

Focusing on Inattention: The Diagnostic Accuracy of Brief Measures of Inattention for Detecting Delirium

Annachiara Marra, MD, PhD^{1,2}, James C. Jackson, PsyD^{1,3,4,5}, E. Wesley Ely, MD, MPH^{1,3,4,5,6}, Amy J. Graves, MS⁷, John F. Schnelle, PhD^{4,5,6,8}, Robert S. Dittus, MD, MPH^{3,5,8}, Amanda Wilson, MD⁹, Jin H. Han, MD, MSc^{3,6,10*}

¹Division of Allergy, Pulmonary, and Critical Care Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; ²Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples, Federico II, Naples, Italy; ³Center for Health Services Research, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; ⁴Research Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, Tennessee; ⁵Geriatric Research, Education and Clinical Center (GRECC) Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, Tennessee; ⁶Center for Quality Aging, Vanderbilt University Medical Center, Nashville, Tennessee; ⁷Department of Urology, Vanderbilt University Medical Center, Nashville, Tennessee; ⁸Division of General Internal Medicine and Public Health, Vanderbilt University Medical Center, Nashville, Tennessee; ⁹Department of Psychiatry, Vanderbilt University Medical Center, Nashville, Tennessee; ¹⁰Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, Tennessee.

BACKGROUND: Delirium is frequently missed in most clinical settings. Brief delirium assessments are needed.

OBJECTIVE: To determine the diagnostic accuracy of reciting the months of year backwards (MOTYB) from December to July (MOTYB-6) and December to January (MOTYB-12) for delirium as diagnosed by a psychiatrist and to explore the diagnostic accuracies of the following other brief attention tasks: (1) spell the word "LUNCH" backwards, (2) recite the days of the week backwards, (3) 10-letter vigilance "A" task, and (4) 5 picture recognition task.

DESIGN: Preplanned secondary analysis of a prospective observational study.

SETTING: Emergency department located within an academic, tertiary care hospital.

PARTICIPANTS: 234 acutely ill patients who were ≥65 years old.

MEASUREMENTS: The inattention tasks were administered by a physician. The reference standard for delirium was a comprehensive psychiatrist assessment using Diagnostic and Statistical Manual of Mental

Disorders, Fourth Edition, Text Revision criteria. Sensitivities and specificities were calculated.

RESULTS: Making any error on the MOTYB-6 task had a sensitivity of 80.0% (95% confidence interval [CI], 60.9%-91.1%) and specificity of 57.1% (95% CI, 50.4%-63.7%). Making any error on the MOTYB-12 task had a sensitivity of 84.0% (95% CI, 65.4%-93.6%) and specificity of 51.9% (95% CI, 45.2%-58.5%). The best combination of sensitivity and specificity was reciting the days of the week backwards task; if the patient made any error, this was 84.0% (95% CI, 65.4%-93.6%) sensitive and 81.9% (95% CI, 76.1%-86.5%) specific.

CONCLUSION: MOTYB-6 and MOTYB-12 had very good sensitivities but had modest specificities for delirium, limiting their use as a standalone assessment. Reciting the days of the week backwards appeared to have the best combination of sensitivity and specificity for delirium. *Journal of Hospital Medicine* 2018;13:551-557. Published online first March 26, 2018. © 2018 Society of Hospital Medicine

Delirium is an acute neurocognitive disorder¹ that affects up to 25% of older emergency department (ED) and hospitalized patients.²⁻⁴ The relationship between delirium and adverse outcomes is well documented.⁵⁻⁷ Delirium is a strong predictor of increased length of mechanical ventilation, longer intensive care unit and hospital stays, increased risk of falls, long-term cognitive impairment, and mortality.⁸⁻¹³ Delirium is frequently missed by healthcare

professionals^{2,14-16} and goes undetected in up to three out of four patients by bedside nurses and medical practitioners in many hospital settings.^{14,17-22} A significant barrier to recognizing delirium is the absence of brief delirium assessments.

In an effort to improve delirium recognition in the acute care setting, there has been a concerted effort to develop and validate brief delirium assessments. To address this unmet need, 4 'A's Test (4AT),²³ the Brief Confusion Assessment Method (bCAM), and the 3-minute diagnostic assessment for CAM-defined delirium (3D-CAM) are 1- to 3-minute delirium assessments that were validated in acutely ill older patients.²³ However, 1 to 3 minutes may still be too long in busy clinical environments, and briefer (<30 seconds) delirium assessments may be needed.

