

Ultrabrief Screens for Detecting Delirium in Postoperative Cognitively Intact Older Adults

Andrea Yevchak Sillner, PhD^{1*} and Long Ngo, PhD², Yoojin Jung, PhD², Sharon K Inouye, MD, MPH^{2,3}, Marie Boltz, PhD¹, Douglas Leslie, PhD⁴, Edward R Marcantonio, MD, SM^{2**} and Donna M Fick, PhD^{1**}

¹College of Nursing, The Pennsylvania State University, University Park, Pennsylvania; ²Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; ³Aging Brain Center, Marcus Institute for Aging Research, Hebrew SeniorLife, Boston, Massachusetts; ⁴Department of Public Health Sciences, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania.

The authors' sought to develop an ultrabrief screen for postoperative delirium in cognitively intact patients older than 70 years undergoing major elective surgery. All possible combinations of one-, two- and three-item screens and their sensitivities, specificities, and 95% confidence intervals were calculated and compared with the delirium reference standard Confusion Assessment Method (CAM). Among the 560 participants (mean age, 77 years; 58% women), delirium occurred in 134 (24%). We considered 1,100 delirium assessments from

postoperative days 1 and 2. The screen with the best overall performance consisted of three items: (1) Patient reports feeling confused, (2) Months of the year backward, and (3) "Does the patient appear sleepy?" with sensitivity of 92% and specificity of 72%. This brief, three-item screen rules out delirium quickly, identifies a subset of patients who require further testing, and may be an important tool to improve recognition of postoperative delirium. *Journal of Hospital Medicine* 2020;15:544-547. © 2020 Society of Hospital Medicine

Delirium is the most common postsurgical complication for older adults, with incidence of 15%-54%, depending on surgery type.¹ Increasing numbers of older adults are undergoing surgery²; and those who develop delirium experience negative consequences including longer lengths of stay, higher likelihood of institutional discharge, and increased morbidity and mortality.³ The American Geriatrics Society Expert Panel on Postoperative Delirium in Older Adults and the European Society of Anaesthesiology⁴ recommend routine screening for delirium in those at risk.

Ultrabrief screens are designed to rule out delirium quickly and identify a subset of patients who require further testing.⁵ Our group, and others, have previously published ultrabrief screens for the general medicine, nonsurgical population and for patients with dementia.^{5,6} The UB-2 is an ultrabrief screen consisting of "Months of the year backward" (MOYB) and "What day of the week is it?", which has a sensitivity of 93% and specificity of 64% in hospitalized older adults and takes less than 40 seconds to administer.⁵ However, no such screens for delirium have been developed for the group with relatively high cognitive and phys-

ical functioning undergoing scheduled major surgery in which delirium may present differently. Thus, the purpose of this study was to develop an ultrabrief screen for postoperative delirium using data from a large study of delirium in cognitively intact, older adults undergoing scheduled major noncardiac surgery.

METHODS

We performed a secondary data analysis on 560 patients enrolled between June 18, 2010, and August 8, 2013, in the Successful Aging After Elective Surgery (SAGES) study,⁷ an ongoing prospective cohort study of older adults undergoing major elective surgeries (eg, total hip or knee replacement; lumbar, cervical, or sacral laminectomy; lower extremity arterial bypass surgery; open abdominal aortic aneurysm repair; and open or laparoscopic colectomy). Exclusion criteria included evidence of dementia, delirium, prior hospitalization within 3 months, legal blindness, severe deafness, terminal condition, history of schizophrenia or psychosis, and history of alcohol abuse or withdrawal. The Institutional Review Boards of Beth Israel Deaconess Medical Center, Brigham and Women's Hospital, and Hebrew SeniorLife, all in Boston, Massachusetts, approved the study.

SAGES Delirium Assessment and Additional Variables

The presence or absence of delirium was based on daily in-hospital assessments by trained research staff using the Confusion Assessment Method (CAM)⁸ long form. The Delirium Symptom Interview (DSI)⁹ and information related to acute changes in mental status were also included as provided by nursing staff and/or family. Delirium severity was determined using the CAM-S.¹⁰ Participants in The SAGES Study had an ini-

*Corresponding Author: Andrea Yevchak Sillner, PhD;
Email: amy139@psu.edu; Telephone: 814-863-4294; Twitter: @amyevchak.

