

BRIEF REPORTS

Validation of the Delirium Observation Screening Scale in a Hospitalized Older Population

Katherine Gavinski, BA¹, Ryan Carnahan, PharmD, MS², Michelle Weckmann, MS, MD^{1,3*}

¹University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, Iowa; ²Department of Epidemiology, University of Iowa College of Public Health, Iowa City, Iowa; ³Department of Family Medicine, University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Delirium is challenging to diagnose in older populations. It is often reversible, and when detected, treatment can improve patient outcomes. Delirium detection currently relies on trained staff to conduct neurocognitive interviews. The Delirium Observation Screening Scale (DOS) is a screen designed to allow faster, easier identification of delirium. In this validation study, conducted at an academic tertiary care center, we attempted to determine the accuracy of the DOS as a delirium screening tool in hospitalized patients

over 64 years old. We compared DOS results to a validated delirium diagnostic tool, the Delirium Rating Scale-Revised-98. We also assess the user-friendliness of the DOS by nurses via electronic survey. In 101 assessments of 54 patients, the DOS had sensitivity of 90% and specificity of 91% for delirium. The DOS is an accurate and easy way to screen for delirium in older inpatients. *Journal of Hospital Medicine* 2016;11:494–497. © 2016 Society of Hospital Medicine

Delirium is a rapidly developing, fluctuating disturbance in consciousness, caused by a medical condition. The diagnosis of delirium is often missed, potentiating negative outcomes.^{1,2} Regular delirium screening by nurses results in increased recognition and treatment.³ Although multiple screening tools exist, many are cumbersome to execute. Efforts have been made to shorten them, but although the screening tools may predict adverse outcomes, there are concerns about their specificity.^{1,2,4–6} The Delirium Observation Screening Scale⁷ (DOS) is a brief screening tool based on observation. It has been validated in several patient populations, but no published studies have taken place in the United States or have focused on an older, general medicine, inpatient population. Given the low numbers of patients in earlier validation studies, the effectiveness of the DOS for screening hospitalized, older patients is not yet fully established.

This study aimed to determine the ability of the DOS to screen hospitalized, older patients for delirium compared to a validated delirium diagnostic tool, the Delirium Rating Scale-Revised-98 (DRS-R-98).⁸ In addition, DOS acceptability, ease of use, and benefit were explored by surveying nurses.

METHODS

Participants

After institutional review board approval, participants were selected by convenience sample from general medicine inpatients at a large, tertiary care, academic hospital. Eligible patients were age 65 years or older, admitted to a medicine inpatient unit, and spoke English. If participants were unable to consent, consent was obtained from the participant's legally authorized representative.

Delirium Observation Screening Scale

The DOS is a 13-point screen for delirium, based on the *Diagnostic and Statistical Manual of Mental Disorders IV* delirium criteria, designed to be completed by a nurse (see Supporting Information, Appendix 1, in the online version of this article). Responses are dichotomous. Scores ≥ 3 were considered positive delirium screens.⁷

Nurses on medicine units attended educational inservices on delirium recognition and use of the DOS. The DOS was embedded in the electronic medical record (EMR) and nurses are electronically prompted to chart DOS results every 12 hours for patients, age 65 years or older. Nursing staff utilized the DOS for 1 year prior to study start.

DRS-R-98

The DRS-R-98 was used as the study reference standard.⁸ Scores ≥ 15 are indicative of delirium.⁹ All assessments were performed by a medical student (K.G.) trained to administer the DRS-R-98.

Data Collection

After consent, hospitalized participants were evaluated daily (Monday–Friday) using the DRS-R-98. Enrollment took place over a 10-week period. Nurses and researchers were blinded to other delirium assessment

*Address for correspondence and reprint requests: Michelle T. Weckmann, MD, Department of Family Medicine, 01102 PFP, Iowa City, IA 52242-1000; Telephone: 319-384-7000; Fax: 319-384-7822; E-mail: michelle-weckmann@uiowa.edu

Additional Supporting Information may be found in the online version of this article.

