor. It successfully differentiates delirium from dementia, depression and schizophrenia. The receiver operator characteristic curve analysis indicated optimal cut-points of 15.25 (sensitivity 0.92, specificity 0.86) or 17.75 (sensitivity 0.92, specificity 0.95). The DRS-R-98 correlated strongly with the DRS (Pearson's $r = 0.83$) but less so with the Clinical Global Impression ($r = 0.61$). It had both good interrater reliability (Cronbach's $a = 0.87$) and internal consistency. The findings of original validation study of the DRS-R-98 have been replicated in subsequent work, and Spanish and Dutch versions have validated.

Delirium Scale

The Delirium Scale (D-Scale) has 58 items in 13 categories, each of which scored on a four-point scale (Lowy, Engelsmann, & Lipowski, 1973). It correlates well with the Brief Psychiatric Rating Scale ($r = 0.68$) and MMSE ($r = 0.83$). However, it is lengthy for use in clinical settings and possibly inappropriate for research purposes due to its unknown psychometric properties (Smith, Breitbart, & Platt, 1995).

Delirium Severity Scale

The Delirium Severity Scale (DSS) is a 10 min assessment of delirium consisting of forward digit span and cognitive similarities (Bettin et al., 1998). It was designed as a scale to measure severity and change. The usefulness of this scale is limited because as it has only been validated in delirious populations without dementia or severe physical illness (Robertsson, 2002).

Delirium Symptom Interview

The Delirium Symptom Interview (DSI) is a structured interview that assesses seven domains from the DSM-III criteria (Albert et al., 1992). It has 107 items of which 63 are interview questions and the remainder observations. Its sensitivity was 0.90 and the specificity was 0.80; the interobserver agreement between two lay interviewers ($k = 0.90$) and between two physicians ($k = 0.92$) was excellent (Albert et al., 1992). The authors recommend that the DSI could be used by lay interviewers to reliably assess the symptoms of delirium.

Delirium-O-Meter

The Delirium-O-Meter is a behavioural observation scale which has 12 items: sustained attention; shifting of attention; orientation; disturbance of consciousness; apathy; hypokinesia/psychomotor retardation; incoherence; fluctuating functioning (diurnal variation/sleep-wake cycle); restlessness (psychomotor agitation); delusions; hallucinations; and anxiety/fear. Each item is scored on a four-point scale (0 = absent; 1 = mild disturbance; 2 = moderate; 3 = severe) (de Jonghe, Kalisvaart, Timmers, Kat, & Jackson, 2005). The scoring of each severity level is described in detail for all items. The total score ranges from 0 to 36. A high degree of correlation with DOS, DRS R98, MMSE and also very good interater agreement and internal consistency were reported (ICC 0.84–0.91 Cronbach's $a = 0.87$). No other study has been identified that further validates this scale.

Memorial Delirium Assessment Scale

The Memorial Delirium Assessment Scale (MDAS) was originally designed for repeated assessment of delirium symptoms over a relatively brief time in cancer patients on opioid infusion (Breitbart et al., 1997). It has 10 items, three of which measure cognition. The primary study was conducted with 33 cancer patients, 17 of whom met the DSM-IIIR/DSM-IV criteria for delirium. It had a high interrater reliability ($r = 0.92$) and internal consistency (Cronbach's $a = 0.91$). Because the MDAS does not include some important phenomenological features of delirium, such as variability of the symptoms or acuteness of onset, caution is needed when using it to diagnose or screen for delirium (Trzepacz, 1994). It has also been suggested that the MDAS is best used to quantify the severity of delirium after initial diagnosis is made by other scales (Roth-Roemer, Fann, & Syrjala, 1997).

NEECHAM Confusion Scale

The NEECHAM confusion scale was developed for use as a rapid bedside assessment by nurses (Neelon, Champagne, Carlson, & Funk, 1996). It has three subscales and is scored between 0 and 30. A score of 30 indicates normal function and 0 severe confusion with a range of 25–26 as 'at risk for confusion'. The sensitivity and specificity of the NEECHAM scale ranges from 30 to 95% and 78 to 92%, respectively (Cacchione, 2002) (Table 1). It incorporates physiological parameters, but it has been argued that this makes the scale unsuitable and adds nothing to its ability to measure severity (Robertsson, 2002; Smith et al., 1995; Trzepacz, 1994). It has been criticized that the NEECHAM scale measures specifically a construct called 'acute confusion' rather than criterion-defined delirium (Rapp et al., 2000). This scale also has been translated into different languages.

