

Does This Patient Have Delirium?

Value of Bedside Instruments

Camilla L. Wong, MD, MHSc, FRCPC

Jayna Holroyd-Leduc, MD, FRCPC

David L. Simel, MD, MHS

Sharon E. Straus, MD, MSc, FRCPC

CLINICAL SCENARIO

An 85-year-old man with non-small cell lung cancer was admitted yesterday to the oncology ward for treatment of pneumonia. He started antibiotics and supplemental oxygen. Prior to hospitalization, he lived independently in the community. An overnight nursing report indicates that he did not sleep and was pacing up and down the ward. During morning rounds, he was easily startled and noted to be picking at the bed sheets. The patient's daughter says, "This is not my father. Right now, he's in his own little world. He never behaves this way!" What instruments could the nurse and physician use to determine if these changes and observations are indicative of a delirium?

BACKGROUND

Patients with delirium have a reduced ability to focus, to sustain or shift attention with an associated change in cognition, or they develop a perceptual disturbance that occurs over a short period of time and tends to fluctuate over the course of the day. There is usually evidence of a medical cause from history, physical examination, or investiga-

Context Delirium occurs in many hospitalized older patients and has serious consequences including increased risk for death and admission to long-term care. Despite its importance, health care clinicians often fail to recognize delirium. Simple bedside instruments may lead to improved identification.

Objective To systematically review the evidence on the accuracy of bedside instruments in diagnosing the presence of delirium in adults.

Data Sources Search of MEDLINE (from 1950 to May 2010), EMBASE (from 1980 to May 2010), and references of retrieved articles to identify studies of delirium among inpatients.

Study Selection Prospective studies of diagnostic accuracy that compared at least 1 delirium bedside instrument to the *Diagnostic and Statistical Manual of Mental Disorders*-based diagnosis made by a geriatrician, psychiatrist, or neurologist.

Data Synthesis There were 6570 unique citations identified with 25 prospectively conducted studies (N=3027 patients) meeting inclusion criteria and describing use of 11 instruments. Positive results that suggested delirium with likelihood ratios (LRs) greater than 5.0 were present for the Global Attentiveness Rating (GAR), Memorial Delirium Assessment Scale (MDAS), Confusion Assessment Method (CAM), Delirium Rating Scale Revised-98 (DRS-R-98), Clinical Assessment of Confusion (CAC), and Delirium Observation Screening Scale (DOSS). Normal results that decreased the likelihood of delirium with LRs less than 0.2 were calculated for the GAR, MDAS, CAM, DRS-R-98, Delirium Rating Scale (DRS), DOSS, Nursing Delirium Screening Scale (Nu-DESC), and Mini-Mental State Examination (MMSE). The Digit Span test and Vigilance "A" test in isolation have limited utility in diagnosing delirium. Considering the instrument's ease of use, test performance, and clinical importance of the heterogeneity in the confidence intervals (CIs) of the LRs, the CAM has the best available supportive data as a bedside delirium instrument (summary-positive LR, 9.6; 95% CI, 5.8-16.0; summary-negative LR, 0.16; 95% CI, 0.09-0.29). Of all scales, the MMSE (score <24) was the least useful for identifying a patient with delirium (LR, 1.6; 95% CI, 1.2-2.0).

Conclusion The choice of instrument may be dictated by the amount of time available and the discipline of the examiner; however, the best evidence supports use of the CAM, which takes 5 minutes to administer.

JAMA. 2010;304(7):779-786

www.jama.com

See also Patient Page.



CME available online at
www.jamaarchivescme.com
and questions on p 805.

tions.¹ Many hospitalized older patients become delirious and delirium is an independent marker for increased mortality during the 12 months after hospital admission.² In addition to death, developing delirium has been associated with longer length of hospital stay, increased hospital-acquired complications, persistent cognitive deficits, and increased discharge rates to long-term care.³⁻⁶

Author Affiliations: Division of Geriatrics, and Li Ka Shing Knowledge Institute, St. Michael's Hospital (Drs Wong and Straus), University of Toronto, Toronto, Ontario, Canada; Divisions of General Internal Medicine and Geriatrics, University of Calgary, Calgary, Alberta, Canada (Dr Holroyd-Leduc); and Durham VA Medical Center, and Duke University, Durham, North Carolina (Dr Simel).
Corresponding Author: Camilla L. Wong, MD, MHSc, FRCPC, 30 Bond St, Toronto, ON M5B 1W8, Canada (camilla.wong@utoronto.ca).

The Rational Clinical Examination Section Editors: David L. Simel, MD, MHS, Durham Veterans Affairs Medical Center and Duke University Medical Center, Durham, NC; Drummond Rennie, MD, Deputy Editor.

The symptoms of delirium may become apparent when a patient has difficulty in carrying on a normal conversation. Because visiting family members or nurses often spend more time with the patient than physicians do, they may be the first to detect delirium. Health care workers fail to recognize more than half of delirium cases.⁷⁻⁹ The *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition) (*DSM-III*) criteria for delirium, when applied systematically to 133 consecutively admitted patients to an acute medical ward, led to a diagnosis of delirium in 20 patients, only 1 of whom was reported as delirious by the primary clinician.⁹

Delirium is often iatrogenic, resulting from a diversity of problems such as adverse drug reactions, the multiple stresses of surgery, complications of procedures, or immobilization.¹⁰ Given that multiple factors usually contribute to the development of delirium, randomized trials have shown multicomponent preventive strategies to be effective in preventing delirium.¹¹⁻¹⁴

