



Published in final edited form as:

Ann Intern Med. 2022 January ; 175(1): 65–73. doi:10.7326/M21-1687.

Comparative Implementation of a Brief App-directed Delirium Identification Protocol by Hospitalists, Nurses, and Nursing Assistants

Edward R. Marcantonio, MD, SM^{a,b,c,*}, Donna M. Fick, RN, PhD^{d,e,*}, Yoojin Jung, PhD^a, Sharon K. Inouye, MD, MPH^{b,c,f}, Marie Boltz, GNP-BC, PhD^d, Douglas L. Leslie, PhD^e, Erica Husser, PhD^d, Priyanka Shrestha, RN, MS^d, Amber Moore, MD, MPH^{a,c,g}, Kimberlyann Sulmonte, DNP, MHA, RN^h, Jonathan Siuta, MDⁱ, Malaz Boustani, MD MPH^j, Long H. Ngo, PhD^{a,c}

^aDivision of General Medicine, Beth Israel Deaconess Medical Center, Boston MA

^bDivision of Gerontology, Department of Medicine, Beth Israel Deaconess Medical Center, Boston MA

^cHarvard Medical School, Boston, MA

^dCollege of Nursing, Penn State University, University Park, Pennsylvania

^eCollege of Medicine, Pennsylvania State University, Hershey, Pennsylvania

^fAging Brain Center, Marcus Institute for Aging Research, Hebrew SeniorLife, Boston MA

^gDivision of General Internal Medicine, Massachusetts General Hospital Boston MA

^hDepartment of Nursing, Beth Israel Deaconess Medical Center, Boston, MA

ⁱDepartment of Medicine, Mount Nittany Medical Center, State College, PA

^jCenter for Health Innovation and Implementation Science, Indiana University School of Medicine.

Abstract

Background: Systematic screening improves delirium identification among hospitalized older adults. Little data exists on how to implement such screening.

Corresponding Author: Edward R. Marcantonio, MD SM, Mailing address: Beth Israel Deaconess Medical Center CO-216, 330 Brookline Ave, Boston MA 02215, emarcant@bidmc.harvard.edu, Twitter handle: @EdMarcantonio, @Agilis.

*Drs. Marcantonio and Fick contributed equally to this paper.

Author Contributions:

Study Concept and Design: Edward Marcantonio, Donna Fick, Sharon Inouye

Acquisition of Data: Donna Fick, Marie Boltz, Erica Husser, Priyanka Shrestha, Amber Moore, Kimberlyann Sulmonte, Jonathan Siuta, Edward Marcantonio

Analysis and Interpretation of Data: Long Ngo, Yoojin Jung, Edward Marcantonio

Drafting of the Manuscript: Edward Marcantonio

Critical Revision of the Manuscript for Important Intellectual Content: Yoojin Jung, Long Ngo, Donna Fick, Sharon Inouye, Douglas Leslie, Marie Boltz, Erica Husser, Priyanka Shrestha, Amber Moore, Kimberlyann Sulmonte, Jonathan Siuta, Malaz Boustani

Conflict of Interest: The authors have no relevant conflicts of interest to report.

Reproducible Research Statement:

Study protocol: Available from Dr. Marcantonio (emarcant@bidmc.harvard.edu)

Statistical code: Available from Dr. Ngo (Ingo@bidmc.harvard.edu)

Data set: Not available

Objective: To test implementation of a brief app-directed delirium identification protocol by physicians, nurses, and certified nursing assistants (CNAs) in real-world practice relative to a research reference standard delirium assessment (RSDA).

Design: Prospective diagnostic test study

Setting: Large urban academic medical center, small rural community hospital

Participants: 527 general medicine inpatients (mean age 80 years, 35% with pre-existing dementia); 399 clinicians (53 physicians, 236 nurses, 110 CNAs)

Measurements: On two study days, enrolled patients underwent an RSDA. Subsequently, CNAs performed an ultra-brief delirium screen (UB-2), while physicians and nurses performed a two-step protocol consisting of the UB-2, followed in “positives” by the 3D-CAM diagnostic assessment.

Results: Delirium was diagnosed in 154/924 RSDAs (17%) and in 114/527 patients (22%). Completion rate for clinician protocols exceeded 97%. The UB-2 was administered in 62 (51) [mean (SD)] seconds by CNAs; two-step protocols were administered in 104 (99) seconds by nurses, and 106 (105) seconds by physicians. The UB-2 had sensitivities (95% CI) of 88 (72–96)%, 87 (73–95)%, and 82 (65–91)% when administered by CNAs, nurses, and physicians, respectively, with specificities of 64–70%. The two-step protocol had overall accuracy of 89 (83–93)% and 87 (81–91)%, with sensitivities of 65 (48–79)% and 63 (46–77)%, and specificities of 93 (88–96)% and 91 (86–95)% for nurses and physicians, respectively. Two-step sensitivity for moderate-severe delirium was 78 (54–91)%.

Limitations: Two sites, limited diversity

Conclusion: An app-directed delirium identification protocol was feasible, brief, and accurate, with CNAs and nurses performing as well as physicians.

Primary Funding Source: National Institute on Aging

Keywords

Delirium; screening; case identification; implementation science

INTRODUCTION:

Delirium (acute confusion) affects 25–30% of older general medical inpatients, and 10–15% of medical-surgical hospitalized patients of any age (1–3). Acutely, delirium is associated with longer lengths of stay, increased post-hospital nursing home use, and higher mortality (1, 4). Longer term, delirium is associated with poor functional recovery, cognitive decline, and incident dementia (5, 6). Since only 12–35% of delirium is identified in routine care (1–3), many hospitals are launching systematic delirium screening; however, evidence supporting such implementation is limited (7). Therefore, we conducted the READI: Researching Efficient Approaches to Delirium Identification study (8).

READI builds on our previous work deriving and validating the 3D-CAM, a 3-Minute Diagnostic Assessment for the Confusion Assessment Method (CAM) (9–11), and the UB-2

ultra-brief screen (12), consisting of the two most sensitive items in the 3D-CAM. In READI, we combine the UB-2 and 3D-CAM into a 2-step delirium identification protocol (8) (Figure 1). The UB-2 quickly rules out delirium in patients who answer both questions correctly. Otherwise, the protocol continues with the rest of the full 3D-CAM, or with a 3D-CAM skip pattern such that the first positive item to trigger a CAM feature ends assessment of that feature. The UB-2, followed by 3D-CAM with skip pattern, is called the ultra-brief CAM (UB-CAM) (13). Paper versions of the 3D-CAM, UB-2, and UB-CAM, along with training materials, are available at help.agscocare.org.