Nursing Delirium Screening Scale

The Nursing Delirium Screening Scale (Nu-DESC) is an adaptation and extension of the CRS (Gaudreau, Gagnon, Harel, & Roy, 2005). It has five items: orientation, behaviour, communication, illusions/hallucinations, and psychomotor retardation. Each item is rated from 0 to 2 at each nursing shift, and the total score is based on 24 h observation. The psychometric properties of the scale are reported in Table 1.

Organic Brain Syndrome

The Organic Brain Syndrome (OBS) was developed in Sweden initially for clinical evaluation in dementia (Berggren et al., 1987). A modified version has been used in elderly population (Berggren et al., 1987; Gustafson et al., 1988), and an English language version was published by Jensen, Dehlin, and Gustafson (1993). This consists of two subscales (OBS1 and OBS2) based on operationalizations of DSM-III criteria. The OBS1 (disorientation subscale) consists of a 16-question interview describing the patient's awareness and orientation and takes approximately 10 min to complete. The patient is assessed according to a four-point ordinal scale with a detailed description given for each level (0–3), where 0 indicates a correct response, while 1, 2 and 3 indicate slightly, moderately or completely wrong answers, respectively. The OBS2 (confusion subscale) is a 30 min observation scale comprising 39 clinical items covering the latest seven day period. The severity of the symptoms is ranked in four ordinal scale steps according to their intensity and frequency: a score of 0 indicates lack of any symptoms and a score of 1, 2 or 3 represents occasional, moderate and obviously constant/recurring symptoms, respectively. Both subscales exist in modified versions with the OBS 1 shortened to 12 items and the OBS 2 to 21 items (Berggren et al., 1987).

The OBS has been criticized for lacking sufficient validation (Robertsson, 2002; Smith et al., 1995). However, a recent systematic review of 30 studies shows that the scale is a valid clinical instrument in various clinical settings and clinical conditions, but its clinical utility in delirium needs further investigation (Bjorkelund, Larsson, Gustafson, & Andersson, 2006).

Saskatoon Delirium Checklist

The Saskatoon Delirium Checklist (SDC) is a checklist based on the DSM-III criteria (Miller et al., 1988). It consists of nine symptoms observed during a 24 h period and one question regarding the physical causes of delirium. Each item is rated from 0 to 4 with a total score 40. The SDC developers recommend caution when using, because whilst the scale is sufficiently sensitive to diagnose delirium, further studies are needed to establish its validity and other psychometric properties. Therefore, the SDS's utility as a delirium scale is uncertain (Smith et al., 1995; Trzepacz, 1994).

The Visual Analog Scale for Acute Confusion

Adapted from the VAS-C (Nagley, 1986), the Visual Analog Scale for Acute Confusion (VASC-AC) uses a 10 cm line upon which the rater estimates the patient's degree of confusion from 'no acute confusion' at the one end to 'severe acute confusion' at the far end (Cacchione, 2002). This scale measures 'confusion' instead of delirium and is very subjective.

Handwriting tests

The Delirium Writing Test, Signature Test

The Delirium Writing Test (Aakerlund & Rosenberg, 1994) and the Abnormal Signature Test (Adamis, Reich, Treloar, Macdonald, & Martin, 2006b) reported a kind of measurement of handwriting with the

Table 2. Comparison of scales.

