Table B.61: Delirium, Screening and Assessment–Single Studies

Note: Full references are available in the [Section 14.1 reference list](#Section14point1refs).

| Author, Year | Description of Patient Safety Practice | Study Design; Sample Size; Patient Population | Setting | Outcomes: Benefits | Outcomes: Harms | Implementation Themes/Findings | Risk of Bias (High, Moderate, Low) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Adamis et al., 20105 | Evaluation of evidence-based assessment tools | Literature review; sample size range 47–432; older adults | Acute care | The Confusion Assessment Method (CAM), Delirium Rating Scale (DRS), DRS-Revised-98 (DRS-R-98), Memorial Delirium Assessment Scale (MDAS), and Neelon and Champagne (NEECHAM) confusion scale are sufficiently validated. | Not provided | Not provided | Low |
| Adamis et al., 201510 | Comparison of four different tools to identify delirium | Prospective observational study; 200 patients; adults aged 70+ | University teaching general hospital | Agreement between Diagnostic and Statistical Manual-5 (DSM-5), DSM-IV, DRS-R-98, and CAM were all significant. Highest agreement was between DSM and DRS-R-98, while lowest agreement was between DSM-IV and DSM-5. | Not provided | Not provided | Low |
| Adamis et al., 201611 | Comparison of clock drawing test as screening tool (with DRS) | Prospective, observational, longitudinal study; 200 patients; adults aged 70+ | Acute medical wards of general hospital | There was a significant negative correlation between the Clock Drawing Test (CDT) and DRS-R-98 (Pearson correlation r=-0.62, p<0.0010), CDT and CAM (Spearman’s rho=-0.40, p<0.001), CDT and Montreal Cognitive Assessment (MoCA) (Pearson’s r=0.69, p<0.001), and CDT and MoCA (Pearson’s r=0.77, p<0.001). | Not provided | Not provided | Moderate |
| Arendts et al., 20174 | Use of Emergency Department (ED) Delirium Screening Form | Prospective three-phase trial; 3,905 patients; adults age 65+ admitted to an inpatient hospital bed from the ED | EDs of two tertiary hospitals | An absolute increase in delirium diagnosis of 2% across study phases was statistically insignificant (Pearson chi-square=2.49, P=0.29). | Not provided | Not provided | Not provided |
| Boettger et al., 201716 | Comparison of CAM and Intensive Care Delirium Screening Checklist (ICDSC) for delirium in intensive care unit (ICU) patients | Prospective, descriptive cohort study; 210 patients; adults under intensive care management for more than 18 hours | Twelve-bed ICU at level one trauma center | Agreement was moderate between the CAM-ICU and DSM-IV-TR (k=0.44, p<0.001), the ICDSC and DSM-IV-TR (k=0.60, p<0.001), and the CAM-ICU and ICDSC (k=0.56, p<0.001). | Not provided | Not provided | Low |
| Bull et al., 201722 | Evaluating telephone- based screening for delirium to be used by family members | Pre-post, quasi-experimental design; 34 family caregiver-older adult dyads; older adults aged 70+ who underwent joint surgery | Orthopedic clinic at a Veterans Affairs Medical Center | There was 94% agreement (32 out of 34) between the Family Confusion Assessment Method (FAM-CAM) and the researcher-led CAM 2 days after the patient’s surgery. Cohen kappa for agreement was moderate (k=0.477; p=0.001). Two family caregivers reported positive FAM-CAM ratings during the 2 weeks after hospitalization, which led to the physician changing the prescribed pain medication. | Not provided | Not provided | Moderate |
| De et al., 201524 | Screening tools for culturally and linguistically different populations | Systematic review; hospitalized adult inpatients | Hospital, excluding ICU | CAM, DRS, Nursing Delirium Screening Scale (NuDESC), sleep quality rating, MDAS, 4 A’s Test (4 AT) | Not provided | Not provided | Moderate |
| Van Eijk et al., 200917 | Comparison of screening tools (CAM-ICU vs. ICDSC) | One hundred twenty-six patients (mean age = 62.4 years) | Thirty-two-bed mixed medical and surgical ICU | The CAM-ICU showed superior sensitivity and negative predictive value (64% and 83%) compared with the ICDSC (43% and 75%). The ICDSC showed higher specificity and positive predictive value (95% and 82% vs. 88% and 72%). The sensitivity of the physician’s view was only 29%. | Not provided | Not provided | Low |