Received: October 9, 2015; Revised: January 18, 2016; Accepted: February 13, 2016

2016 Society of Hospital Medicine DOI 10.1002/jhm.2580

Published online in Wiley Online Library (Wileyonlinelibrary.com).

results until after participant discharge. Following discharge, additional data were collected from the EMR: age, gender, cognitive comorbidities, and nurse-charted DOS score. Cognitive comorbidities were classified as “no impairment,” “dementia,” or “cognitive impairment” based on the problem list and admission note. A psychiatrist (M.W.) confirmed questions of cognitive impairment.

The DOS score closest in time, within 24 hours of DRS-R-98 assessment, was used for comparison. If a DOS score was not charted within 24 hours of the DRS-R-98 evaluation, that assessment was excluded. Partial DRS-R-98 assessments were included only if there was enough information to classify a subject as delirious or not.

Nursing Survey

A 13-question nursing survey was developed and consisted of demographic, Likert-style, and multiple-choice questions, with opportunities for open-ended responses (see Supporting Information, Appendix 2, in the online version of this article). Survey design followed similar surveys investigating staff experiences and clinical functionality of other brief delirium screening tools, such as the Confusion Assessment Method for the Intensive Care Unit.^{10,11} The survey was distributed by e-mail to 435 nurses on 16 units. Coffee gift cards were raffled as participation incentive.

Statistical Analysis

Statistical analysis was completed using SPSS (IBM, Armonk, NY) and SAS (SAS Institute, Inc., Cary, NC) software. DOS results were compared to the DRS-R-98, and validity statistics were calculated for delirium. Confidence intervals were calculated using the Clopper-Pearson method for binomial data. The Spearman rank correlation coefficient between DOS and DRS-98 score was calculated. PROC LOGISTIC (SAS Institute, Inc.) modeled the relationship between positive DOS screens and delirium and created a receiver operating characteristic (ROC) curve using continuous DOS score to predict delirium. Because these models did not control for multiple observations per individual, PROC GENMOD (SAS Institute, Inc.) was used to confirm the relationship between a positive DOS screen and delirium using a marginal logistic regression model accounting for repeated measures. In addition, we selected 10 random samples of 1 observation per person, and validity statistics were calculated for each sample.

The nursing survey results were analyzed using descriptive statistics. Open-ended comments were reviewed in aggregate.

RESULTS

Participant Characteristics

Fifty-four participants enrolled in the study. Fifty-three were able to complete 1 DRS-R-98 and

TABLE 1. Participant Characteristics

Characteristic	No Delirium, n = 45	Delirium, n = 8*
Age, y		
65–74, n = 26	22	4
75–84, n = 15	13	2
85+, n = 12	10	2
Age, y, mean (SD) [range]	77 (10) [65–92]	76 (8.6) [65–92]
Gender		
Female, n = 33	28	5
Male, n = 20	17	3
Cognitive status per chart		
No impairment, n = 45	43	2
Cognitive impairment without dementia, n = 5	1	4
Dementia, n = 3	1	2

*Any Delirium Rating Scale-Revised-98 score ≥ 15 .

comprise the study sample (Table 1). Participants completed 1 to 5 daily DRS-R-98 assessments (mean, 1.94; standard deviation [SD], 0.90; mean length of admission, 6.06 days). Of the 105 DRS-R-98 assessments, 101 were classifiable for delirium. Of the 101 DRS-R-98 assessments classifiable for delirium, 100 had a corresponding DOS score within 24 hours. Participant characteristics are listed in Table 1. Eight of the 53 participants (15%) had at least 1 positive DRS-R-98. Overall, 10 of the 101 delirium assessments diagnosed delirium (DRS-R-98 score ≥ 15).

DOS Validity

The mean and standard deviation of delirium screening scores are as follows: DRS-R-98 (mean, 6.13; SD, 4.74; range, 0–20) and DOS (mean, 1.22; SD, 2.37; range, 0–9). The Spearman correlation coefficient between DOS and DRS-R-98 scores was 0.58. DOS had a sensitivity of 90% (95% confidence interval [CI]: 56%–100%) and specificity of 91% (95% CI: 83%–96%) compared to the DRS-98-R standard. There was only 1 false negative DOS screen out of 83 negative assessments (negative predictive value = 99%, 95% CI: 93%–100%). Out of the 17 positive assessments, 9 were true positives (positive predictive value = 53%, 95% CI: 28%–77%), and 7 scored in the subsyndromal range for delirium (DRS-R-98 score 8–14).