While numerous delirium screening tools have been developed (14), and some have been integrated into electronic health systems, few studies have tested their use by clinicians in real-world practice compared to a Reference Standard Delirium Assessment (RSDA). The specific aims of READI are to prospectively implement and compare the feasibility, speed, and accuracy of delirium identification by three clinical disciplines—hospitalists, nurses, and certified nursing assistants (CNAs). We hypothesized that clinician delirium assessments would: 1) be feasible, with greater than 95% completion; 2) be brief, on average under 1 minute for the UB-2, and under 2 minutes for the 2-step protocols; 3) have accuracy over 85%, sensitivity over 80% and specificity over 90% compared to delirium determination by an RSDA.

METHODS:

The READI app:

We developed the READI app using the REDCap application programming interface (API) for use on the iPad Air® tablet (15). The app executes the adaptive testing of the 2-step protocols described above, with or without the skip pattern. The app allows the user to record accuracy of patient responses and automatically times each question. On subsequent days, the app retrieves information from the previous day's assessment to assess acute change. Once completed, the app reports the presence or absence of delirium, and uploads data into RedCAP, a secure research database.

Patient Participants:

We enrolled patients from a large teaching hospital in Massachusetts and a smaller community hospital in Pennsylvania. Inclusion criteria were: 1) age ≥ 70 years old, 2) admitted to general medicine (hospitalist) service, 3) able to communicate effectively in English, 4) without terminal or major psychiatric/substance abuse conditions, 5) expected hospital stay at least 2 days from enrollment, and 6) not a previous study participant. We overenrolled patients with dementia since they are at higher risk for delirium non-recognition (16). Research staff members screened admission logs and approached eligible patients for informed consent. If the patient lacked capacity, the designated surrogate decision-maker was contacted. The study protocol and informed consent were approved by the Institutional Review Boards of participating institutions.

Clinician Participants and Training:

Clinicians (hospitalists, nurses, CNAs) assigned to general medicine units provided streamlined informed consent followed by a brief survey of demographics, education, primary language, certification in geriatrics/gerontology, and years at the hospital. Subsequently, physicians and nurses were given brief training on the 2-step protocol, while CNAs were trained to administer only the UB-2, which is considered a “mental status vital sign”, and within the scope of CNA practice (17). Total training time was 15 minutes for CNAs, and 20–30 minutes for physicians and nurses. A READI cue card was provided as a refresher tool.

Reference Standard Delirium Assessments (RSDA):

The RSDA diagnosis was established daily using detailed assessment by trained research associates. This included: 1) cognitive assessment using the Mini-Mental State Examination, purchased from Psychological Assessment Resources (18); 2) supplemental attention testing, using digit span up to 5 forwards and 4 backwards, and days of the week backwards (19); 3) a modified 8 question Delirium Symptom Interview to elicit symptoms of delirium (20); 4) the Confusion Assessment Method (CAM) long-form (10), which rates 10 features, and enables delirium diagnosis using the CAM algorithm, and quantification of delirium severity using the CAM-S (21). A 10% subset of RSDAs were conducted simultaneously by two staff members to assess inter-rater reliability. Enrolled patients were assessed on two consecutive hospital days unless discharged, in which case they contributed only 1 day of assessments.

On the first hospital day, enrolled patients were administered a brief questionnaire to assess race, ethnicity, marital status, living situation, years of education, pre-hospital function using basic and instrumental Activities of Daily Living (22, 23). Also on day 1, the medical record was reviewed to quantify comorbidities using the Charlson index (24), and to determine pre-hospital diagnosis of dementia. Additionally, researchers conducted interviews with a patient-designated proxy (usually close family member) to confirm patient-reported demographic and functional data or to obtain these if the patient was unable to answer the questions. Proxies were administered the AD-8, a proxy-based screening questionnaire for dementia (25). Similar to our previous work (26), we considered dementia to be present if: 1) the patient and/or proxy reported a previous clinical diagnosis of dementia or Alzheimer’s Disease; 2) a diagnosis of dementia or Alzheimer’s Disease was in the medical record, or 3) the AD-8 had a score of 4 or higher out of 8. Mini-Mental State scores were not used to determine dementia status, as these might be confounded by delirium, or subtler reversible cognitive impairment (27).

Clinician Delirium Assessment Protocols:

Following the RSDA, enrolled hospitalists, nurses, and CNAs were contacted by research staff to assess patient participants using the READI app. Each clinician was kept strictly blinded to the RSDA, and to the other clinicians’ assessments. Whenever possible, primary hospital team clinicians were used. Clinicians completed the delirium screens within the context of their daily work on the hospital units. Because of work schedules, the same clinician did not necessarily assess the patient on both study days. We aimed to complete the

RSDA and all 3 clinician assessments within a 2-hour window between 11 AM and 3 PM on two consecutive hospital days.

Statistical Analyses:

For descriptive statistics, we present mean and standard deviation for continuous variables, and percentages for categorical variables. When the distribution of the continuous variable is asymmetric, we also report the median, interquartile range, and minimum and maximum value. For Table 1, the unit of observation is the patient. For Tables 2, 3, and 4, the unit of observation is the delirium assessment.

For the UB-2 and 2-step delirium identification protocols, we calculated modeled point estimates and 95% confidence intervals for each test characteristic, with each clinical discipline considered separately. We addressed clustering within our study design as recommended in the literature (28) using SAS Proc GLIMMIX with logit link, and included random effects for patient clustering, and four fixed effects for clinician characteristics: age, gender, hospital and years practicing at the hospital. For further details and justification of this approach, see the Supplement (Technical Appendix). Estimates of accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive (LRP), and likelihood ratio negative (LRN) came from the mean of the fitted probabilities derived from these models. The 95% confidence intervals were computed from taking the mean of the upper and lower bounds of the standard errors of the fitted probabilities.

We performed stratified analyses for patients with and without dementia (based on the RSDA) for delirium identification protocol duration and test characteristics. We performed sensitivity (“per protocol”) analyses examining protocol accuracy for clinician assessments performed within 2 hours of the RSDA, for those performed by the patient’s primary hospital team, for Day 1 vs. Day 2 assessments, and for clinicians who performed 10 or more assessments. To address order effects, we calculated accuracy for the 1st, 2nd, and 3rd assessment of the day independent of clinician type. We calculated sensitivity for the nurse and physician 2-step protocols including only RSDA delirium positive assessments where measured delirium severity was moderate-high (CAM-S = 8). Finally, we present subset analyses of Day 2 assessments to determine how often the Day 1 “look back” feature was used, how often it was positive for acute change, and the percent of positive clinician delirium protocols attributed to this feature. When test characteristics are reported for stratified and sensitivity analyses, modeling methods accounting for clustering as described above were used.