How to Diagnose Delirium

The diagnosis is primarily clinical and based on careful observation of key features. Consensus from an expert panel identified several clinical features of delirium: acute onset and fluctuating course, inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, increased or decreased psychomotor activity, and disturbance of the sleep-wake cycle.¹⁵ The key diagnostic feature that helps to distinguish delirium from dementia is that delirium has an acute and rapid onset, whereas dementia is much more gradual in progression. Alternations in attention and changes in level of consciousness also favor a diagnosis of delirium. Delirium typically presents in 1 of 2 major forms—hypoactive or hyperactive. The hypoactive form is characterized by lethargy and reduced psychomotor functioning; this form often

goes unrecognized by clinicians and caregivers. The hyperactive form is characterized by agitation, increased vigilance, and often hallucinations; this form rarely goes unnoticed by clinicians or caregivers. There is also a mixed form of delirium in which patients fluctuate between the hypoactive and hyperactive forms.¹⁶

The criterion standard for the diagnosis of delirium as defined by the *DSM* continues to evolve.^{1,17,18} A formal assessment using this reference standard, however, may involve an in-depth interview and series of cognitive tests by a clinician familiar with *DSM-IV* criteria. Even in the hands of experts, the reference standard for delirium is a clinical diagnosis based on the *DSM-IV* criteria and may be prone to some subjectivity. For example, one study noted the agreement κ for reliability between 2 geriatricians with each classification of delirium to be 0.74 (*DSM-III*), 0.74 (*DSM-III-R*), and 0.72 (*DSM-IV*).¹⁹ Simpler bedside instruments may better guide which patients should receive formal consultation and intervention. The objective of this review was to determine the diagnostic accuracy of bedside delirium instruments.

METHODS

Literature Search Strategy

Searches of MEDLINE (from 1950 to May 2010) and EMBASE (from 1980 to May 2010) using Ovid were completed to identify studies performed in a clinical setting. The search strategy used terms including *confusion* and *delirium* combined with validated search filters for retrieving articles^{20,21} on the diagnosis of health disorders (eBox available at <http://www.jama.com>). Additional articles were identified from searching the bibliographies of retrieved articles.

Study Selection and Data Extraction

Inclusion criteria were published prospective studies that were conducted in hospitalized patients not in the intensive care unit, described the use of an appropriate reference standard (*DSM-*

III, *DSM-III-R*, or *DSM-IV*),^{1,17,18} had the reference standard performed by a specialist physician (geriatrician, neurologist, or psychiatrist), applied the same index test to most patients (>80%), applied the same reference tests to all patients and all reference test results were available, and included participants with and without delirium. Exclusion criteria were studies involving mostly alcohol-related delirium or a pediatric population, studies in which the index and reference tests were performed by the same individual, and duplicate or non-English-language publication. The bedside index instrument must be feasible in a clinical setting, without requiring special equipment, and may be performed by a nonexpert. Furthermore, primary data or appropriate summary statistics had to be available. When necessary, additional data were obtained by contacting study authors.

Each abstract was reviewed independently by 2 reviewers to select relevant publications that met the inclusion criteria for data extraction. In cases of doubt, full-text articles were retrieved for review and discussion. Two investigators independently reviewed each full-text article to confirm that inclusion and exclusion criteria were met. Disagreements were resolved by discussion, and when necessary, with a third reviewer.

Data were extracted from the included studies independently by 2 reviewers. Disagreements were resolved by consensus with the third reviewer. Information was extracted using a specially designed form based on the principles outlined by the Standards for Reporting of Diagnostic Accuracy (STARD).²² Details pertaining to study quality included study size, participant recruitment method, demographic characteristics of participants, application of reference standard, application of diagnostic test(s), presence of blinding, independence of tests, and attrition rates. Study quality was summarized using a quality checklist designed for the Rational Clinical Examination series in which a threshold

of 100 patients was used to distinguish level 1 from level 2 studies.²³

Statistical Methods

For studies of test accuracy, sensitivity, specificity, and likelihood ratios (LRs) were calculated from the raw data and then rounded for display in the data tables.²⁴ If a study contained any zeros in the 2 × 2 table, resulting in likelihood estimates of zero or infinity, 0.5 was added to all the counts for that study for calculating the LR and respective confidence intervals (CIs). For studies that evaluated the same instrument at different threshold levels, the data were abstracted at each level and then the optimum threshold was selected based on a balance between the diagnostic odds ratio (OR) and our confidence in the findings because the threshold was studied by multiple investigators. The diagnostic OR (positive LR divided by negative LR) is a single indicator of test performance in which higher values indicate better discriminatory test performance. Summary LRs for instruments were derived using a univariate random-effects model (Comprehensive MetaAnalysis, version 2.2046, Biostat Inc, Englewood, New Jersey) since bivariate measures were similar or failed to converge on a solution.²⁵

Heterogeneity was explored using I^2 , which describes the percentage of total variation across studies that is due to heterogeneity rather than chance. Heterogeneity was categorized as low, moderate, and high to I^2 values of 25%, 50%, and 75%.²⁶ If there was at least moderate heterogeneity, summary statistics were recalculated by eliminating the individual studies that contributed most to heterogeneity or by selecting only certain subgroups in a sensitivity analysis. Heterogeneity was assessed using Meta-Disc, version 1.4 (Unit of Clinical Biostatistics, Ramón y Cajal Hospital, Madrid, Spain).²⁷

RESULTS

Study Characteristics

We identified 6570 potential citations of which 79 unique studies were re-

trieved for full-text review. Fifty-four studies were later excluded for a variety of reasons, leaving 25 studies^{15,28-51} suitable for data extraction and synthesis (eFigure 1).