| Flanagan and Spencer, 201632 | Use of CAM in post-acute patients—informal caregivers | Community-dwelling older adults aged 65+ admitted to postacute care (rehabilitation or skilled nursing center) with the intention of returning to community living and their family member/informal caregivers. The participants had to be English-speaking and have a caregiver willing to participate in the study. | Post-acute care | The FAM-CAM highly correlates with the confusion assessment method and diagnostic and statistical manual of mental disorders text revision criteria for detecting delirium in older adults in the postacute care setting. | Not provided | This study was a convenience sample; subjects were not randomized. The sample size was small, which limits generalization of the findings. A replication of this study with a larger sample size, as well as additional sites, would be beneficial. | Moderate |
| Frisch et al, 201331 | Tools for assessing patients in transport by emergency medical services staff; compared CAM to Glasgow Coma Scale (GCS) | A convenience sample of matched dyads of emergency medical services providers and elderly patients (age ≥65 years) | Two academic, tertiary-care EDs | Prehospital providers’ recognition of any delirium symptom resulted in a sensitivity of 0.63 (95% confidence interval [CI] 0.43–0.79) and a specificity of 0.74 (95% CI 0.73–0.84). Prehospital report of a GCS *<*15 has a sensitivity of 0.67 (95% CI 0.47–0.82) and a specificity of 0.85 (95% CI 0.80–0.89). | Not provided | Not provided | Moderate |
| Gelinas et al., 201813 | Evaluation of nursing assessment tools for delirium in ICU | Systematic review;  two independent reviewers analyzed the psychometric properties of five delirium assessment tools by using a standardized scoring system (range, 0–20) to assess the development process, reliability, validity, feasibility, and implementation of each tool | Intensive care | Psychometric properties were very good for the CAM-ICU (19.6) and the ICDSC (19.2), moderate for the NuDSS (13.6), low for the Delirium Detection Score (DDS) (11.2), and very low for the Cognitive Test for Delirium (8.2). | Not provided | Not provided | Low |
| Khan et al., 20127 | Evaluation of Richmond Agitation-Sedation Scale (RASS) and Riker Sedation-Agitation Scale (SAS) in identifying patients eligible for delirium assessment | Quality improvement project; 975 patients; patients aged 18 and older admitted to the ICU | Four hundred fifty-seven-bed university-affiliated urban public hospital | The Spearman rank correlation between the RASS and SAS scores was estimated at 0.91; 70.1% of screens were eligible for CAM-ICU assessment using RASS ≥-3 compared with 72.1% using SAS ≥3. The agreement between RASS and SAS for assessing CAM-ICU eligibility as estimated by the k coefficient was 0.93. | Not provided | Not provided | Not provided |
| Kuczmarska et al.,  20169 | Evaluated CAM-ICU and 3D-CAM for hospitalized general medical/surgical patients | Hospitalized general medicine patients aged ≥75 years | Two non-intensive care general medicine units at a single academic medical center | The sensitivity (95% CI) of delirium detection for the 3D-CAM was 95% (74%, 100%) and for the CAM-ICU was 53% (29%, 76%), while specificity was >90% for both instruments. Subgroup analyses showed that the CAM-ICU had sensitivity of 30% in patients with mild delirium vs. 100% for the 3D-CAM. | Not provided | Not provided | Moderate |
| Marcantonio et al., 20148 | Derivation and validation of 3D-CAM | Prospective validation study; 201 patients; adults aged 75+ admitted to general medicine or geriatric medicine services | Large urban teaching hospital | Compared with the reference standard delirium diagnosis, the 3D-CAM had a sensitivity of 95% (CI 90 to 97%), resulting in a positive likelihood ratio of 16.8 (95% CI 8.9 to 31.8) and a negative likelihood ratio of 0.05 (CI 0.01 to 0.20). In post-hoc analyses, sensitivity of the 3D-CAM improved to 96% and specificity to 98%. | Not provided | Not provided | Not provided |
| Mistarz at al., 201127 | Demonstrated importance of using a structured assessment tool rather than relying on nursing documentation | Bedside nurses assessed 35 patients for delirium during routine patient care throughout their shift; this assessment was then compared to an independent assessment using the CAM-ICU performed by a nurse trained in this delirium detection tool | A 12-bed general ICU | Not provided | There was a significant discrepancy between the ICU bedside nurses’ assessment of delirium and the independent formal delirium assessment using the CAM-ICU. Routine bedside nursing patient interactions do not reliably detect delirium in a critically ill patient. | Not provided | High |