One potential more-rapid method to screen for delirium is to specifically test for the presence of inattention, which is a

*Address for correspondence: Jin H. Han, MD, MSc, Suite 312, 1313 21st Ave S, Nashville, TN 37232; Telephone: 615-322-0253; Fax: 615-936-1316; E-mail: jin.h.han@vanderbilt.edu

Received: August 4, 2017; Revised: November 17, 2017;

Accepted: December 5, 2017

© 2018 Society of Hospital Medicine DOI 10.12788/jhm.2943

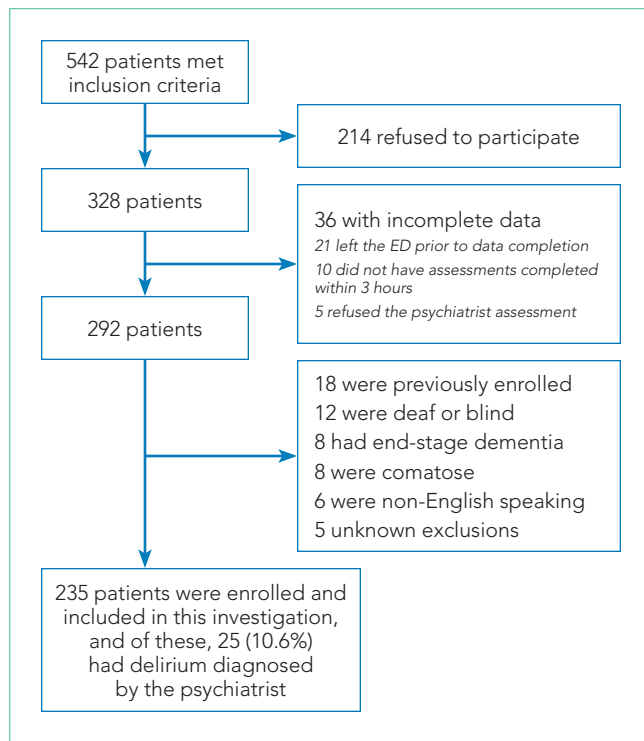


FIG. Enrollment and flow diagram.

cardinal feature of delirium.^{24,25} Inattention can be ascertained by having the patient recite the months backwards, recite the days of the week backwards, or spell a word backwards.²⁶ Recent studies have evaluated the diagnostic accuracy of reciting the months of the year backwards for delirium. O'Regan et al.²⁷ evaluated the diagnostic accuracy of the month of the year backwards from December to July (MOTYB-6) and observed that this task was 84% sensitive and 90% specific for delirium in older patients. However, they performed the reference standard delirium assessments in patients who had a positive MOTYB-6, which can overestimate sensitivity and underestimate specificity (verification bias).²⁸ Fick et al.²⁹ examined the diagnostic accuracy of 20 individual elements of the 3D-CAM and observed that reciting the months of the year backwards from December to January (MOTYB-12) was 83% sensitive and 69% specific for delirium. However, this was an exploratory study that was designed to identify an element of the 3D-CAM that had the best diagnostic accuracy.

To address these limitations, we sought to evaluate the diagnostic performance of the MOTYB-6 and MOTYB-12 for delirium as diagnosed by a reference standard. We also explored other brief tests of inattention such as spelling a word ("LUNCH") backwards, reciting the days of the week backwards, 10-letter vigilance "A" task, and five-picture recognition task.

METHODS

Study Design and Setting

This was a preplanned secondary analysis of a prospective observational study that validated three delirium assessments.^{30,31}

This study was conducted at a tertiary care, academic ED. The local institutional review board (IRB) reviewed and approved this study. Informed consent from the patient or an authorized surrogate was obtained whenever possible. Because this was an observational study and posed minimal risk to the patient, the IRB granted a waiver of consent for patients who were both unable to provide consent and were without an authorized surrogate available in the ED or by phone.

Selection of Participants

We enrolled a convenience sample of patients between June 2010 and February 2012 Monday through Friday from 8 AM to 4 PM. This enrollment window was based upon the psychiatrist's availability. Because of the extensiveness of the psychiatric evaluations, we limited enrollment to one patient per day. Patients who were 65 years or older, not in a hallway bed, and in the ED for less than 12 hours at the time of enrollment were included. We used a 12-hour cutoff so that patients who presented in the evening and early morning hours could be included. Patients were excluded if they were previously enrolled, non-English speaking, deaf or blind, comatose, suffered from end-stage dementia, or were unable to complete all the study assessments. The rationale for excluding patients with end-stage dementia was that diagnosing delirium in this patient population is challenging.

Research assistants approached patients who met inclusion criteria and determined if any exclusion criteria were present. If none of the exclusion criteria were present, then the research assistant reviewed the informed consent document with the patient or authorized surrogate if the patient was not capable of providing consent. If a patient was not capable of providing consent and no authorized surrogate was available, then the patient was enrolled (under the waiver of consent) as long as the patient assented to be a part of the study. Once the patient was enrolled, the research assistant contacted the physician rater and reference standard psychiatrists to approach the patient.