We used SAS software version 9.4 64-bit for Microsoft Windows 10 (SAS Institute, Inc., Cary, NC) for all analyses.

RESULTS:

Enrollment, Patient and Clinician Participant Characteristics, Feasibility:

Figure 2 presents study enrollment at the two study sites. We approached 1,070 eligible patients; 527 (49%) were enrolled and had at least one RSDA. Characteristics of the

participants are described in Table 1, overall, and by delirium status (Supplement Table S1 presents the same data by site). The mean age was 80 years old, 43% were male, and 85% were white, non-Hispanic, which were similar to unenrolled eligible patients (data not shown). Approximately 1/3 were college graduates. Comorbidity and functional impairment were high—over 60% had Charlson scores ≥ 2 , over 1/3 had basic ADL impairments and nearly 2/3 had Instrumental ADL impairments. Dementia was present in 183 (35%). Compared to those without delirium, the 114 (21.6%) patients with delirium by RSDA were slightly older, had less education, higher comorbidity, and higher functional impairments. Most notably, those with delirium were more than twice as likely to have dementia.

The 399 enrolled clinicians—53 physicians, 236 nurses, and 110 CNAs constituted over half of the day-shift staff on the participating units (Supplement Table S2) and were representative of staff at the study hospitals (data not shown). The mean age of all three groups were in the 30's. Men accounted for slightly over half of physicians, but only 10% of nurses and CNAs. CNAs were the most racially and linguistically diverse, with less than half being white, and only 2/3 being native English speakers. All physicians had doctoral degrees; most nurses had college degrees, and most CNAs high school degrees. Two-thirds of each group were in practice less than 5 years, and three-quarters worked at the current hospital for less than five years. Under 5% of each discipline reported certification in geriatrics or gerontology.

Figure 2 shows the completion rate of clinician assessments on each day when a RSDA was performed. Combining both sites, 527 RSDAs were completed on study day 1; subsequently 512 physicians (97%), 518 nurses (99%) and 510 CNAs (97%) completed app-based protocols. On study day 2, 397 RSDAs were performed; subsequently 381 physicians (96%), 384 nurses (97%) and 388 CNAs (98%) completed app-based protocols. Overall, 2693 clinician delirium screens were completed, a 97.1% completion rate. The main reason for non-completion was unavailability of the clinician due to excessive competing workload.

Reference Standard Delirium Assessments (RSDA):

CAM-defined delirium was present in 154/924 RSDAs (17%), including 101/527 (19%) on study day 1 and 53/397 (13%) on day 2. Among 89 paired RSDAs, 88 (99%) independently agreed on presence or absence of delirium, concordance kappa of 0.97 [95% C.I. 0.93–1.0]. As seen in Table 2, cognitive scores were lower, and delirium symptoms much higher in those with delirium vs. without. Patients with delirium had a mean CAM-S score of 8.1 vs. 2.3 for those without delirium, similar to our previous work (19), and representing mild to moderate delirium (range 0–19, 19 worst). Only 29/154 (19%) of RSDAs with delirium had hyperactive psychomotor features. Additionally, 103/154 RSDAs (67%) represented delirium superimposed on dementia. Supplement Table S3 presents additional RSDA data about CAM features present in those with and without delirium.

Clinician Assessments—Time Required:

Table 3 presents time required for clinicians to complete the app-guided protocols (Supplement Table S4 presents the same data with outliers >3 SD from the mean trimmed). The UB-2 was completed in 55–62 seconds by all three types of clinicians. Two-step

protocols took a mean of 104–106 seconds for nurses and physicians to complete, with the skip pattern reducing the mean duration to 83 seconds.

Clinical Assessments—Test Characteristics:

Table 4 presents the test characteristics of the UB-2 and the two-step delirium identification protocols, accounting for clustering in the study design (Supplement Table S5 presents the same data without clustering, Table S6 the raw numbers). The UB-2 had sensitivity (95% CI) of 88 (82–96)% by CNAs, 87 (73–95)% by nurses, and 82 (65–91)% by physicians. The corresponding specificities were 64 (55–72)% for CNAs, 70 (62–76)% for nurses, and 66 (58–74)% for physicians. Overall accuracy was 69 (61–75)% for CNAs, 73 (67–79)% for nurses, and 70 (62–76)% for physicians. Notably, the UB-2 is designed to maximize sensitivity rather than accuracy or specificity.

For the two-step protocols, overall accuracy was 89 (83–93)% for nurses and 87 (81–91)% for physicians. Sensitivities (95% CI) were 65 (48–79)% for nurses and 63 (46–77)% for physicians, with specificities of 93 (88–96)% for nurses, and 91 (86–95)% for physicians. Stratified analyses show that the skip pattern did not appreciably change sensitivity of nurse assessments, sensitivity (95% CI) of 68 (44–85)% for no skip vs. 64 (42–82)% for skip, but did lower sensitivity of physician assessments from 71 (48–86)% for no skip to 56 (35–76)% for skip. Specificities remained in the 90–94% range, and overall accuracy in the 87–89% range regardless of skip pattern for both disciplines.

When we restricted delirium cases to those with an RSDA CAM-S of 8 (moderate to severe), sensitivity of the two-step protocol increased to 78 (54–91)% for nurses and 78 (57–91)% for physicians.

Stratified and Sensitivity Analyses:

The presence of dementia increased mean (SD) UB-2 administration times to slightly over one minute: 71(50), 65(39), 69(56) seconds for CNAs, nurses, and physicians, respectively and increased mean 2-step protocol administration time to about 2.5 minutes: 150(114) and 153(122) seconds for nurses and physicians, respectively. The presence of dementia decreased overall accuracy of both the UB-2 (accuracy %, 95% C.I.) to 58 (45–70)%, 63 (50–73)%, 59 (46–70)% for CNAs, nurses, and physicians, respectively, and the 2-step protocols to 80 (69–88)% and 78 (66–86)% for nurses and physicians, respectively. In general, clinician protocol sensitivity was slightly higher in patients with dementia, while specificity was lower, though it remained above 80% for both nurse and physician 2-step protocols. For details, see Supplement Tables S7, S8.