The included studies ranged in size from 26 to 791 participants (N = 3027; eTable 1). Eleven bedside delirium instruments, with data on diagnostic accuracy, met our inclusion and exclusion criteria (BOX 1).*

All studies provided details on participant recruitment, and 9 enrolled consecutive patients.^{30,33-36,38,41-43} Eighteen studies described the use of independent, blinded assessment of reference and diagnostic tests in a clinical setting.[†] Application of the diagnostic tests was consistent and complete in all but 2 studies.^{39,45} Application of the reference test, a DSM-based diagnosis by a specialist physician, was identical within each study. Overall, 1, 7, 9, and 8 studies were rated levels 1, 2, 3, and 4, respectively, in terms of quality level of evidence using the Rational Clinical Examination quality scale (eTable 1).²³

Prevalence of Delirium

From the 9 studies^{30,33-36,38,41-43} that enrolled patients consecutively, the study prevalence of delirium ranged from 9% to 63% depending on the clinical setting (eTable 1). The prevalent cases in some of these studies likely included patients admitted with delirium, but also patients who developed delirium (incident cases). The highest prevalence, 63%, was in a study of hospitalized patients in an oncology or palliative care ward presenting with mental status change.³³ The prevalence of delirium among patients admitted to a geriatric unit ranged from 9% to 43%^{36,38,42,46} and from 12% to 27% among patients in a postcardiac surgery unit.^{35,41}

Accuracy of Clinical Examination Findings in the Diagnosis of Delirium

Of the 11 bedside instruments, positive findings on the Global Attentive-

ness Rating (GAR), Memorial Delirium Assessment Scale (MDAS), Confusion Assessment Method (CAM), Delirium Rating Scale Revised-98 (DRS-R-98), Clinical Assessment of Confusion (CAC), and Delirium Observation Screening Scale (DOSS) all had an LR of greater than 5.0 for diagnosing delirium (TABLE). A commonly used test for dementia, the Mini-Mental State Examination (MMSE) was the least useful for identifying patients with delirium (LR, 1.6 for a score <24). Negative findings on the GAR, MDAS, CAM, DRS-R-98, DRS, DOSS, Nursing Delirium Screening Scale (Nu-DESC), and MMSE made delirium less likely with LRs of 0.2 or less (Table). Circumstances may only allow a clinician to use a time-efficient bedside instrument that can be completed in 5 minutes or less, such as the CAM, DOSS, GAR, or Nu-DESC.^{15,31,38,48} Also, the choice of instrument may be guided by the background of the examiner. For example, when only considering studies in which a time-efficient instrument was performed by nurses, the CAM had the best diagnostic OR (summary-positive LR, 7.3; 95% CI, 1.9-27; and summary-negative LR, 0.08; 95% CI, 0.03-0.21).^{31,43} A physician may also choose to use the CAM (summary-positive LR, 19; 95% CI, 6.7-51; and summary-negative LR, 0.19; 95% CI, 0.13-0.27)^{15,32,36,37,50,51} rather than the MDAS, DRS, or DRS-R-98 because it can be performed in less than 5 minutes. Lastly, the accuracy of the GAR (positive LR, 65; 95% CI, 9.3-458; and negative LR, 0.06; 95% CI, 0.01-0.38) may not apply to nongeriatricians and has only been assessed in 1 study.³⁸

There was significant heterogeneity among the studies likely owing to differences in the study quality, experience level of the individual performing the index test, and the version of the DSM criteria used for the reference standard. For the CAM, the problem of heterogeneity for the negative LR was resolved ($I^2=0\%$) when only studies in which the index test was performed by a physician were included, but the negative LR was not substan-

*References 15, 29, 31, 38, 39, 47, 48, 52-55.

†References 15, 29-34, 36-41, 45, 47, 48, 50, 51.

Box 1. Bedside Delirium Instruments**Clinical Assessment of Confusion (CAC)⁵²**

A nursing checklist of 25 psychomotor behaviors associated with varying degrees of confusion (range, 0-77). The presence of more behaviors is associated with more severe confusion.

Confusion Assessment Method (CAM)^{15a}

The CAM includes an instrument and diagnostic algorithm for identification of delirium. The instrument assesses the presence, severity, and fluctuation of 9 delirium features: acute onset, inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation or retardation, and altered sleep-wake cycle. The questionnaire can be administered in 5 minutes. The algorithm is based on the cardinal elements of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) (Third Edition, Revised) criteria for delirium: features 1 (acute onset and fluctuating course) and 2 (inattention) are essential features, and feature 3 (disorganized thinking) or 4 (altered level of consciousness) is supported by expert judgment and clinical practice, in which the first 2 and either of the latter 2 are required for diagnosis.

Delirium Observation Screening Scale (DOSS)^{48,53}

The original version consisted of a 25-item scale based on the DSM-IV criteria for delirium. The scale was designed to capture early symptoms of delirium that nurses could observe during regular care. The scale was subsequently reduced to 13 observations that could each be rated as normal (score, 0) or abnormal (score, 1). A total score of 3 or more points indicates delirium. Completion of the instrument requires less than 5 minutes.

Delirium Rating Scale (DRS)⁴⁷

A 10-item observational scale (range, 0-32) that rates patients on the characteristic symptoms of delirium including temporal onset, perceptual disturbance, hallucinations, delusions, psychomotor behavior, cognitive status, physical disorder, sleep-wake cycle disturbance, lability of mood, and variability of symptoms. It is intended to be used by clinicians with psychiatric training.

Delirium Rating Scale-Revised-98 (DRS-R-98)⁵⁴

This is a revised version of the DRS. It is more comprehensive and separates the scale into 2 sections: 3 diagnostic items for initial ratings and a 13-item scale for repeated measures (range, 0-46).

Digit Span Test³⁹

A series of random numbers are presented at a rate of 1 per second. The patient is asked to repeat the sequence. Starting with a 2-number sequence, each correctly repeated series is followed by a sequence with 1 additional digit. The test result is abnormal if the patient cannot repeat at least 5 digits.

Global Attentiveness Rating (GAR)³⁸

A physician-rated 10-cm visual analog scale in which a high rating indicates the patient can be easily engaged and a low rating indicates the patient cannot be aroused or is so agitated that a conversation is not possible. The assessment is based on a minimum of 2 minutes of general conversation with the patient, without any formal cognitive testing or collateral information.

