UCDAVIS

The Department of Emergency Medicine (EM) at the University of California Davis, Health is conducting a faculty search for EM physicians in either a clinician/educator or clinician/researcher track. Candidates must be residency trained in EM (or Pediatric EM training) with board certification/eligibility and be qualified for licensure in California. At least one year of post-residency clinical experience and/or fellowship training is preferred. Candidates are expected to enter at the Assistant/Associate level, commensurate with experience and credentials. EM faculty members at UC Davis who have preference for night shifts work fewer clinical shifts each month.

The University of California, Davis, Medical Center, one of the nation's "Top 50 Hospitals," is a 613 bed academic medical center with approximately 80,000 emergency department visits annually, including approximately 20,000 pediatric visits. The emergency services pavilion opened in 2010 and is state-of-the-art. Our program provides comprehensive emergency services to a large local urban and referral population as a level 1 trauma and burn center, paramedic base station and training center. The department also serves as the primary teaching site for a fully accredited EM residency program and nine different EM fellowship programs. Our residency training program began more than twenty years ago and currently has 54 residents. The Department has a separate area for the care of children and is one of the leading centers in the Pediatric Emergency Care Applied Research Network (PECARN).

Salary and benefits are competitive and commensurate with training and experience. Sacramento is located near the northern end of California's Central Valley, close to Lake Tahoe, San Francisco, and the "wine country" of the Napa and Sonoma Valleys. Sports enthusiasts will find Sacramento's climate and opportunities ideal.

Interested candidates should submit a letter outlining interests and experience, and curriculum vitae to: recruit.ucdavis.edu/apply/JPF02202

Erik Laurin MD, Professor and Search Committee Chair (eglaurin@ucdavis.edu)
UC Davis Department of Emergency Medicine
2315 Stockton Blvd., PSSB 2100
Sacramento, CA 95817

The University of California is an affirmative action/equal opportunity employer.

Article type : Original Contribution

Validation of the Brief Confusion Assessment Method for screening delirium in elderly medical patients in a German emergency department

Verena Baten[1], Hans-Jörg Busch, MD[2], Caroline Busche, MD[3], Bonaventura Schmid, MD[4], Miriam Heupel-Reuter, MD[5], Evgeniy Perlov, MD[6], Jochen Brich, MD[7]*, Stefan Klöppel, MD[8]*

[1] Department of Neurology and Neuroscience, Medical Center, University of Freiburg, Freiburg, Germany. Department of Emergency Medicine, University Hospital of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Email: verena.baten@uniklinik-freiburg.de
[2] Department of Emergency Medicine, University Hospital of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Email: hans-joerg.busch@uniklinik-freiburg.de
[3] Department of Emergency Medicine, University Hospital of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Email: caroline.busche@uniklinik-freiburg.de
[4] Department of Emergency Medicine, University Hospital of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Email: bonaventura-schmid@uniklinik-freiburg.de
[5] Center of Geriatrics and Gerontology Freiburg, Medical Center, University of Freiburg, Freiburg, Germany. Email: miriam.heupel-reuter@uniklinik-freiburg.de
[6] Department of Psychiatry and Psychotherapy, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Clinic for Psychiatry Luzern, St. Urban, Switzerland. Email: evgeniy.perlov@lups.ch
[7] Department of Neurology and Neuroscience, Medical Center, University of Freiburg, Freiburg, Germany. Email: jochen.brich@uniklinik-freiburg.de
[8] Department of Psychiatry and Psychotherapy, Faculty of Medicine, University of Freiburg, Freiburg, Germany. Center for Geriatric Medicine and Gerontology, Faculty of Medicine, University of Freiburg, Freiburg, Germany. University Hospital of Old Age Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland. Email: stefan.kloeppel@upd.ch

* These authors contributed equally to this work.

Corresponding author
Verena Baten
University Medical Center Freiburg, Freiburg i.Br., Germany
Email: verena.baten@uniklinik-freiburg.de

Author contributions:
study concept and design: VB, H-J B, JB, SK
acquisition of data: VB, CB, BS, M H-R, EP, JB, SK
analysis and interpretation of data: VB, JB, SK
drafting of the manuscript: VB, JB
revision of the manuscript: H-J B, SK

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/acem.13449

Disclosure: All authors report no conflict of interest.

Running Title: Validating the Brief Confusion Assessment Method

Key Words: Emergency Department, Geriatric medicine, bCAM, Delirium, Germany, Validity

Acknowledgments: We thank all physicians of the emergency department of the university hospital of Freiburg who performed the bCAM tests. Also we would like to thank Werner Vach, PhD, and Susanne Weber, MSc, from the Institute of Medical Biometry and Statistics, Medical Faculty and Medical Center, University of Freiburg, for their advice on the analysis and interpretation of data.