| Moon et al., 201828 | Building delirium assessment tool into electronic health records; used CAM tool | Participants: a total of 3,284 patients for the development of Auto-DelRAS, 325 for external validation, 694 for validation after clinical applications | Medical and surgical ICUs in two university hospitals in Seoul, Korea. | The predictive validity, analyzed after the clinical application of Auto-DelRAS after 1 year, showed a sensitivity of 0.88, specificity of 0.72, positive predictive value of 0.53, negative predictive value of 0.94, and a Youden index of 0.59. A relatively high level of predictive validity was maintained with the Auto-DelRAS system, even 1 year after it was applied to clinical practice. | Not provided | Not provided | Moderate |
| Neufeld et al., 201118 | CAM-ICU and ICDSC in non-critically ill hospitalized patients | Not provided | Two medical oncology units at a large teaching hospital | Not provided | This study suggests that in non-critically ill hospitalized patients, the CAM-ICU and ICDSC intensive care delirium screening tools are not adequately sensitive for use in routine clinical practice. | Not provided | Low |
| Neufeld et al., 201314 | Comparison of CAM-ICU with NuDESC | Prospective study; 91 patients; adults aged 70+ receiving general anesthesia during surgery | One teaching hospital | CAM-ICU had sensitivity of 28% (95% CI 16 to 45) and specificity of 98% (95% CI 88 to 100). NuDESC (threshold ≥2) had similarly high specificity of 92% (95% CI 80 to 97) and low sensitivity of 32% (95% CI 19 to 48). The NuDESC (threshold ≥1) had improved sensitivity (80%; 95% CI 65 to 91) but reduced specificity (69%; 95% CI 54 to 80). | Not provided | Not provided | Not provided |
| O’Regan et al., 201425 | Spatial Span Forwards (SSF) and months of the year backwards (MOTYB) as bedside screening tests to detect delirium | Cross-sectional study; 265 patients; adult inpatients excluding patients in the ED, ICU, and hematology/burns isolation unit | Large tertiary referral hospital | MOTYB was most accurate of the three, with a sensitivity of 83.3% (95% CI 69.8 to 92.5) and specificity of 90.8% (95% CI 86.1 to 94.3). SSF5 had high sensitivity (91.7%, 95% CI 80 to 97.6) but low specificity (69.12%, 95% CI 62.5 to 75.2). SSF4 had the lowest sensitivity (77.1%, 95% CI 62.7 to 87.9) | Not provided | Not provided | Not provided |
| Radtke et al., 200815 | Use of CAM, NuDESC, and DDS | Observational study; 154 patients; adults aged 18+ admitted to recovery room after general anesthesia | Recovery room of hospital | The CAM had a sensitivity of 0.43 and specificity of 0.985; the DDS had sensitivity of 0.14 and specificity of 0.99; the Nu-DESC had sensitivity of 0.95 and specificity of 0.87. Sensitivity between the CAM and DDS did not differ significantly (p=0.07). The NuDESC was most sensitive compared to the DDS (p<0.001) and CAM (p=0.003). Specificity did not differ significantly between scores. | False positives were 1.5% for CAM, 12.8% for the Nu-DESC, and 0.8% for the DDS. False negative rates were 57% for the CAM, 85% for the DDS, and 5% for the Nu-DESC. | Not provided | Not provided |
| Rainsford et al., 201412 | Compare CAM, DRS-R-98, and chart review | Fifty-one patients; adults aged 18+ with a diagnosis of advanced cancer | Nineteen-bed acute inpatient specialist palliative care unit | The DRS-R-98 identified 21 patients positively for delirium (41.2%) and 30 negatively for delirium (58.8%). The CAM identified 21 patients positively for delirium (41.2%) and 36 negatively for delirium (70.6%). The clinical team identified only 15 patients positively for delirium (29.4%) and 30 negatively (58.8%).  The data are unclear about agreement between the CAM and DRS-R-98. | Not provided | Not provided | Not provided |
| Rice et al., 201126 | CAM performance in practice (nurse vs. researcher rating) | Prospective, descriptive design; 170 patients; adults aged 65+ at risk for delirium | Tertiary care teaching hospital (541 beds) | Sensitivity of nurses’ rating of delirium using the CAM was low for all comparisons with researcher ratings (25% overall, 25% best case, 10% worst case). A significant difference was observed between nurses’ recognition of delirium and that of the researcher, X2 (1, n=170)= 40.21, p<0.001; Fisher exact p<0.001. Specificity was high (99.6% overall, 100% best case, 100% worst case). Agreement beyond chance in detecting delirium was poor for overall (k=0.34), best case (k=0.38) and worst case (k=0.14) comparisons. | Not provided | Not provided | Not provided |