Measures of Inattention

An emergency physician (J.H.H.) who had no formal training in the mental status assessment of elders administered a cognitive battery to the patient, including tests of inattention. The following inattention tasks were administered:

- Spell the word "LUNCH" backwards.³⁰ Patients were initially allowed to spell the word "LUNCH" forwards. Patients who were unable to perform the task were assigned 5 errors.
- Recite the months of the year backwards from December to July.^{23,26,27,30,32} Patients who were unable to perform the task were assigned 6 errors.
- Recite the days of the week backwards.^{23,26,33} Patients who were unable to perform the task were assigned 7 errors.
- Ten-letter vigilance "A" task.³⁴ The patient was given a series of 10 letters ("S-A-V-E-A-H-A-R-T") every 3 seconds and was asked to squeeze the rater's hand every time the patient heard the letter "A." Patients who were unable to perform the task were assigned 10 errors.

- Five-picture recognition task.³⁴ Patients were shown 5 objects on picture cards. Afterwards, patients were shown 10 pictures with the previously shown objects intermingled. The patient had to identify which objects were seen previously in the first 5 pictures. Patients who were unable to perform the task were assigned 10 errors.
- Recite the months of the year backwards from December to January.²⁹ Patients who were unable to perform the task were assigned 12 errors.

Reference Standard for Delirium

A comprehensive consultation-liaison psychiatrist assessment was the reference standard for delirium; the diagnosis of delirium was based on Diagnostic and Statistical Manual of Mental Disorders, *Fourth Edition, Text Revision (DSM-IV-TR)* criteria.³⁵ Three psychiatrists who each had an average of 11 years of clinical experience and regularly diagnosed delirium as part of their daily clinical practice were available to perform these assessments. To arrive at the diagnosis of delirium, they interviewed those who best understood the patient's mental status (eg, the patient's family members or caregivers, *physician, and nurses*). They also reviewed the patient's medical record and radiology and laboratory test results. They performed bedside cognitive testing that included, but was not limited to, the Mini-Mental State Examination, Clock Drawing Test, Luria hand sequencing task, and tests for verbal fluency. A focused neurological examination was also performed (ie, screening for paraphasic errors, tremors, tone, asterixis, frontal release signs, etc.), and they also evaluated the patient for affective lability, hallucinations, and level of alertness. If the presence of delirium was still questionable, then confrontational naming, proverb interpretation or similarities, and assessments for apraxias were performed at the discretion of the psychiatrist. The psychiatrists were blinded to the physician's assessments, and the assessments were conducted within three hours of each other.

Additional Variables Collected

Using medical record review, comorbidity burden, severity of illness, and premorbid cognition were ascertained. The Charlson Comorbidity Index, a weighted index that takes into account the number and seriousness of 19 preexisting comorbid conditions, was used to quantify comorbidity burden; higher scores indicate higher comorbid burden.^{36,37} The Acute Physiology Score of the Acute Physiology and Chronic Health Evaluation II was used to quantify severity of illness.³⁸ This score is based upon the initial values of 12 routine physiologic measurements such as vital sign and laboratory abnormalities; higher scores represent higher severities of illness.³⁸ The medical record was reviewed to ascertain the presence of premorbid cognitive impairment; any documentation of dementia in the patient's clinical problem list or physician history and physical examination from the outpatient or inpatient settings was considered positive. The medical record review was performed by a research assistant and was double-checked for accuracy by one of the investigators (JHH).

Data Analyses

Measures of central tendency and dispersion for continuous variables were reported as medians and interquartile ranges. Categorical variables were reported as proportions. Receiver operating characteristic curves were constructed for each inattention task. Area under the receiver operating characteristic curves (AUC) was reported to provide a global measure of diagnostic accuracy. Sensitivities, specificities, positive likelihood ratios (PLRs), and negative likelihood ratios (NLRs) with their 95% CIs were calculated using the psychiatrist's assessment as the reference standard.³⁹ Cut-points with PLRs greater than 10 (strongly increased the likelihood of delirium) or NLRs less than 0.1 (strongly decreased the likelihood of delirium) were preferentially reported whenever possible.

All statistical analyses were performed with open source R statistical software version 3.0.1 (<http://www.r-project.org/>), SAS 9.4 (SAS Institute, Cary, North Carolina), and Microsoft Excel 2010 (Microsoft Inc., Redmond, Washington).

RESULTS

A total of 542 patients were screened; 214 patients refused to participate, and 93 were excluded, leaving 235 patients. The patient characteristics can be seen in Table 1. Compared with all patients (N = 15,359) who presented to the ED during the study period, enrolled patients were similar in age but more likely to be male, have cardiovascular chief complaints, and be admitted to the hospital. Of those enrolled, 25 (10.6%) were delirious. Delirious patients were older, more likely to be non-white, have a past history of dementia, have a graduate school degree, and have a chief complaint of altered mental status.