Scale	Psychometric properties reported	Validation	Number of publication papers	Feasibility	Type of scale	Comments
BCS	–	Yes	3	Easy to use	Screening/diagnostic	Palliative care
CAC-A	+	Yes	5	Easy to use	Screening/diagnostic	Detects 'confusion'
CAC-B	+++	Yes	1	Lengthy	Screening/diagnostic	Detects 'confusion'
CDT	+	Yes	8	Easy to use	Screening	Measures cognition
CTD	+++	Yes	641	Easy to use	Screening/diagnostic	Measures cognition
CAM	+++	Yes	227	Some training	Screening/diagnostic	Recommended
CRS	–	No	5	Easy to use	Screening	Detects 'confusion'
CSE	++	No	2	Easy to use	Severity	
DAS	++	No	0	Easy to use	Severity	No clear cut-off points
DI	+++	Yes	15	Easy to use	Severity	Based on CAM
DOS	+++	Yes	4	Lengthy	Screening	Not very popular yet.
DRS	+++	Yes	89	Easy to use	Diagnostic/severity	Recommended
DRS-R-98	+++	Yes	12	Easy to use,	Diagnostic/severity	Recommended
D-Scale	–	No	0	lengthy	Diagnostic	
DSS	+	No	2	Easy to use	Severity	
DSI	++	No	2	Very lengthy	Diagnostic	
DOM	++	No	1	Easy to use	Severity	
MDAS	+++	Yes	33	Easy to use	Severity	Recommended
NEECHAM	+++	Yes	33	Easy to use	Screening/diagnostic/ severity	Recommended
Nu-DESC	++	Partial	6	Based in 24h obs	Screening	
OBS	+++	Yes	15	lengthy	Screening/diagnostic	Designed to detect dementia
SDC	–	No	2	Based in 24h obs	Diagnostic	
VAS-AC	+	Yes	1	Very easy	Diagnostic	Subjective, detects 'confusion'
Handwriting tests	+	No	2	Very easy	Screening	Subjective

Notes: −: No psychometric properties reported, +: few psychometric properties reported, ++: Some psychometric properties reported, +++: All or nearly all the psychometric properties reported.
BSC: Bedside Confusion Scale; CAC-A: Clinical Assessment of Confusion – A; CAC-B: Clinical Assessment of Confusion – B; CDT: Clock Drawing Test; CTD: Cognitive Test for Delirium; CAM: Confusion Assessment Method; CRS: Confusion Rating Scale; CSE: Confusion State Evaluation; DAS: Delirium Assessment Scale DI: Delirium Index; DOS: Delirium Observation Severity; DRS: Delirium Rating Scale; DRS-R-98: Delirium Rating Scale-Revised-98; D-Scale: Delirium Scale; DSS: Delirium Severity Scale; DSI: Delirium Symptom Interview; DOM: Delirium-O-Meter; MDAS: Memorial Delirium Assessment Scale; NEECHAM: NEECHAM Confusion Scale; Nu-DESC: Nursing Delirium Screening Scale; OBS: Organic Brain Syndrome; VAS-AC: Visual Analog Scale for Acute Confusion.

purpose of identifying delirium. However, neither study has yet been replicated.

Tests of global cognitive tests (such as the MMSE)

Cognitive tests – not part of this review – have been used for many years for evaluation of delirium, alone or in conjunction with other delirium scales or in repeated assessments (O'Keefe, Mulkerrin, Nayeem, Varughese, & Pillay, 2005). Indeed, cognitive deficits are common in delirium but not specific to it, occurring in a number of other conditions such as dementia. For a recent review, see Woodford and George (2007).

Comparison of scales

A number of studies compared two or more different scales in their psychometric properties. However, most of the studies use the one scale as a 'gold standard' to validate other scales and some other studies evaluate agreement between scales. Those studies and a summary of their findings are presented in the second part of Table 1.

Further, in Table 2, we compared all the reviewed scales regarding their reported psychometric properties, their validation, the popularity of the scale, through an indirect measurement of the number of published papers which used it, in MEDLINE, their feasibility, and if they can be used as a screening/diagnostic or severity measurement.

Discussion

There are a large number of scales designed to detect delirium to choose from. Instruments vary in purpose (e.g. diagnosis vs. symptoms severity), nature of data on which the rating is based (e.g. observation vs. interview), the rater, the number of items and the rating time required. The majority of the most frequently cited instruments, such as the CAM and

the DRS, are based on DSM criteria for delirium. Brief scales with fewer items have inherent problems with lower sensitivity and specificity. Also, a number of shorter instruments are simply checklists and are less suitable for use as repeated measurements or follow-ups. Scales with more items have better accuracy, superior interrater reliability, less variability, and higher sensitivity and specificity.

Many scale development studies employed a heterogeneous group of cognitively disordered elderly, and the large number of different scales available also reflect the use of the less clearly defined, or less meaningful, concept of 'confusion', which is a broader term than delirium and does not necessarily include it. These factors together raise the issue of interrater reliability with some authors suggesting that reported estimations are spuriously high (Bhat & Rockwood, 2005). Moreover, most of the scales in this review are not actually validated, i.e. following the primary development study there have been few further studies replicating the scale's results in a similar population. This is an important point, as the study from which a scale is derived cannot be assumed to be a validation study. Subsequently, many have never been used again since their primary study.