| Ringdal et al., 201119 | Compare CAM with DSM-IV; evaluate Mini-Mental State Exam (MMSE) as screening tool | Mokken nonparametric latent trait model for unidimensional scaling; 365 patients; adults aged 65+ acutely admitted for hip fracture for at least 24 hours | Two hospitals in Oslo, Norway | The MMSE cutpoint of 24 had 84% agreement with the CAM for patients diagnosed with delirium. Using the total MMSE score had a sensitivity of 46% and specificity of 96%. Using step-wise logistic regression to locate a subset of MMSE items that may function as a screening tool resulted in a sensitivity of 51% and specificity of 95%. | Using the MMSE cutpoint of 24 had low agreement with the CAM for identifying negative cases (54% agreement), indicating a very high rate of false positives. | Not provided | Not provided |
| Rippon et al., 201620 | Development and evaluation of Delirium Early Monitoring System (DEMS) (two versions) | Observational study; 501 and 474 participants; healthcare assistants and support workers | Acute ward for patients with moderate to severe dementia in North East of England | Seventy-nine percent of staff completed the DEMS-CAM and 68% completed the DEMS-DOSS (Delirium Observation Screening Scale). Completion rates relating to the number of occasions that completion of the DEMS-CAM/DEMS-DOSS led to appropriate clinical action was 46% of the time for DEMS-CAM and 54% of the time for DEMS-DOSS. | Not provided | An end of study questionnaire completed by 10 of the non-medically trained staff found the DEMS-CAM was easier to understand than the DEMS-DOSS. | Not provided |
| Ryan et al., 200930 | CAM in palliative care | One hundred six patients; patients admitted to specialist palliative care unit study | Thirty-bed specialist palliative care unit in Mid-West region of Ireland | The sensitivity of the CAM in the pilot phase was 0.5 (0.22 to 0.78) and specificity was 1.0 (0.81 to 1.0). In the main study, the sensitivity of the CAM was 0.88 (0.62 to 0.98) and the specificity was 1.0 (0.88 to 1.0). | In the pilot phase, the non-consultant hospital doctors (NCHDs) made six false negative diagnoses of delirium. In the main study, the NCHDs made two false negative diagnoses of delirium. | A significant difference in the sensitivity of the CAM in the pilot phase and the main study was found (Χ2=5.15, p<0.05), demonstrating that the performance of the CAM was improved when the NHCDs received the “enhanced” training module. | Not provided |
| Shulman et al., 201621 | Sour Seven questionnaire as screening tool for delirium | Pilot study; 80 patients; adults aged 65+ admitted to either the medical or surgical units of the study hospital and in the hospital for at least 1 day | Large academically affiliated community hospital in Canada. | Agreement between geriatric psychiatrist on Sour Seven questionnaire and untrained nurses ranged from 64.3 to 92.8%, between geriatric psychiatrist and caregivers ranged from 44 to 84%. For each of the seven questions, the Fisher exact test analysis had a p value greater than 0.05, suggesting there was no difference between the questionnaire posed to nurses versus informal caregivers. Out of a possible maximum total score of 18 on the Sour Seven Questionnaire, a score of 4 was selected as the screening cut-off and a score of 9 was selected as diagnostic of delirium because of its specificity of 100% and high Youden Index. | Not provided | Not provided | Not provided |
| Steis et al., 201223 | Convergent validation of FAM-CAM and CAM by family caregivers | Exploratory analysis of agreement between two primary studies: the eCare for Eldercare pilot study and the Hospital to Home: Cognitively Impaired Elders/Caregivers study; 52 paired assessments from patient-caregiver dyads; adults aged 65+ with preexisting cognitive impairment. | Communities across Pennsylvania | Overall agreement between the CAM and FAM-CAM was 96%. Compared with the original CAM algorithm, the FAM-CAM had a sensitivity of 88% (95% CI=47 to 99) and specificity of 98% (95% CI=86 to 100). | Not provided | Not provided | Not provided |
| Vasilevskis et al., 201129 | Evaluate performance of CAM-ICU (nurse vs. researcher) | Prospective cohort study; 510 patients; critically ill patients admitted to the ICU | Nine hundred-bed teaching hospital | Substantial agreement between bedside and research nurses on measures done within 2 hours of each other (CAM-ICU weighted kappa=0.67, 95% CI=0.66 to 0.70; RASS weighted kappa=0.66, 95% CI=0.64 to 0.68). Of 3,856 paired assessments for delirium within 2 hours, bedside nurses identified delirium with a sensitivity of 0.81 (95% CI=0.78 to 0.83) and specificity of 0.81 (95% CI=0.78 to 0.85) compared with research nurse reference standard. | Agreement between research and bedside nurses was slightly lower for mechanically ventilated patients and in nurses assessing delirium in patients aged 65+ compared to in assessments in patients younger than 65. | Not provided | Not provided |