Making any error on the MOTYB-6 task had a sensitivity of 80.0% (95% CI, 60.9%-91.1%), specificity of 57.1% (95% CI, 50.4%-63.7%), PLR of 1.87 (95% CI, 1.45-2.40) and NLR of 0.35 (95% CI, 0.16-0.77) for delirium as diagnosed by a psychiatrist. Making any error on the MOTYB-12 task had a sensitivity of 84.0% (95% CI, 65.4%-93.6%), specificity of 51.9% (95% CI, 45.2%-58.5%), PLR of 1.75 (95% CI, 1.40-2.18), and NLR of 0.31 (95% CI, 0.12-0.76) for delirium. The AUCs for the MOTYB-6 and MOTYB-12 tasks were 0.77 and 0.78, respectively, indicating very good diagnostic performance.

The diagnostic performances of the other inattention tasks and additional cutoff values for the MOTYB-6 and MOTYB-12 tasks can be seen in Table 2. Increasing the MOTYB-6 cut-off to two or more errors and MOTYB-12 cut-off to three or more errors increased the specificity to 70.0% and 70.5%, respectively, without decreasing their sensitivity. The best combination of sensitivity and specificity was reciting the days of the week backwards task; if the patient made any error, this was 84.0% (95% CI, 65.4%-93.6%) sensitive and 81.9% (95% CI, 76.1%-86.5%) specific for delirium. The inattention tasks that strongly increased the likelihood of delirium (PLR > 10) were the vigilance "A" and picture recognition tasks. If the patient made two or more errors on the vigilance task or three or more errors on the picture recognition task, then the likelihood of delirium strongly increased, as evidenced by a PLR of 16.80 (95% CI, 8.01-35.23) and 23.10 (95% CI, 7.95-67.12), respectively. No other inattention tasks

TABLE 1. Demographic and Clinical Characteristics of Patients

Patient Characteristics	Enrolled Patients (n = 235)	All Potentially Eligible Patients (N = 15,359)	Nondelirious N = 210	Delirious N = 25
Median age (IQR)	74 (69, 79)	74 (69, 81)	74 (69, 79)	84 (67, 86)
Female gender	107 (45.5%)	8,198 (53.4%)	96 (45.7%)	11 (44.0%)
Nonwhite race	31 (13.9%)	-	24 (11.4%)	7 (28.0%)
Residence		-		
Home	215 (91.5%)		192 (91.4%)	23 (92.0%)
Assisted living	15 (6.4%)		14 (6.7%)	1 (4.0%)
Rehabilitation/postacute	2 (0.9%)		2 (1.0%)	0 (0.0%)
Nursing home	3 (1.3%)		2 (1.0%)	1 (4.0%)
Education		-		
Elementary or below	6 (2.6%)		5 (2.4%)	1 (4.0%)
Middle school	25 (10.6%)		21 (10.0%)	4 (16.0%)
High school	95 (40.4%)		86 (41.0%)	9 (36.0%)
College	71 (30.2%)		65 (31.0%)	6 (24.0%)
Graduate school	38 (16.2%)		33 (15.7%)	5 (20.0%)
Dementia in medical record	17 (7.2%)	-	10 (4.8%)	7 (28.0%)
Median Charlson (IQR)	2 (1, 4)	-	2 (1, 4)	3 (2, 6)
Median APS (IQR)	3 (1, 5)	-	3 (1, 5)	4 (1, 5)
ED chief complaint				
Abdominal pain	8 (3.4%)	854 (5.6%)	7 (3.3%)	1 (4.0%)
Altered mental status	14 (6.0%)	617 (4.0%)	2 (1.0%)	12 (48.0%)
Chest pain	43 (18.3%)	1575 (10.3%)	39 (18.6%)	4 (16.0%)
General weakness	21 (8.9%)	1,101 (7.2%)	18 (8.6%)	3 (12.0%)
Shortness of breath	31 (13.2%)	1,377 (9.0%)	29 (13.8%)	2 (8.0%)
Syncope	14 (6.0%)	422 (2.3%)	14 (6.7%)	0 (0.0%)
Admitted to the hospital	168 (71.5%)	9491 (61.8%)		

NOTE: The APS is part of the APACHE II. Abbreviations: APACHE II, Acute Physiology and Chronic Health Evaluation II; APS, Acute Physiology Score; ED, emergency department; IQR, interquartile range.

were able to achieve a PLR of greater than 10, regardless of what cutoff was used. No inattention task was able to achieve a NLR of less than 0.10, which would have strongly decreased the likelihood of delirium. The best NLRs were if the patient made no errors spelling the word "LUNCH" backwards (NLR, 0.16; 95% CI, 0.04-0.60), no errors on the vigilance "A" task (NLR, 0.18; 95% CI, 0.07-0.43), and no errors on the days of the week backwards task (NLR, 0.20; 95% CI, 0.08-0.48).