Comparison of scales

For the above reasons, it is difficult to directly compare many of the scales. However, scales like CRS, CSE, DAS, D-Scale, DSS, DSI, DOM, Nu-DESK and SDC have not been validated further, irrespective of whether the psychometric properties in their primary study from where they were derived are good or not. Although have been used in delirium studies and especially the later in ICU delirium, they measure mostly cognition than delirium. Similarly CAC-A and CAC-B diagnose 'confusion' rather than delirium. Scales like VAS-AC and, the two handwriting tests, although very easy to use, are subjective in their interpretation, not validated and their psychometric properties not fully evaluated. BSC has only been used in palliative care until now and its psychometric properties also have not been evaluated. OBS is a lengthy scale designated for dementia, not very popular (except Scandinavia) and not fully validated as delirium scale. DI is in fact a delirium severity CAM, despite its good psychometric properties not being largely used. Similarly DOS has not yet gain in popularity.

It is the authors' views that a small number of scales are not affected by these, above-discussed, methodological weaknesses and are sufficiently validated (Table 2); they are the CAM, the DRS, the DRS-R-98, the MDAS and the NEECHAM. CAM can be used either as diagnostic or screening toll, is easy to administer, although for inexperienced administrators some training is necessary. For research purposes the use of a cognitive test may be required. DRS can be used as a severity scale for repeated measurements but also for diagnostic purposes if a cut-off point be used. Like the CAM, the DRS is necessary to be administered together with a cognitive test. In contrast, the DRS-R-98 does not require a cognitive test. MDAS can be used as a severity scale for repeated measurements over the time but it cannot be used as a screening or diagnostic test and has been recommended to be used after initial diagnosis is made by other scales. NEECHAM is a very popular scale mainly among nurses, easy to administer and can be used mainly as a screening tool (using cut-off points can be a diagnostic or even severity measurement but because it has been suggested that it measures 'acute confusion' rather than criterion-defined, delirium is better to be used for screening purposes).

Choice of scales

Estimates of validity and reliability have been reported for all the scales in this review (Table 1). However, these data do not allow an unequivocal endorsement that a given scale is the best. In choosing a scale, factors other than psychometrics are important, such as administration time constraints, learning effects and, most importantly, the level of rater expertise. Different skill levels and training of the assessor can affect the scoring of delirium scales (Schuurmans, 2003a) and this makes the results of different studies less comparable.

There are a number of different factors that influence the choice of scale:

- *Purpose of the scale*. Different scale properties are required depending on whether it is being used for clinical or research purposes. Certain scales are better screening instruments (e.g. the NEECHAM), some are more suited to making a diagnosis (e.g. the CAM), whilst others are good for measuring symptom severity (e.g. the DRS).
- *Clinical setting*. Different settings require different scales. In this review, primary scales for elderly medical in-patients were reviewed, whereas ICU or alcohol treatment centers will require different scales.
- *Properties of the scale*. General properties (e.g. administration time) and psychometric properties (e.g. validity, reliability) for the given population are pertinent to choice.
- *Popularity of the scale*. Scales that are frequently used in other delirium studies permit inter-scale comparison.
- *Familiarity with the scale*. The previous experience of the clinician or researcher will influence scale choice.
- *Organizational culture*. Within various academic departments, hospitals or research teams there will be a preference towards using a certain scale based on need and the preference of influential individuals with the organization.

Conclusions

There are a large number of delirium scales. Most show promising results but need further research to replicate their findings. Studies are needed in different samples and on a broader range of aspects with regard to reliability and validity. However, a small number of scales may be considered already to be robust and useable. These include the CAM, an excellent diagnostic instrument, the DRS and its revision, and the MDAS, which are good at measuring symptom severity, and the NEECHAM, a quality screening tool.

References

Aakerlund, L.P., & Rosenberg, J. (1994). Writing disturbances: An indicator for postoperative delirium. *International Journal of Psychiatry in Medicine, 24*(3), 245–257.