**Drs Marcantonio and Fick contributed equally as senior authors to this manuscript.

Published online first August 19, 2020.

Find additional supporting information in the online version of this article.

Received: December 4, 2019; Revised: March 9, 2020;

Accepted: March 10, 2020

© 2020 Society of Hospital Medicine DOI 10.12788/jhm.3410

TABLE. Baseline Characteristics of the Study Cohort (N = 560)

Characteristics	Full Sample N = 560	Delirium n = 134	No Delirium n = 426
Age - mean years (SD)	76.6 (5.2)	77.4 (5.0)	76.4 (5.2)
Male - n (%)	234 (41.8)	53 (39.6)	181 (42.5)
White - n (%)	523 (93.4)	125 (93.3)	398 (93.4)
Education - mean years (SD)	15.0 (2.9)	14.7 (3.0)	15.1 (2.9)
Married - n (%)	332 (59.3)	79 (59.0)	253 (59.4)
Lives Alone - n (%)	167 (29.8)	39 (29.1)	128 (30.0)
Charlson Comorbidity Index - n (%)			
0	257 (45.9)	54 (40.3)	203 (47.7)
1	139 (24.8)	23 (17.2)	116 (27.2)
2+	164 (29.3)	57 (42.5)	107 (25.1)
GDS 15 score - mean (SD)	2.5 (2.5)	3.0 (2.8)	2.3 (2.4)
GCP score - mean (SD)	57.6 (7.3)	54.7 (6.5)	58.5 (7.3)
3MS score - mean (SD)	93.5 (5.4)	91.6 (5.8)	94.1 (5.1)
Impaired in ADL - n (%)	42 (7.5)	12 (9.0)	30 (7.0)
Impaired in IADL - n (%)	152 (27.1)	48 (35.8)	104 (24.4)
Surgery type - n (%)			
Orthopedic			
Vascular	454 (81.1)	105 (78.4)	349 (81.9)
Gastrointestinal	35 (6.3)	11 (8.2)	24 (5.6)
	71 (12.7)	18 (13.4)	53 (12.4)

Abbreviations: ADL, activities of daily living; GCP, General Cognitive Function Composite; GDS, Geriatric Depression Scale, range = 0-15, higher is worse; IADL, instrumental activities of daily living; 3MS, Modified Mini-Mental State examination, range = 0-100, higher is better; SD, standard deviation.

tial baseline, presurgical assessment in their homes. Cognitive and physical functioning, depression, comorbidities, laboratory, and self-reported demographic data were collected.

Statistical Analyses

We included CAM delirium data from postoperative days (POD) 1 and 2 for each participant, if available; postoperative day 0 was not included because of potential residual anesthetic effects. We chose these days because most delirium began on POD1 or 2, and patients started being discharged on POD3. We considered all one-, two-, and three-item combinations of the 12 cognitive items of the 3D-CAM¹¹ because of their demonstrated high information content for CAM diagnostic features per Item Response Theory.¹² There were 12 possible one-item screens, 66 two-item screens, and 220 three-item screens. Sensitivity, specificity, and 95% confidence intervals for each were compared with CAM delirium determination. An ideal ultrabrief screen for delirium has high sensitivity with moderate specificity; general guidelines considered based on investigator consensus included screens with a sensitivity higher than 0.90 and specificity greater than 0.70. Because these screens are used to quickly rule out delirium, we also pres-

ent the percent positive screen among the entire population (whether delirium is present or not). Screens with a positive screen rate of more than 50% are unlikely to be helpful in ruling out delirium quickly in a large enough fraction of the population. We also required that in multiple item screens, no two items should assess the same CAM feature. For instance, we would eliminate a two-item screen with MOYB and four-digit span since both items measure CAM Feature 2 (Inattention). Finally, we evaluated screen performance separately on POD1 and POD2. Switching screens by POD can be confusing, so we chose a single best screen that retained excellent performance over both days. Data analyses used SAS version 9.4 (SAS Institute, Cary, North Carolina).