DISCUSSION

Delirium is frequently missed by healthcare providers because it is not routinely screened for in the acute care setting. To help address this deficiency of care, we evaluated several brief measures of inattention that take less than 30 seconds to complete. We observed that any errors made on the MOTYB-6 and MOTYB-12 tasks had very good sensitivities (80% and 84%) but were limited by their modest specificities (approximately 50%) for delirium. As a result, these assessments have limited clinical utility as standalone delirium screens. We also explored other commonly used brief measures of inattention and at a variety of error cutoffs. Reciting the days of the week backwards ap-

peared to best balance sensitivity and specificity. None of the inattention measures could convincingly rule out delirium (NLR < 0.10), but the vigilance "A" and picture recognition tasks may have clinical utility in ruling in delirium (PLR > 10). Overall, all the inattention tasks, including MOTYB-6 and MOTYB-12, had very good diagnostic performances based upon their AUC. However, achieving a high sensitivity often had to be sacrificed for specificity or, alternatively, achieving a high specificity had to be sacrificed for sensitivity.

Inattention has been shown to be the cardinal feature for delirium,⁴⁰ and its assessment using cognitive testing has been recommended to help identify the presence of delirium according to an expert consensus panel.²⁶ The diagnostic performance of the MOTYB-12 observed in our study is similar to a study by Fick et al., who reported that MOTYB-12 had very good sensitivity (83%) but had modest specificity (69%) with a cutoff of 1 or more errors. Hendry et al. observed that the MOTYB-12 was 91% sensitive and 50% specific using a cutoff of 4 or more errors. With regard to the MOTYB-6, our reported specificity was different from what was observed by O'Regan et al.²⁷ Using 1 or more errors as a cutoff, they observed a much higher specificity for delir-

TABLE 2. Diagnostic Accuracy of Inattention Measures

Inattention Test	AUC	Sensitivity	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)
MOTYB-6					
1 or more errors	0.77	80.0% (60.9%-91.1%)	57.1% (50.4%-63.7%)	1.87 (1.45-2.40)	0.35 (0.16-0.77)
2 or more errors		80.0% (60.9%-91.1%)	70.0% (63.5%-75.8%)	2.67 (2.01-3.55)	0.29 (0.13-0.63)
3 or more errors		60.0% (40.7%-76.6%)	80.0% (74.1%-84.9%)	3.00 (1.97-4.56)	0.50 (0.31-0.81)
4 or more errors		56.0% (37.1%-73.3%)	86.2% (80.9%-90.2%)	4.06 (2.50-6.58)	0.51 (0.33-0.88)
5 or more errors		52.0% (33.5%-70.0%)	89.6% (83.6%-92.2%)	4.55 (2.67-7.75)	0.54 (0.36-0.82)
6 errors or UTD		44.0% (26.7%-62.9%)	91.9% (87.4%-94.8%)	4.48 (2.46-8.17)	0.62 (0.44-0.88)
MOTYB-12					
1 or more errors	0.78	84.0% (65.4%-93.6%)	51.9% (45.2%-58.5%)	1.75 (1.40-2.18)	0.31 (0.12-0.76)
2 or more errors		84.0% (65.4%-93.6%)	65.2% (58.6%-71.4%)	2.42 (1.88-3.11)	0.25 (0.10-0.61)
3 or more errors		84.0% (65.4%-93.6%)	70.5% (64.0%-76.2%)	2.85 (2.17-3.73)	0.23 (0.09-0.56)
4 or more errors		76.0% (56.6%-88.5%)	73.8% (67.5%-79.3%)	2.90 (2.11-3.98)	0.33 (0.16-0.66)
12 errors or UTD		48.0% (30.0%-66.5%)	90.0% (85.2%-93.4%)	4.80 (2.70-8.53)	0.58 (0.40-0.84)
Spelling "LUNCH" backwards					
1 or more errors	0.81	92.0% (75.0%-97.8%)	50.5% (43.8%-57.2%)	1.86 (1.55-2.22)	0.16 (0.04-0.60)
2 or more errors		88.0% (70.0%-95.8%)	64.8% (58.1%-70.9%)	1.57 (1.32-1.87)	0.27 (0.09-0.80)
3 or more errors		80.0% (60.9%-91.1%)	73.3% (67.0%-78.9%)	3.00 (2.23-4.04)	0.27 (0.12-0.60)
4 or more errors		56.0% (37.1%-73.3%)	81.9% (76.1%-86.5%)	3.39 (2.15-5.33)	0.53 (0.34-0.82)
5 errors or UTD		56.0% (37.1%-73.3%)	87.1% (81.9%-91.0%)	4.56 (2.78-7.49)	0.50 (0.32-0.78)
Days of the week backwards					
1 or more errors	0.85	84.0% (65.4%-93.6%)	81.9% (76.1%-86.5%)	4.64 (3.32-6.49)	0.20 (0.08-0.48)
2 or more errors		60.0% (40.7%-76.6%)	90.5% (85.8%-93.8%)	6.30 (3.72-10.66)	0.44 (0.27-0.72)
3 or more errors		56.0% (37.1%-73.3%)	91.4% (86.9%-94.5%)	6.53 (3.72-11.46)	0.48 (0.31-0.75)
4 or more errors		56.0% (37.1%-73.3%)	91.9% (87.4%-94.9%)	6.92 (3.90-12.27)	0.48 (0.31-0.75)
5 or more errors		56.0% (37.1%-73.3%)	92.4% (88.0%-95.3%)	7.35 (4.09-13.20)	0.48 (0.31-0.74)
6 or more errors		56.0% (37.1%-73.3%)	92.9% (88.6%-95.6%)	7.84 (4.31-14.27)	0.47 (0.30-0.74)
7 errors or UTD		44.0% (26.7%-62.9%)	95.2% (91.5%-97.4%)	9.24 (4.37-19.55)	0.59 (0.41-0.83)
10-letter vigilance "A" task					
1 or more errors	0.84	84.0% (65.4%-93.6%)	63.8% (57.1%-70.0%)	9.80 (6.10-15.74)	0.18 (0.07-0.43)
2 or more errors		64.0% (44.5%-79.8%)	91.4% (85.9%-94.5%)	16.80 (8.01-35.23)	0.37 (0.22-0.63)
3 or more errors		60.0% (40.7%-76.7%)	96.2% (92.7%-98.1%)	126.00 (17.37-913.82)	0.40 (0.25-0.65)
Picture recognition task					
1 or more errors	0.81	64.0% (44.5%-79.8%)	80.0% (74.1%-84.9%)	1.40 (1.02-1.92)	0.66 (0.39-1.13)
2 or more errors		60.0% (40.7%-76.7%)	93.8% (89.7%-96.4%)	9.69 (5.23-17.95)	0.43 (0.26-0.69)
3 or more errors		44.0% (27.7%-62.9%)	98.1% (95.2%-99.3%)	23.10 (7.95-67.12)	0.57 (0.40-0.81)