None of the following sensitivity analyses appreciably changed clinician protocol accuracy (see Supplement Tables S9, S10 for details): 1) restricting to protocols performed within 2 hours of the RSDA (12% CNA, 11% nurse, and 23% hospitalist protocols dropped); 2) restricting to protocols performed by the patient's primary hospital team; 3) Day 1 vs. Day 2 protocols, 4) order of protocol administration on each day, and 5) including only clinicians who performed 10 or more protocols.

Finally, we examined the impact of the automatic look-back feature of the app to detect acute change on Day 2 by comparing responses to Day 1. Of 389 Day 2 nurse assessments, the acute change feature was used 29 times (7.5%) and was positive 15 times (3.9%). Of 381 physician Day 2 assessments, the feature was used 30 times (7.9%) and was positive 21 times (5.5%). The app-directed acute change detection accounted for 15/59 (25.4%) and 21/68 (30.9%) of cases of delirium diagnosed by nurses and physicians, respectively, on Day 2.

DISCUSSION:

Our study demonstrated that an iPad app-directed adaptive delirium identification protocol conducted by typical clinicians as part of routine daily workflow was feasible, quick and accurate. Clinicians completed over 97% of protocols, with an average completion time of under two minutes. Relative to the RSDA, sensitivity of the UB-2 ranged from 82–88%, with CNAs and nurses performing similarly to physicians. Sensitivity of the two step protocols was 63–65%, with specificity exceeding 90%, and overall accuracy in the high 80%’s. Sensitivity for moderate-severe delirium was 78%. Our approach directly addresses two major barriers to screening, lack of knowledge and lack of time, by providing a structured assessment with specified cognitive testing and an app-facilitated adaptive design to efficiently complete the CAM, the most widely used and recommended bedside algorithm to assess delirium (11). Since only 12–35% of delirium cases are currently detected in clinical practice, our protocol can substantially improve delirium identification with a low false positive rate among vulnerable hospitalized adults.

A major strength of our study was simultaneous comparison of protocol implementation by three clinical disciplines—physicians, nurses, and CNAs—incorporated into their daily workflow, relative to an RSDA. Prior studies have focused on delirium screening by nurses and physicians; fewer studies have employed CNAs, especially in inpatient hospital care (29–34). Contrary to our expectations, protocol accuracy was similar across all clinical disciplines. In the two-step protocol, nurses had equal or slightly higher point estimates of sensitivity and specificity than physicians. For the UB-2 screen, CNAs had slightly higher sensitivity than nurses or physicians, though lower specificity. Confidence intervals overlapped for all test characteristics across all three disciplines. This suggests that delirium screening can be performed efficiently by nurses and CNAs operating within their scope of practice, using for example the nurse two-step protocol with skip (UB-CAM), or the CNA UB-2 followed by nurse 3D-CAM with skip (not explicitly tested in our study). Notably, physician sensitivity was lower in skip pattern protocols, possibly because they were more reliant on time spent with the patient during the assessment to inform coding of observational items.

Another strength is our CAM-based, app-directed delirium identification protocol that uses adaptive testing to improve efficiency (15). Numerous tools exist for delirium case identification including both CAM and non-CAM-based methods (14, 35–40). Our protocol provides a standardized approach to score the CAM, and the app compares previous day’s results to assess “acute change”, a key feature of delirium. Subset analyses demonstrated that this “look back” feature accounted for 25–30% of positive nurse or physician delirium

assessments on Day 2. Our app was designed as a research-based tool that sent data to RedCAP. For clinical implementation, apps could input data into the electronic health record. Importantly, to improve patient outcomes, delirium identification tools such as our app should be coupled with recommendations for delirium prevention and/or management (1).

The sensitivity of our app-directed protocol was lower than previously reported 3D-CAM or simulated UB-CAM validation results (9, 13). These prior studies, typical of initial validation studies of delirium instruments, used highly trained research staff (9,11,14, 35–38). READI results come from nearly 400 diverse front-line clinicians without prior delirium assessment training during the course of their daily work. In this context, two-step sensitivities in the mid-60's and approaching 80% for moderate-severe delirium, with specificities above 90%, compare favorably to those reported in the literature (41–43). In a simultaneous qualitative analysis, clinicians reported positive experiences with the app-directed protocols, and expressed interest in using them in practice (15, 44).

Strengths of our study were the large sample of over 500 patients, nearly 400 clinicians, and over 900 days of delirium assessment. Another strength was use of a rigorous RSDA, which involved direct patient assessment, medical record review and proxy interview to characterize patients' pre-hospital status. Clinician protocols were performed strictly blinded to the RSDA. We purposely enrolled a challenge sample, with mean age of 80 years and high prevalence of dementia. Moreover, over 2/3 of delirium occurred in patients with dementia, and over 80% was hypoactive; these are the two strongest risk factors for delirium non-recognition (42).

We also had some study limitations. Given our study design, an individual clinician did relatively few assessments. Clinician training was quite limited in scope and duration; more training with skill validation might have improved performance. RSDAs were performed by trained research associates and not physicians; however, this approach has been utilized in our prior studies (45). Clinicians administered the protocol once daily, more frequent administration may have improved sensitivity. Several factors may impact generalizability of our findings, including enrolling primarily older, white, non-Hispanic medical patients, and less than 50% of patients approached. Future studies should evaluate protocol implementation in diverse locations and patients, including surgical patients, skilled nursing facilities, and resource-poor settings.

In conclusion, we found that an iPad app-directed, CAM-based adaptive delirium identification protocol was feasible, brief, and accurate, with CNAs and nurses performing as well as physicians. Future studies should evaluate costs of systematic screening across hospital systems, and identify implementation strategies that promote adoption and fidelity of delirium identification, coupled with prevention or management strategies, among vulnerable hospitalized older adults.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding Sources:

NIA Grants: R01AG030618 (Marcantonio, Fick), R24AG054259 (Inouye), and K24AG035075 (Marcantonio)

We would like to thank the older adult participants and the staff members at the study hospitals, without whom the study would not have been possible. We would also like to thank the following study staff members: Jacqueline Gallagher, Shannon Malloy, Angelee Butters, Brett Armstrong, Abigail Overstreet, Michele Henry, Jess Garrity, Meghan McGraw, Logan Foreman, Janelle Bessette, Caroline Madrigal.

Sponsor's Role:

The National Institute on Aging had no role in the design, conduct, analysis, or decision to submit the manuscript for publication.