Memorial Delirium Assessment Scale (MDAS)²⁹

A 10-item, 4-point clinician-rated scale (range, 0-30) designed to quantify the severity of delirium. Scale items reflect the DSM criteria for delirium, such as assessing for disturbances in arousal and level of consciousness, as well as several areas of cognitive functioning. It is intended for repeated administration for measurement of change. Completion of the instrument takes approximately 10 minutes.

Mini-Mental State Examination (MMSE)⁵⁵

A widely used brief objective measure of cognitive function in older adults. It is a 30-item instrument that includes tests of orientation, memory, and concentration.

Nursing Delirium Screening Scale (Nu-DESC)³¹

A screening scale (range, 0-10) designed to be administered by a nurse based on clinical observation in routine practice. Five symptoms are rated: disorientation, inappropriate behavior, inappropriate communication, hallucination, and psychomotor retardation. The threshold for delirium is a score of 2 or more. On average, the instrument can be completed in 1 minute.

Vigilance "A" Test³⁹

A list of 60 letters in which 18 of them are the letter A is read to the patient at a rate of 1 letter per second. The patient is asked to indicate to the examiner each time the letter A is heard. More than 2 errors is considered abnormal.

^aThe copyright of the CAM is owned by Sharon K. Inouye, Yale University School of Medicine. A training manual and permission to use the CAM are available online at <http://hospitalelderlifeprogram.org/public/public-main.php>.

tially different (summary-negative LR, 0.19; 95% CI, 0.13-0.27).^{15,32,36,37,50,51} When a similar sensitivity analysis was performed for the positive LR, the positive LR increased (summary-positive LR, 19; 95% CI, 6.7-51), but the heterogeneity persisted ($I^2=67\%$).^{15,32,36,37,50,51} When only levels 1 and 2 evidence stud-

ies were included, heterogeneity disappeared for the negative LR ($I^2=0\%$), but without any substantial change in the negative LR (summary-negative LR, 0.20; 95% CI, 0.14-0.30).^{34,36,37,50,51} Sub-analysis by language or DSM version did resolve heterogeneity. Interrater reliability data were available for some of

the studies for the CAM, DRS, DRS-R-98, and MDAS. All had a substantial degree of agreement (eTable 2).

COMMENT**Limitations**

The results of this review should be interpreted within the context of the

Table. Summary Data for Diagnostic Accuracy of Bedside Instruments for Diagnosing Delirium^a

Source	Sample Size	Delirium Prevalence, %	Examiner Specialty	% (95% Confidence Interval)		Likelihood Ratio (95% Confidence Interval)	
				Sensitivity	Specificity	Positive	Negative
CAC							
Pompei et al, ³⁹ 1995	428	15	Research assistant	36 (24-49)	95 (92-97)	7.4 (4.2-13)	0.67 (0.56-0.81)
CAM ^b							
Farrell and Ganzini, ³⁴ 1995	72	45	Osteopathic physician	93 (77-98)	93 (80-98)	13 (4.4-39)	0.07 (0.02-0.28)
Gaudreau et al, ³¹ 2005	59	32	Nurse	98 (70-100)	95 (82-99)	16 (4.8-53)	0.03 (0.00-0.42)
Gonzalez et al, ³² 2004	123	24	General physician or psychiatrist	90 (73-97)	99 (92-100)	167 (10-2654)	0.11 (0.04-0.31)
Hestermann et al, ⁵⁰ 2009 ^c	39	33	Gerontologist or resident physician in geriatric medicine	77 (48-92)	96 (77-99)	20 (2.9-140)	0.24 (0.09-0.65)
Inouye et al, ¹⁵ 1990 (site 1)	30	33	Geriatrician	95 (55-100)	95 (72-99)	13 (2.8-63)	0.05 (0.00-0.74)
Inouye et al, ¹⁵ 1990 (site 2)	26	62	Geriatrician	94 (66-99)	90 (53-99)	9.4 (1.5-60)	0.07 (0.01-0.47)
Laurila et al, ³⁶ 2002 ^d	81	40	Geriatrician	81 (64-91)	84 (71-92)	5.0 (2.6-9.6)	0.22 (0.11-0.47)
Leung et al, ³⁷ 2008	100	25	Family physician	76 (56-89)	99 (90-100)	114 (7.1-1822)	0.25 (0.13-0.49)
Pompei et al, ³⁹ 1995	428	15	Research assistant	46 (34-58)	92 (89-94)	5.8 (3.7-9.0)	0.59 (0.47-0.74)
Rolfson et al, ⁴¹ 1999 ^e	30	27	Nurse	13 (0-53)	100 (85-100)	7.7 (0.34-171)	0.85 (0.63-1.1)
Ryan et al, ⁵¹ 2009 ^f	52	33	Hospitalist	88 (63-97)	99 (82-100)	65 (4.1-1034)	0.14 (0.04-0.45)
Zou et al, ⁴³ 1998	87	49	Nurse	93 (80-98)	77 (63-87)	4.1 (2.4-7.1)	0.09 (0.03-0.27)
Pooled ^g	1036			86 (74-93)	93 (87-96)	9.6 (5.8-16.0)	0.16 (0.09-0.29)
DOSS (13-Item) ^b							
van Gerner and Schuurmans, ⁴⁵ 2007	87	10	Nurse	89 (50-98)	88 (79-94)	7.6 (4.0-15)	0.13 (0.02-0.80)
Schuurmans et al, ⁴⁸ 2003	92	20	Nurse	94 (69-99)	76 (65-84)	3.9 (2.6-5.9)	0.07 (0.01-0.50)
Pooled ^g	178			92 (74-98)	82 (66-92)	5.2 (2.7-9.9)	0.10 (0.03-0.37)
DRS ≥ 10 ^{b,h}							
Grassi et al, ³³ 2001	105	63	Research psychologist	95 (87-99)	62 (46-75)	2.5 (1.7-3.7)	0.07 (0.02-0.23)
Rockwood et al, ⁴⁰ 1996 ⁱ	67	37	Resident physician in geriatric medicine or psychiatry	82	94	14	0.19
Rosen et al, ⁴² 1994	791	9	Research clinician	94 (86-98)	82 (79-84)	5.1 (4.3-6.0)	0.07 (0.03-0.18)
Trzepacz et al, ⁴⁷ 1988	47	43	Psychiatrist	98 (71-100)	98 (77-100)	55 (3.5-853)	0.02 (0.00-0.38)
Pooled ^g	943			95 (90-98)	79 (58-91)	4.3 (2.1-9.1)	0.07 (0.03-0.13)
DRS-R-98 >20 ^{b,h}							
de Negreiros et al, ⁴⁴ 2008	64	40	Researcher	93 (75-98)	95 (81-99)	17 (4.4-66)	0.08 (0.02-0.30)
de Rooij et al, ³⁰ 2006	65	35	Geriatrician or psychiatrist	93 (65-99)	82 (69-90)	5.2 (2.8-9.5)	0.08 (0.01-0.54)
Pooled ^g	129			93 (80-98)	89 (68-97)	8.0 (2.6-25)	0.08 (0.03-0.24)
Digit Span Test							
Pompei et al, ³⁹ 1995	419	15	Research assistant	34 (22-48)	90 (87-93)	3.4 (2.1-5.5)	0.73 (0.61-0.89)
GAR <7							
O'Keefe and Gosney, ³⁸ 1997	87	21	Geriatrician	94 (73-100)	99 (92-100)	65 (9.3-458)	0.06 (0.01-0.38)
MDAS ≥ 10 ^{b,h}							
Breitbart et al, ²⁹ 1997	33	52	Psychiatrist	82 (57-94)	75 (49-90)	3.3 (1.4-7.9)	0.24 (0.08-0.68)
Kazmierski et al, ³⁵ 2008	260	12	Psychiatrist	97 (80-100)	96 (92-98)	22 (12-41)	0.04 (0.01-0.24)
Matsuoka et al, ⁴⁶ 2001	37	43	Psychiatrist	97 (66-100)	98 (72-100)	43 (2.8-662)	0.03 (0.00-0.46)
Pooled ^g	330			92 (75-98)	92 (70-98)	12 (2.4-58)	0.09 (0.02-0.38)
MMSE <24							
Grassi et al, ³³ 2001	105	63	Trained research psychologist	96 (87-99)	38 (23-55)	1.6 (1.2-2.0)	0.12 (0.04-0.38)
Nu-DESC >0							
Leung et al, ³⁷ 2008	100	25	Nurse	96 (80-100)	69 (59-79)	3.1 (2.3-4.4)	0.06 (0.01-0.40)
Vigilance "A" Test							
Pompei et al, ³⁹ 1995	421	15	Research assistant	61 (47-74)	77 (73-81)	2.7 (2.0-3.5)	0.50 (0.36-0.71)