Adamis, D., Morrison, C., Treloar, A., Macdonald, A.J., & Martin, F.C. (2005b). The performance of the Clock Drawing Test in elderly medical inpatients: Does it have utility in the identification of delirium? *Journal of Geriatric Psychiatry and Neurology, 18*(3), 129–133.

Adamis, D., Reich, S., Treloar, A., Macdonald, A.J., & Martin, F.C. (2006b). Dysgraphia in elderly delirious medical inpatients. *Aging Clinical and Experimental Research, 18*(4), 334–339.

Adamis, D., Treloar, A., MacDonald, A.J., & Martin, F.C. (2005a). Concurrent validity of two instruments (the Confusion Assessment Method and the Delirium Rating Scale) in the detection of delirium among older medical inpatients. *Age and Ageing, 34*(1), 72–75.

Adamis, D., Treloar, A., Martin, F.C., & Macdonald, A.J.D. (2006a). Recovery and outcome of delirium in elderly medical inpatients. *Archives of Gerontology and Geriatrics, 43*(2), 289–298.

Albert, M.S., Levkoff, S.E., Reilly, C., Liptzin, B., Pilgrim, D., Cleary, P.D., et al. (1992). The delirium symptom interview: An interview for the detection of delirium symptoms in hospitalized patients. *Journal of Geriatric Psychiatry and Neurology, 5*(1), 14–21.

Andrew, M.K., Bhat, R., Clarke, B., Freter, S.H., Rockwood, M.R., & Rockwood, K. (2009). Inter-rater reliability of the DRS-R-98 in detecting delirium in frail elderly patients. *Age and Ageing, 38*(2), 241–244.

Berggren, D., Gustafson, Y., Eriksson, B., Bucht, G., Hansson, L.I., Reiz, S., et al. (1987). Postoperative confusion after anesthesia in elderly patients with femoral neck fractures. *Anesthesia & Analgesia, 66*(6), 497–504.

Bettin, K.M., Maletta, G.J., Dysken, M.W., Jilk, K.M., Weldon, D.T., Kuskowski, M., et al. (1998). Measuring delirium severity in older general hospital inpatients without dementia. The Delirium Severity Scale. *American Journal of Geriatric Psychiatry, 6*(4), 296–307.

Bhat, R., & Rockwood, K. (2005). Inter-rater reliability of delirium rating scales. *Neuroepidemiology, 25*(1), 48–52.

Bjorkelund, K.B., Larsson, S., Gustafson, L., & Andersson, E. (2006). The Organic Brain Syndrome (OBS) scale: A systematic review. *International Journal of Geriatric Psychiatry, 21*(3), 210–222.

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*(3), 128–137.

Cacchione, P.Z. (2002). Four acute confusion assessment instruments: Reliability and validity for use in long-term care facilities. *Journal of Gerontological Nursing, 28*(1), 12–19.

Csokasy, J. (1999). Assessment of acute confusion: Use of the NEECHAM Confusion Scale. *Applied Nursing Research, 12*(1), 51–55.

de Jonghe, J.F., Kalisvaart, K.J., Timmers, J.F., Kat, M.G., & Jackson, J.C. (2005). Delirium-O-Meter: A nurses' rating scale for monitoring delirium severity in geriatric patients. *International Journal of Geriatric Psychiatry, 20*(12), 1158–1166.

Fisher, B.W., & Flowerdew, G. (1995). A simple model for predicting postoperative delirium in older patients undergoing elective orthopedic surgery. *Journal of the American Geriatrics Society, 43*(2), 175–178.

Gaudreau, J.D., Gagnon, P., Harel, F., & Roy, M.A. (2005). Impact on delirium detection of using a sensitive instrument integrated into clinical practice. *General Hospital Psychiatry, 27*(3), 194–199.

Grassi, L., Caraceni, A., Beltrami, E., Borreani, C., Zamorani, M., Maltoni, M., et al. (2001). Assessing delirium in cancer patients: The Italian versions of the delirium rating scale and the memorial delirium assessment scale. *Journal of Pain and Symptom Management, 21*(1), 59–68.

Gustafson, Y., Berggren, D., Brannstrom, B., Bucht, G., Norberg, A., Hansson, L.I., et al. (1988). Acute confusional states in elderly patients treated for femoral neck fracture. *Journal of the American Geriatrics Society, 36*(6), 525–530.