NOTE: Sensitivities, specificities, PLRs, and NLRs of several brief measures of inattention with their AUC. Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; MOTYB-6, months of the year backwards December to July; MOTYB-12, months of the year backwards December to January; NLR, negative likelihood ratio; PLR, positive likelihood ratio; UTD, unable to do.

ium than we did (90% vs 57%). Discordant observations regarding the diagnostic accuracy for other inattention tasks also exist. We observed that making any error on the days of the week backwards task was 84% sensitive and 82% specific for delirium, whereas Fick et al. observed a sensitivity and specificity of 50% and 94%, respectively. For the vigilance "A" task, we observed that making two or more errors over a series of 10 letters was 64.0% sensitive and 91.4% specific for delirium, whereas Pompei et al.⁴¹ observed that making two or more errors over a series of 60 letters was 51% sensitive and 77% specific for delirium.

The abovementioned discordant findings may be driven by spectrum bias, wherein the sensitivities and specificities for each inattention task may differ in different subgroups. As a result, differences in the age distribution, proportion of college graduates, history of dementia, and susceptibility to delirium

can influence overall sensitivity and specificity. Objective measures of delirium, including the inattention screens studied, are particularly prone to spectrum bias.^{31,34} However, the strength of this approach is that the assessment of inattention becomes less reliant upon clinical judgment and allows it to be used by raters from a wide range of clinical backgrounds. On the other hand, a subjective interpretation of these inattention tasks may allow the rater to capture the subtleties of inattention (ie, decreased speed of performance in a highly intelligent and well-educated patient without dementia). The disadvantage of this approach, however, is that it is more dependent on clinical judgment and may have decreased diagnostic accuracy in those with less clinical experience or with limited training.^{14,42,43} These factors must be carefully considered when determining which delirium assessment to use.

Additional research is required to determine the clinical utility of these brief inattention assessments. These findings need to be further validated in larger studies, and the optimal cutoff of each task for different subgroup of patients (eg, demented vs nondemented) needs to be further clarified. It is not completely clear whether these inattention tests can serve as standalone assessments. Depending on the cutoff used, some of these assessments may have unacceptable false negative or false positive rates that may lead to increased adverse patient outcomes or increased resource utilization, respectively. Additional components or assessments may be needed to improve the diagnostic accuracy of these assessments. In addition to understanding these inattention assessments' diagnostic accuracies, their ability to predict adverse outcomes also needs to be investigated. While a previous study observed that making any error on the MOTYB-12 task was associated with increased physical restraint use and prolonged hospital length of stay,⁴⁴ these assessments' ability to prognosticate long-term outcomes such as mortality or long-term cognition or function need to be studied. Lastly, studies should also evaluate how easily implementable these assessments are and whether improved delirium recognition leads to improved patient outcomes.