Abbreviations: CAC, Clinical Assessment of Confusion; CAM, Confusion Assessment Method; DOSS, Delirium Observation Screening Scale; DRS, Delirium Rating Scale; DRS-R, Delirium Rating Scale-Revised; GAR, Global Attention Rating; MDAS, Memorial Delirium Assessment Scale; MMSE, Mini-Mental State Examination; Nu-DESC, Nursing Delirium Screening Scale.

^aReliability measures are available at eTable 2, see <http://www.jama.com>.

^bHeterogeneity was explored using I^2 , which describes the percentage of total variation across studies that is due to heterogeneity rather than chance. Heterogeneity was categorized as low, moderate, and high, respective to I^2 values of 25%, 50%, and 75%. Positive and negative F , % values were 65 and 85 for CAM, 65 and 0 for DOSS, 14 and 0.19 for DRS, 73 and 0 for DRS-R-98, and 85 and 69 for MDAS.

^cCalculated using the more conservative results of the 2 examiners.

^dUsing *Diagnostic and Statistical Manual of Mental Disorders* Fourth Edition criteria as the reference standard.

^eStudy not included in the summary data because results were based only on 8 patients with delirium and only 1 patient in the entire study population received positive screening results using the CAM instrument.

^fThere was also a pilot study with 32 patients in the same publication.

^gUnivariate random effects.

^hSpecific threshold chosen based on summary diagnostic odds ratio and validation in multiple studies.

ⁱIndividual data not available for confidence interval calculations or inclusion in meta-analysis.

Box 2. The Confusion Assessment Method (CAM) Diagnostic Algorithm**Features and Descriptions**

Feature 1: Acute onset and fluctuating course

Is there an acute change in mental status from the patient's baseline? Did this behavior fluctuate during the past day (tend to come and go or increase and decrease in severity)?

Feature 2: Inattention

Does the patient have difficulty focusing attention, for example, being easily distractible, or having difficulty keeping track of what was being said?

Feature 3: Disorganized thinking

Is the patient's speech disorganized or incoherent, such as rambling or irrelevant conversation, unclear or illogical flow of ideas, or unpredictable switching from subject to subject?

Feature 4: Altered level of consciousness

Overall, how would you rate this patient's level of consciousness?

Alert (normal)

Vigilant (hyperalert)

Lethargic (drowsy, easily aroused)

Stupor (difficult to arouse)

Coma (unarousable)

The diagnosis of delirium requires a present/abnormal rating for features 1 and 2 and also for either feature 3 or 4.