Hart, R.P., Levenson, J.L., Sessler, C.N., Best, A.M., Schwartz, S.M., & Rutherford, L.E. (1996). Validation of a cognitive test for delirium in medical ICU patients. *Psychosomatics, 37*(6), 533–546.

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

Jagmin, M.G. (1998). Postoperative mental status in elderly hip surgery patients. *Orthopaedic Nursing, 17*(6), 32–42.

Jensen, E., Dehlin, O., & Gustafson, L. (1993). A comparison between three psychogeriatric scales. *International Journal of Geriatric Psychiatry, 8*(3), 215–229.

Laurila, J.V., Pitkala, K.H., Strandberg, T.E., & Tilvis, R.S. (2002). Confusion assessment method in the diagnostics of delirium among aged hospital patients: Would it serve better in screening than as a diagnostic instrument? *International Journal of Geriatric Psychiatry, 17*(12), 1112–1119.

Leung, J.M., Leung, V.W., Leung, C.M., & Pan, P.C. (2008). Clinical utility and validation of two instruments (the Confusion Assessment Method Algorithm and the Chinese version of Nursing Delirium Screening Scale) to detect delirium in geriatric inpatients. *General Hospital Psychiatry, 30*(2), 171–176.

Levkoff, S., Liptzin, B., Cleary, P., Wetle, T., Evans, D., Rowe, J., et al. (1996). Subsyndromal delirium. *American Journal of Geriatric Psychiatry, 4*(4), 320–329.

Lowy, F.H., Engelsmann, F., & Lipowski, Z.J. (1973). Study of cognitive functioning in a medical population. *Comprehensive Psychiatry, 14*(4), 331–338.

Marcantonio, E., Ta, T., Duthie, E., & Resnick, N.M. (2002). Delirium severity and psychomotor types: Their relationship with outcomes after hip fracture repair. *Journal American Geriatric Society, 50*(5), 850–7.

McCusker, J., Cole, M., Bellavance, F., & Primeau, F. (1998). Reliability and validity of a new measure of severity of delirium. *International Psychogeriatrics, 10*(4), 421–433.

McCusker, J., Cole, M.G., Dendukuri, N., & Belzile, E. (2004). The delirium index, a measure of the severity of delirium: New findings on reliability, validity, and responsiveness. *Journal of the American Geriatrics Society, 52*(10), 1744–1749.

Meagher, D.J., Moran, M., Raju, B., Gibbons, D., Donnelly, S., Saunders, J., et al. (2007). Phenomenology of delirium. Assessment of 100 adult cases using standardised measures. *British Journal of Psychiatry, 190*, 135–141.

Miller, P.S., Richardson, J.S., Jyu, C.A., Lemay, J.S., Hiscock, M., & Keegan, D.L. (1988). Association of low serum anticholinergic levels and cognitive impairment in elderly presurgical patients. *American Journal of Psychiatry, 145*(3), 342–345.

Nagley, S.J. (1986). Predicting and preventing confusion in your patients. *Journal of Gerontological Nursing, 12*(3), 27–31.

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*(6), 324–330.

O'Keeffe, S.T. (1994). Rating the severity of delirium: The delirium assessment scale. *International Journal of Geriatric Psychiatry, 9*, 551–556.

O'Keeffe, S.T., Mulkerrin, E.C., Nayeem, K., Varughese, M., & Pillay, I. (2005). Use of serial Mini-Mental State Examinations to diagnose and monitor delirium in elderly hospital patients. *Journal of the American Geriatrics Society, 53*(5), 867–870.

Pompei, P., Foreman, M., Cassel, C.K., Alessi, C., & Cox, D. (1995). Detecting delirium among hospitalized older patients. *Archives of Internal Medicine, 155*(3), 301–307.

Radtke, F.M., Franck, M., Schneider, M., Luetz, A., Seeling, M., Heinz, A., et al. (2008). Comparison of three scores to screen for delirium in the recovery room. *British Journal of Anaesthesia, 101*(3), 338–343.

Rapp, C.G., Wakefield, B., Kundrat, M., Mentes, J., Tripp-Reimer, T., Culp, K., et al. (2000). Acute confusion assessment instruments: Clinical versus research usability. *Applied Nursing Research, 13*(1), 37–45.

Rateau, M.R. (2000). Confusion and aggression in restrained elderly persons undergoing hip repair surgery. *Applied Nursing Research, 13*(1), 50–54.