REFERENCES

1. Marcantonio ER. Delirium in hospitalized older adults. *N Engl J Med* 2017;377:1456–1466. [PubMed: 29020579]
2. Ryan DJ, O'Regan NA, Caoimh RÓ, Clare J, O'Connor M, Leonard M, McFarland J, Tighe S, O'Sullivan K, Trzepacz PT, Meagher D, Timmons S. Delirium in an adult acute hospital population: predictors, prevalence and detection. *BMJ Open*. 2013;3:e001772. doi: 10.1136/bmjopen-2012-001772.
3. Schubert M, Schürch R, Boettger S, Garcia Nuñez D, Schwarz U, Bettex D, Jenewein J, Bogdanovic J, Staehli ML, Spirig R, Rudiger A. A hospital-wide evaluation of delirium prevalence and outcomes in acute care patients - a cohort study. *BMC Health Serv Res*. 2018;18:550. [PubMed: 30005646]
4. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in older persons: advances in diagnosis and treatment. *JAMA*. 2017;318:1161–1174. [PubMed: 28973626]
5. Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA*. 2010;304:443–51. [PubMed: 20664045]
6. Goldberg TE, Chen C, Wang Y, Jung E, Swanson A, Ing C, Garcia PS, Whittington RA, Moitra V. Association of delirium with long-term cognitive decline: a meta-analysis. *JAMA Neurol*. 2020;77:1–9.
7. Wilson JE, Mart MF, Cunningham C, Shehabi Y, Girard TD, MacLulich AMJ, Slooter AJC, Ely EW. Delirium. *Nat Rev Dis Primers*. 2020 Nov 12;6:90. doi: 10.1038/s41572-020-00223-4. [PubMed: 33184265]
8. Fick DM, Inouye SK, McDermott C, et al. Pilot study of a two-step delirium detection protocol administered by certified nursing assistants, physicians, and registered nurses. *J Gerontol Nurs* 2018;44(5):18–24. [PubMed: 29596707]
9. Marcantonio ER, Ngo LH, O'Connor M, et al. 3D-CAM: Derivation and validation of a 3-Minute diagnostic interview for CAM-defined delirium. *Ann Int Med* 2014;161:554–561. [PubMed: 25329203]
10. Inouye SK, Van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: The confusion assessment method: A new method for detection of delirium. *Ann Int Med* 1990;113:941–948. [PubMed: 2240918]
11. Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. *JAMA*. 2010;304:779–86. [PubMed: 20716741]
12. Fick DM, Inouye SK, Guess J, et al. Preliminary development of an ultrabrief two-item bedside test for delirium. *J Hosp Med* 2015;10:645–650. [PubMed: 26369992]
13. Motyl CM, Ngo L, Zhou W, Jung Y, Leslie D, Boltz M, Husser E, Inouye SK, Fick D, Marcantonio ER. Comparative accuracy and efficiency of four delirium screening protocols. *J Am Geriatr Soc*. 2020;68:2572–2578. [PubMed: 32930409]
14. Helfand BKI, D'Aquila ML, Tabloski P, Erickson K, Yue J, Fong TG, Hshieh TT, Metzger ED, Schmitt EM, Boudreaux ED, Inouye SK, Jones RN. Detecting delirium: a systematic review of

identification instruments for non-ICU settings. *J Am Geriatr Soc.* 2021; 69: 547–555. [PubMed: 33135780]

15. Armstrong B, Habtemariam D, Husser E, Leslie DL, Boltz B, Jung Y, Fick DM, Inouye SK, Marcantonio ER, Ngo LH. A mobile app for delirium screening. *JAMIA Open*, 4(2), 2021, 1–9. doi: 10.1093/jamiaopen/ooab027.
16. Fick DM, Agostini JV, Inouye SK. Delirium superimposed on dementia: a systematic review. *J Am Geriatr Soc.* 2002;50:1723–32. [PubMed: 12366629]
17. McMullen TL, Resnick B, Hansen JC, Miller N, Rubinstein R. Certified nurse aides and scope of practice: clinical outcomes and patient safety. *J Gerontol Nurs.* 2015;41:32–9.
18. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975; 12:189–98. [PubMed: 1202204]
19. Yang FM, Jones RN, Inouye SK, Tommet D, Crane PK, Rudolph JL, et al. Selecting optimal screening items for delirium: an application of item response theory. *BMC Med Res Methodol.* 2013;13:8. [PubMed: 23339752]
20. Albert MS, Levkoff SE, Reilly C, Liptzin B, Pilgrim D, Cleary PD, Evans D, Rowe JW. The delirium symptom interview: an interview for the detection of delirium symptoms in hospitalized patients. *J Geriatr Psychiatry Neurol.* 1992;5:14–21. [PubMed: 1571069]
21. Inouye SK, Kosar CM, Tommet D, Schmitt EM, Puelle MR, Saczynski JS, Marcantonio ER, Jones RN. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Ann Intern Med.* 2014;160:526–533. [PubMed: 24733193]
22. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA.* 1963;185:914–9. [PubMed: 14044222]
23. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9:179–86. [PubMed: 5349366]
24. Charlson ME, Pompei P, Ales KL, MaKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40:373–83. [PubMed: 3558716]
25. Galvin JE, Roe CM, Powlishta KK, Coats MA, Muich SJ, Grant E, et al. The AD8: a brief informant interview to detect dementia. *Neurology.* 2005;65:559–64. [PubMed: 16116116]
26. Steensma E, Zhou W, Ngo L, Gallagher J, Inouye S, Leslie D, Boltz M, Kolanowski A, Mion L, Marcantonio ER, Fick D. Ultra-brief screeners for detecting delirium superimposed on dementia. *J Am Med Dir Assoc.* 2019;20:1391–1396. [PubMed: 31279670]
27. Inouye SK, Zhang Y, Han L, Leo-Summers L, Jones R, Marcantonio E. Recoverable cognitive dysfunction at hospital admission in older persons during acute illness. *J Gen Intern Med* 2006;21:1276–81. [PubMed: 16965558]
28. Kahan BC, Morris TP. Assessing potential sources of clustering in individually randomised trials. *BMC Medical Research Methodology.* 2013;13:1–9. [PubMed: 23297754]
29. Solberg LM, Campbell CS, Jones K, Vaughn I, Suryadevara U, Fernandez C, Shorr R. Training hospital inpatient nursing to know (THINK) delirium: A nursing educational program. *Geriatr Nurs.* 2020;42:16–20. [PubMed: 33197702]
30. Boockvar KS, Judon KM, Eimicke JP, Teresi JA, Inouye SK. Hospital Elder Life Program in long-term care (HELP-LTC): a cluster randomized controlled trial. *J Am Geriatr Soc.* 2020;68:2329–2335. [PubMed: 32710658]
31. Hasemann W, Tolson D, Godwin J, Spirig R, Frei IA, Kressig RW. Nurses’ Recognition of hospitalized older patients with delirium and cognitive impairment using the delirium observation screening scale: a prospective comparison study. *J Gerontol Nurs.* 2018;44:35–43. [PubMed: 30484846]
32. Hullick C, Conway J, Higgins I, Hewitt J, Stewart B, Dilworth S, Attia J. An assistant workforce to improve screening rates and quality of care for older patients in the emergency department: findings of a pre-post, mixed methods study. *BMC Geriatr.* 2018;18:126. doi: 10.1186/s12877-018-0811-6. [PubMed: 29843623]
33. Teale E, Young J, Siddiqi N, Munyombwe T, Harrison J, Schuurmanns M. Study protocol-investigation of the Delirium Observation Screening Scale (DOSS) for the routine detection of