This study has several notable limitations. Though planned a priori, this was a secondary analysis of a larger investigation designed to validate three delirium assessments. Our sample size was also relatively small, causing our 95% CIs to overlap in most cases and limiting the statistical power to truly determine whether one measure is better than the other. We also asked the patient to recite the months backwards from December to July as well as recite the months backwards from December to January. It is possible that the patient may have performed better at going from December to January because of learning effect. Our reference standard for delirium was based upon DSM-IV-TR criteria. The new DSM-V criteria may be more restrictive and may slightly change the sensitivities and specificities of the inattention tasks. We enrolled a convenience sample and enrolled patients who were more likely to be male, have cardiovascular chief complaints, and be admitted to the hospital; as a result, selection bias may have been introduced. Lastly, this study was conducted in a single center and enrolled patients who were 65 years and older. Our findings may not be generalizable to other settings and in those who are less than 65 years of age.

CONCLUSIONS

The MOTYB-6 and MOTYB-12 tasks had very good sensitivities but modest specificities (approximately 50%) using any error made as a cutoff; increasing cutoff to 2 errors and 3 errors, respectively, improved their specificities (approximately 70%) with minimal impact to their sensitivities. Reciting the days of the week backwards, spelling the word "LUNCH" backwards, and the 10-letter vigilance "A" task appeared to perform the best in ruling out delirium but only moderately decreased the likelihood of delirium. The 10-letter Vigilance "A" and picture recognition task appeared to perform the best in ruling in de-

lirium. Days of the week backwards appeared to have the best combination of sensitivity and specificity.

Disclosure: The authors report no financial conflicts of interest.

Funding: This study was funded by the Emergency Medicine Foundation Career Development Award, National Institutes of Health K23AG032355, and National Center for Research Resources, Grant UL1 RR024975-01.

References

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, DC: American Psychiatric Association; 2013.
2. Hustey FM, Meldon SW, Smith MD, Lex CK. The effect of mental status screening on the care of elderly emergency department patients. *Ann Emerg Med*. 2003;41(5):678-684.
3. Inouye SK, Rushing JT, Foreman MD, Palmer RM, Pompei P. Does delirium contribute to poor hospital outcomes? A three-site epidemiologic study. *J Gen Intern Med*. 1998;13(4):234-242.
4. Pitkala KH, Laurila JV, Strandberg TE, Tilvis RS. Prognostic significance of delirium in frail older people. *Dement Geriatr Cogn Disord*. 2005;19(2-3):158-163.
5. Han JH, Shintani A, Eden S, et al. Delirium in the emergency department: an independent predictor of death within 6 months. *Ann Emerg Med*. 2010;56(3):244-252.
6. Gross AL, Jones RN, Habtemariam DA, et al. Delirium and long-term cognitive trajectory among persons with dementia. *Arch Intern Med*. 2012;172(17):1324-1331.
7. Davis DH, Muniz-Terrera G, Keage H, et al. Delirium is a strong risk factor for dementia in the oldest-old: a population-based cohort study. *Brain*. 2012;135(Pt 9):2809-2816.
8. Ely EW, Baker AM, Dunagan DP, et al. Effect on the duration of mechanical ventilation of identifying patients capable of breathing spontaneously. *N Engl J Med*. 1996;335(25):1864-1869.
9. Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004;291(14):1753-1762.
10. Lin SM, Liu CY, Wang CH, et al. The impact of delirium on the survival of mechanically ventilated patients. *Crit Care Med*. 2004;32(11):2254-2259.
11. Salluh JI, Wang H, Schneider EB, et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. *BMJ*. 2015;350:h2538.
12. Klein Klouwenberg PM, Zaal IJ, Spitoni C, et al. The attributable mortality of delirium in critically ill patients: prospective cohort study. *BMJ*. 2014;349:g6652.
13. Han JH, Vasilevskis EE, Chandrasekhar R, et al. Delirium in the Emergency Department and Its Extension into Hospitalization (DELINEATE) Study: Effect on 6-month Function and Cognition. *J Am Geriatr Soc*. 2017;65(6):1333-1338.
14. Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM Jr. Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. *Arch Intern Med*. 2001;161(20):2467-2473.
15. Han JH, Zimmerman EE, Cutler N, et al. Delirium in older emergency department patients: recognition, risk factors, and psychomotor subtypes. *Acad Emerg Med*. 2009;16(3):193-200.
16. Elie M, Cole MG, Primeau FJ, Bellavance F. Delirium risk factors in elderly hospitalized patients. *J Gen Intern Med*. 1998;13(3):204-212.
17. Spronk PE, Riekerk B, Hofhuis J, Rommes JH. Occurrence of delirium is severely underestimated in the ICU during daily care. *Intensive Care Med*. 2009;35(7):1276-1280.
18. van Eijk MM, van Marum RJ, Klijn IA, de Wit N, Kesecioglu J, Slooter AJ. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med*. 2009;37(6):1881-1885.
19. Devlin JW, Fong JJ, Schumaker G, O'Connor H, Ruthazer R, Garpestad E. Use of a validated delirium assessment tool improves the ability of physicians to identify delirium in medical intensive care unit patients. *Crit Care Med*. 2007;35(12):2721-2724.
20. Han JH, Eden S, Shintani A, et al. Delirium in Older Emergency Department Patients Is an Independent Predictor of Hospital Length of Stay. *Acad Emerg Med*. 2011;18(5):451-457.
21. Pun BT, Gordon SM, Peterson JF, et al. Large-scale implementation of sedation and delirium monitoring in the intensive care unit: a report from two medical centers. *Crit Care Med*. 2005;33(6):1199-1205.