RESULTS

The dataset included 560 adults who had an average age of 76.6 years (SD = 5.2), were 58% women, and were highly educated (15.0 years; SD = 2.9; Table). Postoperative delirium occurred during one or more days in 134 individuals (24%). A total of 1,100 delirium assessments were used, with 113 that were CAM positive (10.3%). For POD1, we used 551 assessments, 61 of which were positive (11.1%); for POD2, 549 assessments were used,

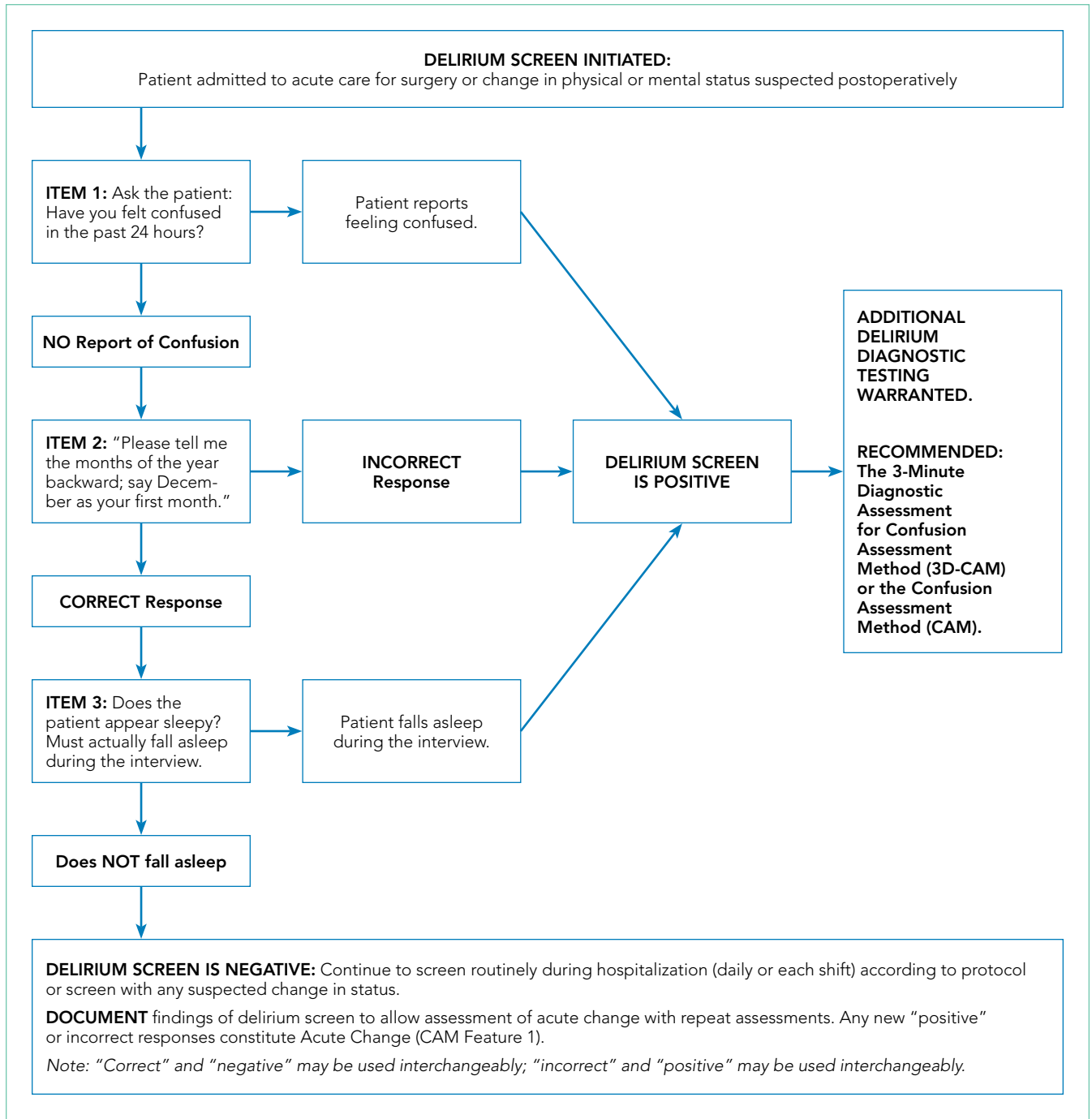


FIG. Flow Diagram of Delirium Screening Process Using Best Performing Three-Item Delirium Screen.