- delirium in the care home setting: a prospective cohort study. *BMJ Open*. 2016;6:e009615. doi: 10.1136/bmjopen-2015-009615.
34. Steis MR, Behrens L, Colancecco EM, Mogle J, Mulhall PM, Hill NL, Fick DM, Kolankowski AM. Licensed nurse and nursing assistant recognition of delirium in nursing home residents with dementia. *Ann Long Term Care*. 2015;23:15–20.
 35. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: Validation of the Confusion Assessment Method for the intensive care unit (CAM-ICU). *Crit Care Med* 2001;29:1370–1379. [PubMed: 11445689]
 36. 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;62:457–465. [PubMed: 23916018]
 37. Schuurmans MJ, Shortridge-Baggett LM, Duursma SA. The delirium observation screening scale: a screening instrument for delirium. *Res Theory Nurs Pract*. 2003 Spring;17:31–50. [PubMed: 12751884]
 38. Bellelli G, Morandi A, Davis DHJ, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing* 2014;43:496–502. [PubMed: 24590568]
 39. Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S. The Memorial delirium assessment scale. *J Pain Symptom Manage*. 1997;13:128–37. [PubMed: 9114631]
 40. Trzepacz PT, Mittal D, Torres R, Canary K, Norton J, Jimerson N. 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 Spring;13:229–42. [PubMed: 11449030]
 41. Marcantonio ER, Bergmann MA, Kiely DK, Orav EJ, Jones RN. Randomized trial of a delirium abatement program for post-acute skilled nursing facilities. *J Am Geriatr Soc* 2010; 58: 1019–26. [PubMed: 20487083]
 42. 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;161:2467–73. [PubMed: 11700159]
 43. Lemiengre J, Nelis T, Joosten E, Braes T, Foreman M, Gastmans C, Milisen K. Detection of delirium by bedside nurses using the confusion assessment method. *J Am Geriatr Soc*. 2006;54:685–9. [PubMed: 16686883]
 44. Husser EK, Fick DM, Boltz M, Shrestha P, Siuta J, Malloy S, Overstreet A, Leslie DL, Ngo L, Jung Y, Inouye SK, Marcantonio ER. Implementing a Rapid, Two-Step Delirium Screening Protocol in Acute Care: Barriers and Facilitators. *J Am Geriatr Soc*. 2021; 69: 1349–1356. [PubMed: 33474729]
 45. Hshieh TT, Fong TG, Schmitt EM, Marcantonio ER, Aquila ML, Gallagher J, Xu G, Guo YR, Abrantes TF, Bertrand SE, Jones RN, Inouye SK; for the BASIL Study Group. The Better Assessment of Illness Study for Delirium Severity: Study Design, Procedures, and Cohort Description. *Gerontology*. 2019;65:20–29. [PubMed: 30032141]

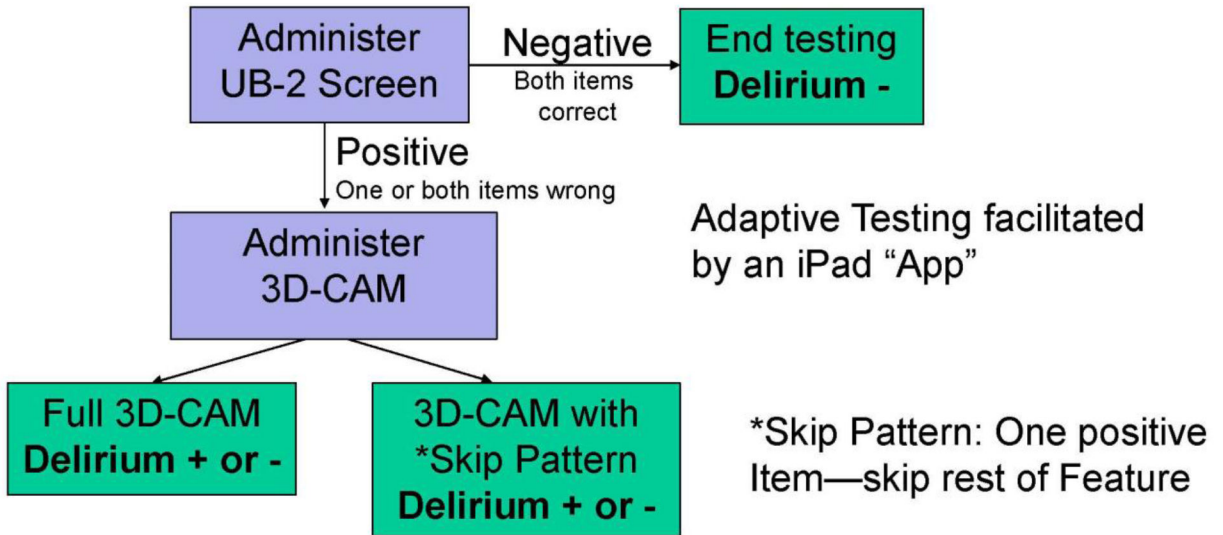


Figure 1. Two-Step Delirium Identification Protocol

This figure depicts the flow of the two-step app-directed, adaptive delirium assessment protocol. The protocol begins with the UB-2 screen, which consists of two items, “What is the day of the week?” and “Please recite the months of the year backwards”. If the screen is negative (both items are correct), the protocol ends and delirium is not present. If one or both of the UB-2 items is incorrect, the protocol proceeds to the second step, the 3D-CAM. Additionally, the 3D-CAM is randomized to using or not using a skip pattern. In the no-skip 3D-CAM, all of the remaining items are administered (the UB-2 items are part of the 3D-CAM). In the skip 3D-CAM, designed to further shorten the protocol, once a single item is incorrect or positive, assessment for that item’s CAM feature ends (the threshold for triggering the presence of a CAM feature on the 3D-CAM is a single positive item). The final determination of delirium presence or absence is based on the Confusion Assessment Method (CAM) diagnostic algorithm, which is built into the 3D-CAM assessment. The adaptive, interactive testing is facilitated by use of an iPad App.