22. Grossmann FF, Hasemann W, Graber A, Bingisser R, Kressig RW, Nickel CH. Screening, detection and management of delirium in the emergency department - a pilot study on the feasibility of a new algorithm for use in older emergency department patients: the modified Confusion Assessment Method for the Emergency Department (mCAM-ED). *Scand J Trauma Resusc Emerg Med.* 2014;22:19.
23. Marcantonio ER, Ngo LH, O'Connor M, et al. 3D-CAM: Derivation and Validation of a 3-Minute Diagnostic Interview for CAM-Defined Delirium: A Cross-sectional Diagnostic Test Study. *Ann Intern Med.* 2014;161(8):554-561.
24. Blazer DG, van Nieuwenhuizen AO. Evidence for the diagnostic criteria of delirium: an update. *Curr Opin Psychiatry.* 2012;25(3):239-243.
25. Meagher DJ, MacLulich AM, Laurila JV. Defining delirium for the International Classification of Diseases, 11th Revision. *J Psychosom Res.* 2008;65(3):207-214.
26. Huang LW, Inouye SK, Jones RN, et al. Identifying indicators of important diagnostic features of delirium. *J Am Geriatr Soc.* 2012;60(6):1044-1050.
27. O'Regan NA, Ryan DJ, Boland E, et al. Attention! A good bedside test for delirium? *J Neurol Neurosurg Psychiatry.* 2014;85(10):1122-1131.
28. Richardson ML, Petscavage JM. Verification bias: an under-recognized source of error in assessing the efficacy of MRI of the meniscii. *Acad Radiol.* 2011;18(11):1376-1381.
29. Fick DM, Inouye SK, Guess J, et al. Preliminary development of an ultrabrief two-item bedside test for delirium. *J Hosp Med.* 2015;10(10):645-650.
30. Han JH, Wilson A, Vasilevskis EE, et al. Diagnosing delirium in older emergency department patients: validity and reliability of the delirium triage screen and the brief confusion assessment method. *Ann Emerg Med.* 2013;62(5):457-465.
31. Han JH, Wilson A, Graves AJ, et al. Validation of the Confusion Assessment Method for the Intensive Care Unit in older emergency department patients. *Acad Emerg Med.* 2014;21(2):180-187.
32. Inouye SK. Delirium in hospitalized older patients. *Clin Geriatr Med.* 1998;14(4):745-764.
33. Hamrick I, Hafiz R, Cummings DM. Use of days of the week in a modified mini-mental state exam (M-MMSE) for detecting geriatric cognitive impairment. *J Am Board Fam Med.* 2013;26(4):429-435.
34. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA.* 2001;286(21):2703-2710.
35. American Psychiatric Association. Task Force on DSM-IV. *Diagnostic and statistical manual of mental disorders: DSM-IV-TR.* 4th ed. Washington, DC: American Psychiatric Association; 2000.
36. Murray SB, Bates DW, Ngo L, Ufberg JW, Shapiro NI. Charlson Index is associated with one-year mortality in emergency department patients with suspected infection. *Acad Emerg Med.* 2006;13(5):530-536.
37. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373-383.
38. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med.* 1985;13(10):818-829.
39. 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.
40. Blazer DG, Wu LT. The epidemiology of substance use and disorders among middle aged and elderly community adults: national survey on drug use and health. *Am J Geriatr Psychiatry.* 2009;17(3):237-245.
41. Pompei P, Foreman M, Cassel CK, Alessi C, Cox D. Detecting delirium among hospitalized older patients. *Arch Intern Med.* 1995;155(3):301-307.
42. Lemiengre J, Nelis T, Joosten E, et al. Detection of delirium by bedside nurses using the confusion assessment method. *J Am Geriatr Soc.* 2006;54(4):685-689.
43. Ryan K, Leonard M, 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.
44. Yevchak AM, Doherty K, Archambault EG, Kelly B, Fonda JR, Rudolph JL. The association between an ultrabrief cognitive screening in older adults and hospital outcomes. *J Hosp Med.* 2015;10(10):651-657.