In analyses using 10 samples, with 1 randomly selected observation per person, the mean sensitivity was 84.6%, ranging from 80% (95% CI: 28%–99%) to 87.5% (95% CI: 47%–100%). The mean specificity in these samples was 92%, ranging from 87% (95% CI: 74%–95%) to 96% (95% CI: 85%–99%).

Logistic Regression Models

All models confirmed that positive DOS screens significantly predicted delirium. The traditional logistic regression model produced an odds ratio (OR) estimate of 92 (95% CI: 10–824, $P < 0.0001$) for a positive DOS screen predicting delirium. The marginal logistic regression model accounting for repeated

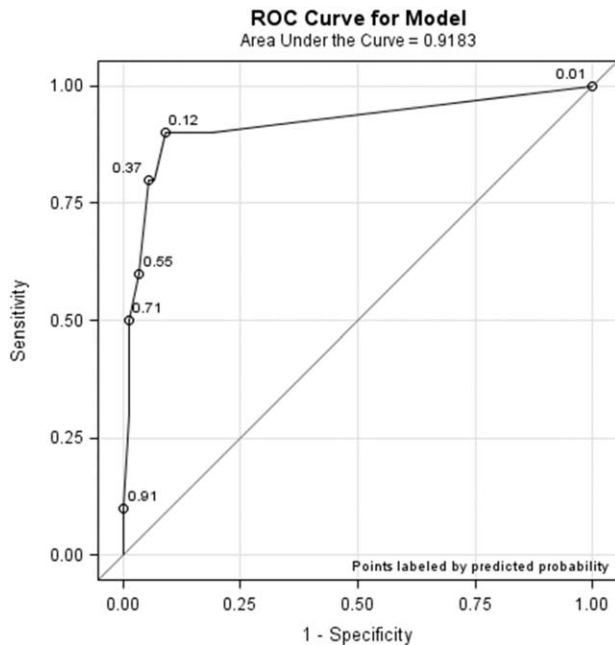


FIG. 1. The Delirium Observation Screening Scale receiver operating characteristic (ROC) curve.

measures produced a consistent estimate (OR: 93, 95% CI: 11-800, $P < 0.0001$). Continuous DOS scores predicted delirium (OR: 2.1, 95% CI: 1.5-2.9, $P < 0.0001$), and the ROC curve supported the cutoff of $\text{DOS} \geq 3$, corresponding to a predicted probability of 0.12 (Figure 1).

Nursing Survey

The nursing survey had a response rate of 23% ($N = 98$). The most robust results related to DOS administration were 87% ($N = 83$) of nurses were confident in DOS administration, 92% ($N = 86$) could complete the DOS in under 3 minutes, and 79% ($N = 74$) agreed that performing the DOS is easy. There was less agreement on the value of the DOS; 37% agreed that the DOS is worth the time to perform, 25% agreed that the DOS enhances patient care, and 36% agreed that the DOS provides valuable information for patient care. Over half the nurses (55%) reported that they perform the DOS 75% to 100% of the prompted times, and 62% stated if the DOS was no longer required, they would not use it. Open-ended questions generated a wide range of responses, from supportive to critical of delirium screening and the DOS (see Supporting Information, Appendix 3, in the online version of this article).

DISCUSSION

This study demonstrated the effectiveness, efficiency, and ease of use of the DOS as a delirium screening tool. The DOS exhibited high sensitivity (90%) and specificity (91%). Similar to previous findings, the positive predictive value was only 53%, but the negative predictive value was 99%.¹² These results support

that the DOS is consistently able to rule out delirium, with only 1 false negative in this study.

Nursing responses regarding user-friendliness are consistent with other studies; however, there was a knowledge gap related to how positive delirium screens can inform and change care for patients.⁷ Education is a known barrier to integrating delirium screening tools secondary to the need for regular and extensive education, frequent reminders to screen, and regular evaluations of assessment quality.^{11,13-15} Developing guidelines for responding to positive DOS screens and documenting its impact on care may incentivize use.