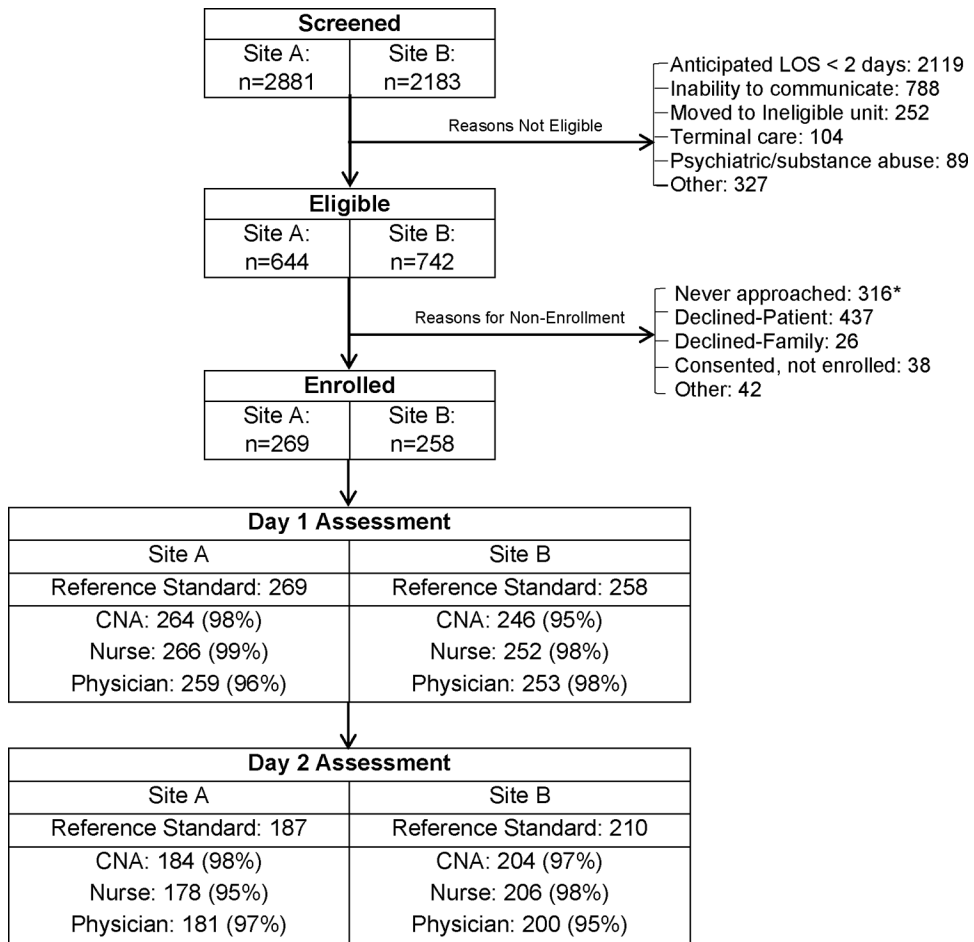


Figure 2. Study flow diagram

The numbers in the parentheses are the percent of completed delirium screening protocols by certified nursing assistants (CNAs), nurses, and physicians following the reference standard delirium assessment (RSDA)

Footnote: *Reasons for non-approach: Reference standard not available = 229, Patient not available (off unit) = 61, Droplet precautions = 15, Other = 11

Table 1.

Patient participant characteristics

	Overall (n = 527)	Without Delirium (n = 413)	With Delirium (n = 114)
Age – mean years (SD)	79.7 (6.6)	79.5 (6.6)	80.5 (6.6)
Male – n (%)	226 (43)	169 (41)	57 (50)
Race* – n (%)			
White	458 (88)	356 (86)	102 (91)
Black or African American	43 (8)	36 (9)	7 (6)
Others	22 (4)	19 (5)	3 (3)
Hispanic or Latino* – n (%)	11 (2)	8 (2)	3 (3)
Education* – n (%)			
Less than high school	54 (10)	37 (9)	17 (15)
High school graduate	191 (37)	143 (35)	48 (44)
Some college	101 (19)	87 (21)	14 (13)
College graduate	73 (14)	62 (15)	11 (10)
Master’s degree	70 (14)	52 (13)	18 (16)
Doctoral degree	33 (6)	31 (7)	2 (2)
Married* – n (%)	224 (43)	173 (42)	51 (46)
Lives Alone* – n (%)	194 (37)	166 (40)	28 (25)
Charlson Score – n (%)			
0	36 (7)	29 (7)	7 (6)
1	172 (33)	143 (35)	29 (25)
2+	319 (60)	241 (58)	78 (69)
ADL* – sum score (SD)	0.9 (1.4)	0.7 (1.3)	1.4 (1.8)
Impaired in ADL* – n (%)	202 (38)	140 (34)	62 (54)
IADL* – sum score (SD)	1.8 (1.8)	1.5 (1.6)	2.7 (2.1)
Impaired in IADL* – n (%)	333 (63)	244 (59)	89 (78)
Dementia – n (%)	183 (35)	110 (27)	73 (64)

Table Abbreviations and Footnotes:

SD=standard deviation, Charlson= Charlson comorbidity score, ADL=basic Activities of Daily Living, scored 0–6, 6 worst, IADL=Instrumental Activities of Daily Living, 0–7, 7 worst

* Number of missing data: race (4), ethnicity (2), education (5), marital status (2), living situation (2), ADL (2), IADL (2).

Table 2.