with 51 positive (9.3%). Appendix Tables present the positive screen rates, sensitivities, specificities, and 95% confidence intervals of all 12 one-item screens and the 12 best performing two- and three-item screens in order of decreasing sensitivity.

The best ultrabrief screen from POD1 included the following three items: “Does the patient report feeling confused?”, MOYB, and “Does the patient appear sleepy?”, with a sensitivity of 0.95 (95% CI, 0.87-0.99) and specificity of 0.73 (95% CI, 0.69-0.77). The same combination of items has a sensitivity of 0.88 (95% CI, 0.77-0.96) and a specificity

of 0.70 (95% CI, 0.66-0.74) on POD2. When POD1 and POD2 are combined, the sensitivity is 0.92 (95% CI, 0.85-0.96) and specificity is 0.72 (95% CI, 0.69-0.74). We consider this to be our best screen overall.

DISCUSSION

We identified a three-item screen for delirium after elective surgery consisting of “Does the patient report feeling confused?”, MOYB, and “Does the patient appear sleepy?” In our own prior work, we identified a two-item screen consist-

ing of MOYB and “What is the day of the week?” as the best ultrascreen for delirium in general medicine populations (termed the “UB-2”)⁵ and a subsequent screen for patients with delirium superimposed on dementia (DSD) including “What type of place is this?”, Days of the Week Backward, and “Does the patient appear sleepy?”⁶ All three contain a test of attention (a cardinal feature of delirium) and a test of orientation, although the specific test for that varies. Both the surgical and DSD screens include “Does the patient appear sleepy?”, which addresses a reduced level of consciousness. This might be particularly important in the postoperative setting because of residual effects of anesthesia and/or postoperative analgesic medications contributing to delirium. Work done by others confirms our current findings, which is that MOYB is the best single item for most groups. Belleli et al¹³ and Han et al¹⁴ included MOYB as the single attentional item in the 4AT and B-CAM, respectively. The Nu-DESC has been used as a screen in surgical patients; however, it involves only nursing observations and no direct questioning of the patient.¹⁵

The Figure describes how our “best screen” could be integrated into clinical care. One or more “positive” or incorrect responses on these three items constitutes a positive screen that should be further evaluated with the CAM or 3D-CAM. If all three items are correct or negative, this effectively rules out delirium; however, continued periodic screening on a daily (or per shift) basis is indicated. On repeat testing, if any of the previously negative or correct items becomes positive or incorrect, this would be evidence for Acute Change, CAM Feature 1. Finally, it should be noted that, if all three items in our best screen are positive, full CAM criteria for delirium diagnosis are met within the screen itself, and no further testing is required. We envision this process being facilitated by use of an app-based program that generates optimal screening items based on patient and setting characteristics.

There are several limitations that must be noted. First, our three-item screen may not generalize to nonsurgical candidates or those undergoing emergent surgery and should be tested in these groups. Second, the SAGES sample is relatively homogenous with respect to racial and ethnic diversity and was highly educated with little functional impairment and no dementia. Therefore, results may not be generalizable to populations with lower educational attainment and/or preexisting mental and physical disabilities. A third limitation is that screen items were included in the reference standard delirium assessment, leading to a potential bias toward increased sensitivity. Finally, all screens were derived from secondary data analysis and further research will be needed to prospectively validate the results. Despite these limitations, this study has several strengths including the use of a well-characterized surgical population and a rigorous approach to delirium measurement. It is one of the first studies to identify a screening tool targeted to identifying delirium in postoperative older adults.