The algorithm is based on the cardinal elements of the *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition, Revised) criteria for delirium. Adapted from Inouye (1990) and Ely (2001).^{15,56} A space for clinicians to indicate each of the 4 features as either absent or present appears on the actual algorithm.

included studies. The reference standard for delirium is a clinical diagnosis based on the *DSM* criteria and may be prone to some subjectivity. For example, one study noted the agreement κ for reliability between 2 geriatricians with each classification of delirium to be 0.74 (*DSM-III*), 0.74 (*DSM-III-R*), and 0.72 (*DSM-IV*).¹⁹ Although this level of agreement would be considered substantial, less trained clinicians may have lower levels of agreement when evaluating patients for delirium.

Most studies occurred in university hospitals and the data represented heterogeneous patient populations with different prevalence rates ranging from patients in postcardiac surgery wards to those in palliative care wards. Thus, the data likely include both prevalent cases (patients admitted with delirium) and incident cases (delirium developed

during the hospital admission) with a wide spectrum of pretest probabilities. The pretest probability depends on the clinical setting (eg, whether the patient is postoperative) and underlying diagnoses. This broad spectrum of probabilities can lead to varying estimates of the probability of delirium after using the screening tests.

For most of the LRs with heterogeneity, the clinical impact can still be inferred from the CIs. While heterogeneity was still present for instruments with many studies such as the CAM, these studies also allow the instrument to be tested across multiple populations to ensure its adequate performance and generalizability across different settings. Clinicians should be less certain when interpreting LRs based on single, small studies. For instance, the extremely high, positive LR and extremely low, negative LR for certain

thresholds on the MDAS and DRS have not been validated with further studies. Some studies may have been missed given we limited the search terms to *confusion* and *delirium*. Like other types of systematic reviews, systematic reviews of diagnostic tests are subject to publication bias and may exaggerate the summary estimate of test accuracy if publication is related to the positivity of results. There were insufficient studies to look for funnel plot asymmetry.

How to Learn the Method for Diagnosing Delirium

CAM. Some training is recommended for optimal use and an instruction manual is available online (<http://hospitalelderlifeprogram.org/pdf/>). The validity of the CAM for unstandardized observations is poor and performance is linked to the quality of training.^{8,51} Thus, the CAM was designed to be scored based on observations made during formal cognitive assessment with brief instruments like the MMSE and the Digit Span test, which were used in the original validation study.⁵⁵ Although there are several clinical features of delirium, the CAM diagnostic algorithm is based on only 4 cardinal elements of the *DSM-III-R* criteria for delirium: acute onset and fluctuating course, inattention, disorganized thinking, and altered level of consciousness (BOX 2; eFigure 2). The algorithm should be based on observations during any contact with the patient and need not be limited to the interview period alone, but the algorithm should be applied immediately following the interview to ensure accuracy. Although not included in this review, the CAM has also been adapted for other clinical settings such as the intensive care unit,⁵⁶ emergency department,⁵⁷ and nursing home.⁵⁸

GAR. The global assessment is based on a minimum of 2 minutes of general conversation with the patient, without necessarily any formal cognitive testing or corroborative information. The question to be answered by the clinician is, "How well did the patient keep his mind on interacting with you dur-

ing the interview?" The answer is rated on an uninterrupted 10-cm visual analog scale.

We systematically determined the accuracy of bedside delirium instruments in diagnosing the presence of delirium and found several sensitive, specific, rapid, and simple instruments which may be useful for assessment of delirium. Although there have been other systematic reviews of bedside delirium instruments,^{59,60} to our knowledge, this is the first study to include such a breadth of instruments and which also quantifies diagnostic accuracy with summary LRs.

The prevalence of delirium is high enough that a bedside delirium instrument should be used in patients 65 years old or older, particularly if there is an impression of a change in mental status. Because of its accuracy, brevity, and ease of use by clinical and lay interviewers, the CAM has become the most widely used standardized delirium instrument for clinical and research purposes over the past 2 decades.⁶⁰ It has been translated into 10 languages for which published articles are available.⁶⁰

Since a busy physician may not have 5 minutes to spend assessing a patient's mental status, the most time-efficient screening approach may be to simply have a 2-minute general conversation with the patient. Although the overall impression of the quality and ease with which the physician can easily keep the patient engaged throughout a minimum of 2 minutes of general conversation may be a very valuable guide to delirium (GAR), this is supported by only 1 study with geriatricians. Thus, generalizability may be limited to physicians with expertise in this patient population.

This review highlights some gaps in the literature pertaining to bedside delirium instruments. Future areas of study include testing educational interventions to improve the ability of health care professionals to identify delirium with the available instruments, correlating results of the instruments with clinical outcomes, and validating

these instruments outside of the hospital setting.

SCENARIO RESOLUTION

The older hospitalized patient should be evaluated for delirium, especially if there is an impression of mental status change. The pretest probability of delirium in this patient was 63% based on the clinical setting (oncology ward patient with pneumonia).³³ After conducting a formal cognitive assessment with the MMSE, both the physician and nurse used the CAM and agreed the change in mental status was acute and fluctuating in course, the patient was inattentive, and there was an altered level of consciousness. The posttest probability of a delirium was 93% and 97% (pretest odds \times positive LR = posttest odds) for the nurse and physician, respectively. The causes for his delirium require prompt assessment.

Author Contributions: Dr Wong had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Wong, Holroyd-Leduc, Straus.

Acquisition of data: Wong, Straus.

Analysis and interpretation of data: Wong, Holroyd-Leduc, Simel, Straus.

Drafting of the manuscript: Wong, Straus.

Critical revision of the manuscript for important intellectual content: Wong, Holroyd-Leduc, Simel, Straus.

Statistical analysis: Wong, Simel.

Study supervision: Holroyd-Leduc, Simel, Straus.

Financial Disclosures: None reported.

Funding/Support: Dr Straus was supported by a Tier 2 Canada Research Chair and a Health Scholar Award from the Alberta Heritage Foundation for Medical Research.