Reference standard delirium assessment cognitive and delirium data

		Overall (n = 924)	Without Delirium (n = 770)	With Delirium (n = 154)
MMSE Total Score* 0–30, 30 best	mean (SD)	24.9 (4.9)	26.1 (3.6)	18.8 (5.7)
	median (Q1, Q3)	26 (23, 29)	27 (25, 29)	20 (15, 23)
	min, max	3, 30	12, 30	3, 30
MMSE Orientation* 0–10, 10 best	mean (SD)	8.3 (2.1)	8.8 (1.6)	5.7 (2.6)
	median (Q1, Q3)	9 (8, 10)	9 (8, 10)	6 (4, 8)
	min, max	0, 10	2, 10	0, 10
MMSE Registration/Recall* 0–6, 6 best	mean (SD)	4.9 (1.3)	5.1 (1.0)	3.7 (1.7)
	median (Q1, Q3)	5 (4, 6)	5 (5, 6)	4 (3, 5)
	min, max	0, 6	1, 6	0, 6
Digit Span* # Correct out of 5	mean (SD)	4.0 (1.0)	4.1 (0.9)	3.2 (1.0)
	median (Q1, Q3)	4 (3, 5)	4 (4, 5)	3 (3, 4)
	min, max	0, 5	0, 5	0, 5
Days of Week Backwards* n correct (%)		780 (85)	696 (91)	84 (56)
DSI* 0–8, 8 worst	mean (SD)	0.6 (1.3)	0.3 (0.8)	2.3 (2.1)
	median (Q1, Q3)	0 (0, 1)	0 (0, 0)	2 (0, 4)
	min, max	0, 8	0, 6	0, 8
CAM-S Score 0–19, 19 worst	mean (SD)	3.3 (2.9)	2.3 (1.8)	8.1 (2.0)
	median (Q1, Q3)	2 (1, 5)	2 (1, 3)	8 (7, 9)
	min, max	0, 14	0, 8	4, 14
Psychomotor Summary	Hyperactive, n(%)	24 (3)	9 (1)	15 (10)
	Hypoactive, n(%)	86 (8)	31 (4)	55 (36)
	Mixed, n(%)	14 (2)	0 (0)	14 (9)
	Normal, n(%)	800 (87)	730 (95)	70 (45)

Table Abbreviations and Footnotes:

MMSE=Mini-Mental State Examination, DSI=Delirium Symptom Interview, CAM-S=CAM-Severity Score, SD=Standard deviation, Q1=quartile 1, Q3=quartile 3, Min=minimum, Max=Maximum

[†]Digit Span= Number Correct out of 3, 4, and 5 forwards, plus 3, 4 backwards

* Number of missing data: MMSE total score (3), MMSE orientation (1), MMSE registration/recall (1), digit span (8), days of week backwards (10), DSI (4).

Table 3.

Duration of clinician delirium identification protocols

	UB-2		2-Step		2-Step, No Skip		2-Step, Skip (UB-CAM)		
	CNA (n = 862)	Nurse (n = 873)	Physician (n = 856)	Nurse (n = 866)	Physician (n = 854)	Nurse (n = 423)	Physician (n = 419)	Nurse (n = 443)	Physician (n = 435)
Mean (SD), sec	62 (51)	55 (34)	55 (44)	104 (99)	106 (105)	125 (124)	131 (124)	83 (61)	83 (76)
Median (Q1,Q3), sec	49 (35, 74)	46 (32, 70)	44 (31, 66)	60 (34, 143)	62 (33, 155)	58 (32, 212)	66 (32, 204)	62 (36, 112)	58 (33, 104)
Min. sec	6	6	6	6	6	6	6	9	6
Max. sec	862	326	775	674	936	674	936	393	801

Table Abbreviations and Footnotes:

UB-2=Ultra-brief 2-Item Screen, 2-Step=2-Step Delirium Identification Protocol, UB-CAM=Ultra-brief CAM, CNA=Certified Nursing Assistant, SD=Standard Deviation, sec=Seconds, Q1=quartile 1, Q3=quartile 3, Min=Minimum, Max=Maximum

Table 4.

Test characteristics of clinician delirium identification protocols*

	UB-2		2-Step		2-Step, No Skip		2-Step, Skip (UB-CAM)		
	CNA (n = 898)	Nurse (n = 910)	Physician (n = 895)	Nurse (n = 902)	Physician (n = 893)	Nurse (n = 441)	Physician (n = 440)	Nurse (n = 461)	Physician (n = 453)
Accuracy, %	68.6 (60.8–75.4)	73.4 (66.5–79.2)	69.7 (62.3–76.2)	88.9 (83.2–92.7)	86.8 (81.0–90.9)	89.0 (81.0–93.7)	86.6 (77.9–92.1)	89.0 (81.1–93.7)	87.2 (79.0–92.2)
Sensitivity, %	88.2 (71.6–95.5)	87.1 (73.1–94.6)	81.9 (64.7–91.1)	64.8 (47.5–79.0)	62.7 (46.1–76.9)	68.1 (43.8–85.2)	70.7 (47.9–86.2)	64.4 (41.6–82.3)	56.3 (35.4–76.1)
Specificity, %	63.7 (55.0–71.7)	69.6 (61.8–76.4)	66.3 (57.9–73.7)	92.9 (87.6–96.0)	91.2 (85.5–94.7)	92.7 (83.2–96.7)	90.0 (83.3–94.1)	93.9 (86.7–97.2)	93.1 (85.0–96.8)
PPV, %	32.0 (22.4–43.4)	35.4 (25.2–47.2)	32.5 (22.5–44.5)	65.0 (46.6–79.3)	58.5 (41.4–73.8)	63.0 (39.2–80.7)	59.3 (37.1–78.1)	68.8 (43.9–85.8)	60.8 (36.5–80.7)
NPV, %	96.3 (90.3–98.6)	96.6 (91.8–98.6)	94.7 (87.7–97.5)	93.4 (87.9–96.4)	92.3 (86.7–95.6)	94.2 (86.7–97.5)	93.7 (84.8–97.5)	92.8 (84.8–96.7)	91.7 (83.5–95.8)
LRP	2.5 (1.6–3.5)	3.1 (2.0–4.6)	2.3 (1.5–3.3)	8.4 (3.5–19.0)	8.1 (3.6–17.0)	8.8 (2.4–36.7)	7.5 (3.2–15.1)	8.8 (2.7–26.3)	12.2 (3.1–44.6)
LRN	0.2 (0.1–0.5)	0.2 (0.1–0.4)	0.3 (0.1–0.6)	0.4 (0.2–0.6)	0.4 (0.2–0.6)	0.3 (0.2–0.7)	0.3 (0.1–0.6)	0.4 (0.2–0.7)	0.5 (0.2–0.7)

Table Abbreviations and Footnotes:

UB-2=Ultra-brief 2-Item Screen, 2-Step=2-Step Delirium Identification Protocol, UB-CAM=Ultra-brief CAM, PPV=Positive Predictive Value, NPV=Negative Predictive Value, LRP=Likelihood Ratio Positive, LRN=Likelihood Ratio Negative

* The values in the parentheses are 95% confidence intervals. All test characteristics and 95% confidence interval estimates are calculated accounting for clustering by patients and clinician characteristics. See text and technical appendix for details.