Future research should prospectively validate our screening tool and test its implementation in a real-world clinical environment. As part of this process, clinicians should document barriers and facilitators to widespread implementation.

The goal of such screens is to facilitate early identification of postoperative delirium, which will allow timely intervention to address underlying causes and prevent adverse consequences, thereby improving the outcomes of vulnerable older surgical patients.

Disclosures: The authors have nothing to disclose.

Funding: This manuscript was supported by the following grants: R01AG030618 (Marcantonio, Fick)—Researching Efficient Approaches to Delirium Identification Study (READI); P01AG031720 (Marcantonio, Inouye)—Successful Aging after Elective Surgery Study (SAGES); National Institute on Aging grant K24AG035075 (Marcantonio); R24AG054259 (Marcantonio, Fick, Inouye)—Network for Investigation of Delirium across the US (NIDUS). Dr. Sillner reported an Early Career Award from the Gordon & Betty Moore Foundation payable to her institution.

References

- Marcantonio ER. Postoperative delirium: a 76-year-old woman with delirium following surgery. *JAMA*. 2012;308(1):73-81. <https://doi.org/10.1001/jama.2012.6857>
- Seib CD, Rochefort H, Chomsky-Higgins K, et al. Association of patient frailty with increased morbidity after common ambulatory general surgery operations. *JAMA Surg*. 2018;153(2):160-168. <https://doi.org/10.1001/jamasurg.2017.4007>
- Gleason LJ, Schmitt EM, Kosar CM, et al. Effect of delirium and other major complications after elective surgery in older adults. *JAMA Surg*. 2015;150(12):1134-1140. <https://doi.org/10.1001/jamasurg.2015.2606>
- Aldecoa C, Bettelli G, Bilotta F, et al. European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium. *Eur J Anaesthesiol*. 2017;34(4):192-214. <https://doi.org/10.1097/EJA.0000000000000594>
- Fick DM, Inouye SK, Guess J, et al. Preliminary development of an ultrascreen two-item bedside test for delirium. *J Hosp Med*. 2015;10(10):645-650. <https://doi.org/10.1002/jhm.2418>
- Steenma E, Zhou W, Ngo L, et al. Ultra-brief screeners for detecting delirium superimposed on dementia. *J Am Med Dir Assoc*. 2019;20(11):1391-1396.e1. <https://doi.org/10.1016/j.jamda.2019.05.011>
- Schmitt EM, Marcantonio ER, Alsop DC, et al. Novel risk markers and long-term outcomes of delirium: the Successful Aging after Elective Surgery (SAGES) study design and methods. *J Am Med Dir Assoc*. 2012;13(9):818.e1-818.e810. <https://doi.org/10.1016/j.jamda.2012.08.004>
- Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method. a new method for detection of delirium. *Ann Intern Med*. 1990;113(12):941-948. <https://doi.org/10.7326/0003-4819-113-12-941>
- Albert MS, Levkoff SE, Reilly C, et al. The delirium symptom interview: an interview for the detection of delirium symptoms in hospitalized patients. *J Geriatr Psychiatry Neurol*. 1992;5(1):14-21. <https://doi.org/10.1177/002383099200500103>
- Inouye SK, Kosar CM, Tommet D, et al. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Ann Intern Med*. 2014;160(8):526-533. <https://doi.org/10.7326/M13-1927>
- 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. <https://doi.org/10.7326/M14-0865>
- Yang FM, Jones RN, Inouye SK, et al. Selecting optimal screening items for delirium: an application of item response theory. *BMC Med Res Methodol*. 2013;13(1):8. <https://doi.org/10.1186/1471-2288-13-8>
- Belleli G, Morandi A, Davis DH, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing*. 2014;43(4):496-502. <https://doi.org/10.1093/ageing/afu021>
- 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. <https://doi.org/10.1016/j.annemergmed.2013.05.003>
- Gaudreau JD, Gagnon P, Harel F, Tremblay A, Roy MA. Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale. *J Pain Symptom Manage*. 2005;29(4):368-375. <https://doi.org/10.1016/j.jpainsymman.2004.07.009>