Role of the Sponsor: Supporters for the Tier 2 Canada Research Chair and the Health Scholar Award, and the Alberta Heritage Foundation had no role in the design and conduct of the study, the collection, management, analysis, or for the interpretation of the data; or for the preparation, review, or approval of the manuscript.

Disclaimer: Dr Simel had no role in the peer review of or the editorial decision to accept this article for publication.

Online-Only Material: The eBox, eFigures, eTables, and eAppendix are available online at <http://www.jama.com>.

Additional Contributions: We thank Laure Perrier, MEd, MLIS, Faculty of Medicine, Continuing Education & Professional Development, University of Toronto, for her assistance in the literature search; and Rhianna Hibberd, BA, Alberta Health Services, Calgary Health Region, for retrieval of relevant articles as part of their normal duties. We also thank Linda Ganzini, MD, MPH, Department of Psychiatry, Oregon Health & Science University; Marieke Schuurmans, PhD, RN, Division of Nursing Science, University Medical Center Utrecht; Karen Ryan, MB BCH, BAO, BMed Sc, MRCPI, Department of Palliative Medi-

cine, Mater Misericordiae University Hospital; and Melissa Andrew, MD, MSc(PH), Division of Geriatric Medicine, Dalhousie University, for providing additional data on their studies. We also thank Barbara Kamholz, MD, and Tracey Holsinger, MD, Department of Psychiatry, and John Williams, MD, Department of Medicine, Duke University, for their valuable comments on previous drafts of the manuscript. None of the acknowledged individuals received compensation for their contributions to this article.

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 2000.
2. McCusker J, Cole M, Abrahamowicz M, Primeau F, Belzile E. Delirium predicts 12-month mortality. *Arch Intern Med*. 2002;162(4):457-463.
3. Francis J, Kapoor WN. Prognosis after hospital discharge of older medical patients with delirium. *J Am Geriatr Soc*. 1992;40(6):601-606.
4. Gustafson Y, Berggren D, Brännström B, et al. Acute confusional states in elderly patients treated for femoral neck fracture. *J Am Geriatr Soc*. 1988;36(6):525-530.
5. Inouye SK, Rushing JT, Foreman MD, Palmer RM, Pompei P. Does delirium contribute to poor hospital outcomes? *J Gen Intern Med*. 1998;13(4):234-242.
6. O'Keefe ST, Lavan J. The prognostic significance of delirium in older hospital patients. *J Am Geriatr Soc*. 1997;45(2):174-178.
7. Gustafson Y, Brännström B, Norberg A, Bucht G, Winblad B. Underdiagnosis and poor documentation of acute confusional states in elderly hip fracture patients. *J Am Geriatr Soc*. 1991;39(8):760-765.
8. Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM Jr. Nurses' recognition of delirium and its symptoms. *Arch Intern Med*. 2001;161(20):2467-2473.
9. Cameron DJ, Thomas RI, Mulvihill M, Bronheim H. Delirium: a test of the Diagnostic and Statistical Manual III criteria on medical inpatients. *J Am Geriatr Soc*. 1987;35(11):1007-1010.
10. Inouye SK, Schlesinger MJ, Lydon TJ. Delirium: a symptom of how hospital care is failing older persons and a window to improve quality of hospital care. *Am J Med*. 1999;106(5):565-573.
11. Lundström M, Olofsson B, Stenvall M, et al. Post-operative delirium in old patients with femoral neck fracture: a randomized intervention study. *Aging Clin Exp Res*. 2007;19(3):178-186.
12. Inouye SK, Bogardus STJ Jr, Charpentier PA, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med*. 1999;340(9):669-676.
13. Marcantonio ER, Flacker JM, Wright RJ, Resnick NM. Reducing delirium after hip fracture: a randomized trial. *J Am Geriatr Soc*. 2001;49(5):516-522.
14. Vidán M, Serra JA, Moreno C, Riquelme G, Ortiz J. Efficacy of a comprehensive geriatric intervention in older patients hospitalized for hip fracture. *J Am Geriatr Soc*. 2005;53(9):1476-1482.
15. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion. *Ann Intern Med*. 1990;113(12):941-948.
16. Liptzin B, Levkoff SE. An empirical study of delirium subtypes. *Br J Psychiatry*. 1992;161:843-845.
17. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Rev 3rd ed. Washington, DC: American Psychiatric Association; 1980.
18. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Washington, DC: American Psychiatric Association; 1987.