Study strengths include strong evaluator consistency, blinding of evaluator and nurses, and responses from a broad range of nurses (14 of 16 units represented). Additionally, this study demonstrated the efficacy and ease of use of an EMR-prompted delirium screen. However, this study had several limitations, including a small sample size and a low incidence of delirium. The lower incidence is likely secondary to selection bias that resulted from difficulty consenting delirious subjects. The discordant time between DOS and DRS-R-98 assessments may have also influenced results; however, inclusion of data from the previous 8 to 24 hours in both tools makes the temporal separation of assessments less impactful.

The ability of the DOS to accurately identify patients at high risk of delirium is useful for health-care staff. Future work will include nurse and physician education to emphasize delirium understanding, the importance of regular screening, and the use of nonpharmacological interventions. Additional studies will include examination of the interventions and outcomes of patients who screen positive for delirium to determine the long-term impact of delirium screening.

Acknowledgements

The study authors would like to thank the University of Iowa Hospitals and Clinics, the Department of Family Medicine, and the University of Iowa Hospitals and Clinics nursing managers and nursing staff.

Disclosures: This study was funded by the Summer Research Fellowship program sponsored by the University of Iowa Carver College of Medicine. Drs. Weckmann and Carnahan were supported by the Health Resources and Services Administration, Iowa Geriatric Education Center (UB4 HP19054) as well as the US Department of Health and Human Services, Agency for Healthcare Research and Quality (AHRQ 1 R18 HS022666-01).

References

1. Featherstone I, Hopton A, Siddiqi N. An intervention to reduce delirium in care homes. *Nurs Older People*. 2010;22(4):16-21.
2. Devlin JW, Fong JJ, Howard EP, et al. Assessment of delirium in the intensive care unit: nursing practices and perceptions. *Am J Crit Care*. 2008;17(6):555-565.
3. Vasilevskis EE, Morandi A, Boehm L, et al. Delirium and sedation recognition using validated instruments: reliability of bedside intensive care unit nursing assessments from 2007 to 2010. *J Am Geriatr Soc*. 2011;59(suppl 2):S249-S255.
4. 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.
5. Yevchak AM, Doherty K, Archambault EG, et al. The association between an ultrabrief cognitive screening in older adults and hospital outcomes. *J Hosp Med*. 2015;10(10):651-657.

6. Zadavec FJ, Tien L, Robertson-Dick BJ, et al. Comparison of mental-status scales for predicting mortality on the general wards. *J Hosp Med.* 2015;10(10):658–663.
7. Schuurmans MJ, Shortridge-Baggett LM, Duursma SA. The Delirium Observation Screening Scale: a screening instrument for delirium. *Res Theory Nurs Pract.* 2003;17(1):31–50.
8. Trzepacz PT. Validation of the Delirium Rating Scale-Revised-98: comparison with the Delirium Rating Scale and the Cognitive Test for Delirium. *J Neuropsychiatry Clin Neurosci.* 2001;13(2):229–242.
9. Franco JG, Trzepacz PT, Mejia MA, et al. Factor analysis of the Colombian translation of the Delirium Rating Scale (DRS), Revised-98. *Psychosomatics.* 2009;50(3):255–262.
10. Soja SL, Pandharipande PP, Fleming SB, et al. Implementation, reliability testing, and compliance monitoring of the Confusion Assessment Method for the Intensive Care Unit in trauma patients. *Intensive Care Med.* 2008;34(7):1263–1268.
11. Riekerk B, Pen EJ, Hofhuis JG, et al. Limitations and practicalities of CAM-ICU implementation, a delirium scoring system, in a Dutch intensive care unit. *Intensive Crit Care Nurs.* 2009;25(5):242–249.
12. Gemert van LA, Schuurmans MJ. The Neecham Confusion Scale and the Delirium Observation Screening Scale: capacity to discriminate and ease of use in clinical practice. *BMC Nurs.* 2007;6:3.
13. Schuurmans MJ, Duursma SA, Shortridge-Baggett LM. Early recognition of delirium: review of the literature. *J Clin Nurs.* 2001;10(6):721–729.
14. Gesin G, Russell BB, Lin AP, et al. Impact of a delirium screening tool and multifaceted education on nurses' knowledge of delirium and ability to evaluate it correctly. *Am J Crit Care.* 2012;21(1):e1–e11.
15. Devlin JW, Brummel NE, Al-Qadheeb NS. Optimising the recognition of delirium in the intensive care unit. *Best Pract Res Clin Anaesthesiol.* 2012;26(3):385–393.