Robertsson, B. (2002). The instrumentation of delirium. In J. Lindesay, K. Rockwood, & A. Macdonald (Eds.), *Delirium in old age* (pp. 9–26). Oxford, New York: Oxford University Press.

Robertsson, B., Karlsson, I., Styrud, E., & Gottfries, C.G. (1997). Confusional State Evaluation (CSE): An instrument for measuring severity of delirium in the elderly. *British Journal of Psychiatry, 170*, 565–570.

Rockwood, K., Cosway, S., Stolee, P., Kydd, D., Carver, D., Jarrett, P., et al. (1994). Increasing the recognition of delirium in elderly patients. *Journal of the American Geriatrics Society, 42*(3), 252–256.

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*(7), 839–842.

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. *Journal of Neuropsychiatry and Clinical Neurosciences, 6*(1), 30–35.

Roth-Roemer, S., Fann, J., & Syrjala, K. (1997). The importance of recognizing and measuring delirium. *Journal of Pain and Symptom Management, 13*(3), 125–127.

Sarhill, N., Walsh, D., Nelson, K.A., LeGrand, S., & Davis, M.P. (2001). Assessment of delirium in advanced cancer: The use of the bedside confusion scale. *Am J Hosp Palliat Care, 18*(5), 335–341.

Schmidt, N.B., Kotov, R., & Joiner, T.E. (2004). *Taxometrics: Toward a new diagnostic scheme for psychopathology*. Washington, DC, London: American Psychological Association.

Schuurmans, M., Donders, R., Shortridge-Bagget, L., & Duursma, S. (2002). Delirium case finding: Pilot testing of a new screening scale for nurses. *Journal of the American Geriatrics Society, 50*(supplement A8), S3.

Schuurmans, M.J., Deschamps, P.I., Markham, S.W., Shortridge-Baggett, L.M., & Duursma, S.A. (2003a). The measurement of delirium: Review of scales. *Research and Theory for Nursing Practice, 17*(3), 207–224.

Schuurmans, M.J., Shortridge-Baggett, L.M., & Duursma, S.A. (2003b). The Delirium Observation Screening Scale: A screening instrument for delirium. *Research and Theory for Nursing Practice, 17*(1), 31–50.

Smith, M.J., Breitbart, W.S., & Platt, M.M. (1995). A critique of instruments and methods to detect, diagnose, and rate delirium. *Journal of Pain and Symptom Management, 10*(1), 35–77.

Stillman, M.J., & Rybicki, L.A. (2000). The bedside confusion scale: Development of a portable bedside test for confusion and its application to the palliative medicine population. *Journal of Palliative Medicine, 3*(4), 449–456.

Trzepacz, P.T. (1994). A review of delirium assessment instruments. *General Hospital Psychiatry, 16*(6), 397–405.

Trzepacz, P.T. (1999). The Delirium Rating Scale. Its use in consultation-liaison research. *Psychosomatics, 40*(3), 193–204.

Trzepacz, P.T., Baker, R.W., & Greenhouse, J. (1988). A symptom rating scale for delirium. *Psychiatry Research, 23*(1), 89–97.

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 and Clinical Neurosciences, 13*(2), 229–242.

Vermeersch, P.E. (1990). The clinical assessment of confusion-A. *Applied Nursing Research, 3*(3), 128–133.

Vermeersch, P. (1992). Clinical assessment of confusion. In S. Funk, M. Tornquist, M. Champagne, & R. Wiese (Eds.), *Key aspects of elder care managing falls, incontinence and cognitive impairment* (pp. 251–262). New York: Springer.

Williams, M.A. (1991). Delirium/acute confusional states: Evaluation devices in nursing. *International Psychogeriatrics, 3*(2), 301–308.

Woodford, H.J., & George, J. (2007). Cognitive assessment in the elderly: A review of clinical methods. *Quarterly Journal of Mathematics, 100*(8), 469–484.

Zou, Y., Cole, M.G., Primeau, F.J., McCusker, J., Bellavance, F., & Laplante, J. (1998). Detection and diagnosis of delirium in the elderly: Psychiatrist diagnosis, confusion assessment method, or consensus diagnosis? *International Psychogeriatrics, 10*(3), 303–308.

Copyright of Aging & Mental Health is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.