19. Laurila JV, Pitkala KH, Strandberg TE, Tilvis RS. Delirium among patients with and without dementia: does the diagnosis according to the DSM-IV differ from the previous classifications? *Int J Geriatr Psychiatry*. 2004;19(3):271-277.
20. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ*. 2004;328(7447):1040-1044.
21. Wilczynski NL, Haynes RB; Hedges Team. EMBASE search strategies for identifying methodologically sound diagnostic studies for use by clinicians and researchers. *BMC Med*. 2005;3:7.
22. Bossuyt PM, Reitsma JB, Bruns DE, et al; Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med*. 2003;138(1):40-44.
23. Simel DL, Rennie D, eds. Update: primer on precision and accuracy. In: *Rational Clinical Examination: The Evidence-Based Clinical Diagnosis*. New York, NY: McGraw-Hill; 2008:9-16.
24. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol*. 1991;44(8):763-770.
25. Simel DL, Bossuyt PM. Differences between univariate and bivariate models for summarizing diagnostic accuracy may not be large. *J Clin Epidemiol*. 2009;62(12):1292-1300.
26. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-560.
27. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
28. Andrew MK, Bhat R, Clarke B, Freter SH, Rockwood MR, Rockwood K. Inter-rater reliability of the DRS-R-98 in detecting delirium in frail elderly patients. *Age Ageing*. 2009;38(2):241-244.
29. Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S. The Memorial Delirium Assessment Scale. *J Pain Symptom Manage*. 1997;13(3):128-137.
30. de Rooij SE, van Munster BC, Korevaar JC, et al. Delirium subtype identification and the validation of the Delirium Rating Scale—Revised-98 (Dutch version) in hospitalized elderly patients. *Int J Geriatr Psychiatry*. 2006;21(9):876-882.
31. Gaudreau JD, Gagnon P, Harel F, Tremblay A, Roy MA. Fast, systematic, and continuous delirium assessment in hospitalized patients. *J Pain Symptom Manage*. 2005;29(4):368-375.
32. González M, de Pablo J, Fuente E, et al. Instrument for detection of delirium in general hospitals. *Psychosomatics*. 2004;45(5):426-431.
33. Grassi L, Caraceni A, Beltrami E, et al. Assessing delirium in cancer patients. *J Pain Symptom Manage*. 2001;21(1):59-68.
34. Farrell KR, Ganzini L. Misdiagnosing delirium as depression in medically ill elderly patients. *Arch Intern Med*. 1995;155(22):2459-2464.
35. Kazmierski J, Kowman M, Banach M, et al. 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*. 2008;49(1):73-76.
36. Laurila JV, Pitkala KH, Strandberg TE, Tilvis RS. Confusion assessment method in the diagnostics of delirium among aged hospital patients. *Int J Geriatr Psychiatry*. 2002;17(12):1112-1119.
37. Leung JM, Leung VW, Leung CM, Pan PC. 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. *Gen Hosp Psychiatry*. 2008;30(2):171-176.
38. O'Keeffe ST, Gosney MA. Assessing attentiveness in older hospital patients. *J Am Geriatr Soc*. 1997;45(4):470-473.
39. Pompei P, Foreman M, Cassel CK, Alessi C, Cox D. Detecting delirium among hospitalized older patients. *Arch Intern Med*. 1995;155(3):301-307.
40. Rockwood K, Goodman J, Flynn M, Stolee P. Cross-validation of the Delirium Rating Scale in older patients. *J Am Geriatr Soc*. 1996;44(7):839-842.
41. Rolfson DB, McElhane JE, Jhangri GS, Rockwood K. Validity of the confusion assessment method in detecting postoperative delirium in the elderly. *Int Psychogeriatr*. 1999;11(4):431-438.
42. Rosen J, Sweet RA, Mulsant BH, Rifai AH, Pasternak R, Zubenko GS. The Delirium Rating Scale in a psychogeriatric inpatient setting. *J Neuropsychiatry Clin Neurosci*. 1994;6(1):30-35.
43. Zou Y, Cole MG, Primeau FJ, McCusker J, Bellavance F, Laplante J. Detection and diagnosis of delirium in the elderly. *Int Psychogeriatr*. 1998;10(3):303-308.
44. de Negreiros DP, da Silva Meleiro AMA, Furlanetto LM, Trzepacz PT. Portuguese version of the Delirium Rating Scale-Revised-98: reliability and validity. *Int J Geriatr Psychiatry*. 2008;23(5):472-477.
45. Gemert van LA, Schuurmans MJ. The Neeham Confusion Scale and the Delirium Observation Screening Scale. *BMC Nurs*. 2007;6:3.
46. Matsuoka Y, Miyake Y, Arakaki H, Tanaka K, Saeki T, Yamawaki S. Clinical utility and validation of the Japanese version of Memorial Delirium Assessment Scale in a psychogeriatric inpatient setting. *Gen Hosp Psychiatry*. 2001;23(1):36-40.
47. Trzepacz PT, Baker RW, Greenhouse J. A symptom rating scale for delirium. *Psychiatry Res*. 1988;23(1):89-97.
48. Schuurmans MJ, Shortridge-Baggett LM, Duursma SA. The Delirium Observation Screening Scale. *Res Theory Nurs Pract*. 2003;17(1):31-50.
49. Huang MC, Lee CH, Lai YC, Kao YF, Lin HY, Chen CH. Chinese version of the Delirium Rating Scale-Revised-98. *Compr Psychiatry*. 2009;50(1):81-85.
50. Hestermann U, Backenstrass M, Gekle I, et al. Validation of a German version of the Confusion Assessment Method for delirium detection in a sample of acute geriatric patients with a high prevalence of dementia. *Psychopathology*. 2009;42(4):270-276.
51. Ryan K, Leonard N, Guerin S, Donnelly S, Conroy M, Meagher D. Validation of the confusion assessment method in the palliative care setting. *Palliat Med*. 2009;23(1):40-45.
52. Vermeersch PE. The clinical assessment of confusion-A. *Appl Nurs Res*. 1990;3(3):128-133.
53. Schuurmans MJ, Donders RT, Shortridge-Baggett LM, Duursma SA. Delirium case finding: pilot testing of a new screening scale for nurses. *J Am Geriatr Soc*. 2002;50(5):A8.
54. Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-revised-98. *J Neuropsychiatry Clin Neurosci*. 2001;13(2):229-242.
55. Folstein MF, Folstein SE, McHugh PR. "Mental state." *J Psychiatr Res*. 1975;12(3):189-198.
56. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients. *JAMA*. 2001;286(21):2703-2710.
57. Lewis LM, Miller DK, Morley JE, Nork MJ, Lasater LC. Unrecognized delirium in ED geriatric patients. *Am J Emerg Med*. 1995;13(2):142-145.
58. Dosa D, Intrator O, McNicoll L, Cang Y, Teno J. Preliminary derivation of a Nursing Home Confusion Assessment Method based on data from the Minimum Data Set. *J Am Geriatr Soc*. 2007;55(7):1099-1105.
59. Schuurmans MJ, Deschamps PI, Markham SW, Shortridge-Baggett LM, Duursma SA. The measurement of delirium. *Res Theory Nurs Pract*. 2003;17(3):207-224.
60. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. *J Am Geriatr Soc*. 2008;56(5):823-830.