Abstract

BACKGROUND:

Delirium is frequent in elderly patients presenting in the emergency department (ED). Despite the severe prognosis, the majority of delirium cases remain undetected by emergency physicians (EPs). At the time of our study there was no valid delirium screening tool available for EDs in German-speaking regions. We aimed to evaluate the brief Confusion Assessment Method (bCAM) for a German ED during the daily work routine.

METHODS:

We implemented the bCAM into practice in a German interdisciplinary high volume ED, and evaluated the bCAM's validity in a convenience sample of medical patients aged ≥70 years. The bCAM, which assesses four core features of delirium, was performed by EPs during their daily work routine and compared to a gold standard based on the criteria for delirium as described in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

RESULTS:

Compared to the gold standard, delirium was found to be present in 46 (16.0%) of the 288 non-surgical patients enrolled. The bCAM showed 93.8% specificity (95% CI 90.0 – 96.5) and 65.2% sensitivity (95% CI 49.8 – 78.7). Positive and negative likelihood ratios were 10.5 and 0.37, respectively, while the odds ratio was 28.4. Delirium was missed in 10 / 16 cases, since the bCAM did not indicate altered levels of consciousness and disorganized thinking. The level of agreement with the gold standard increased for patients with low cognitive performance.

CONCLUSION:

This was the first study evaluating the bCAM for a German ED and when performed by EPs during routine work. The bCAM showed good specificity, but only moderate sensitivity. Nevertheless, application of the bCAM most likely improves the delirium detection rate in German EDs. However, it should only be applied by trained physicians in order to maximize diagnostic accuracy, and hence improve the bCAM's sensitivity. Future studies should refine the bCAM.

Introduction

Delirium is a frequent problem in emergency departments (EDs) worldwide, but is often missed by physicians and nurses, potentially leading to inadequate medical care and adverse outcomes.[1–3] Delirium is clinically defined as an acutely-developing and fluctuating syndrome with impaired attention and awareness (defined as orientation to the environment), as well as additional changes to the cognitive state.[4] It is especially prevalent in older patients, since they generally have lower cognitive reserve and are more vulnerable to risk

factors;[5] this is also particularly true if there is pre-existing neurocognitive impairment such as dementia.[1,6,7]

As both the gateway to and gatekeeper of hospital care, the ED plays a crucial role in the diagnosis of delirious patients. Eight – 17% of older patients in American and German EDs present with symptoms of delirium.[7–9] Rapid on-site diagnosis and subsequent identification of the underlying health problem is critical to each patient's ongoing medical care. Many causes of delirium are reversible if they are recognized early, and can even be resolved in the ED without the need for hospital admission.[10] Prompt diagnosis and intervention in the ED is particularly important if delirium is triggered by a life-threatening condition such as alcohol withdrawal, hypoglycemia, intoxication, intracerebral hemorrhage, or meningitis.[5,11]

Despite there being evidence for the importance of already recognizing patients with delirium in the ED and the presence of various validated delirium screening tools for the ED setting,[12] the implementation of these screenings is still lacking in many EDs. As a result, delirium is only detected in 11 - 46% of ED cases.[13] In particular, hypoactive delirium, which is the most frequent subtype in older patients, often goes undetected.[7,14] Up to 26% of elderly delirious patients are even sent home from the ED,[15,16] while in the patients who are admitted to hospital, delirium remains unrecognized in 90% of those in whom it was missed in the ED.[7]

Delirium as defined by the Diagnostic and Statistical Manual of Mental Disorder, Fifth Edition (DSM-5) is defined clinically and various screening tools have been developed to operationalize the diagnostic criteria.[17] Given that the ED differs strongly from other hospital departments in terms of its high patient turnover, fast workflow, and short patient stays,

adjustments have to be made to the common delirium screening tools in order to carry out proper screening in this setting. In particular, information about the onset and fluctuation of symptoms is often unavailable in the ED.

The brief Confusion Assessment Method (bCAM), developed by Han et al.,[18] is an adaption of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU),[19] both of which are based on the Confusion Assessment Method (CAM).[20] In contrast to the CAM, which takes at least five minutes to be completed,[20] the bCAM can be performed in less than one minute.[18] Like the CAM-ICU, the bCAM consists of objective assessments with determinate cutoffs[18] that still allow its application even if the physician barely knows the patient. The bCAM has previously been validated in 406 patients recruited in an ED in the US, where it showed a sensitivity level of 84.0%, with 95.8% specificity.[18] However, the assessment was performed in English and the evidence for its validity in German speaking EDs is therefore unclear.[21] Moreover, the enrollment was limited to one patient per day, and the bCAM was only partially conducted by physicians, with research personnel also performing part of the bCAMs.[18]

At the time of our study there was no valid delirium screening tool available for use in German EDs. The aim of this study was therefore to investigate the diagnostic strength of the German-language version of the bCAM in an interdisciplinary ED. The bCAM was applied by physicians under routine work conditions and compared to a psychiatrist's or neurologist's assessment, which served as the gold standard.

Methods

Study Design

We used a prospective observational study design that followed STARD standards [22] to evaluate implementation of the bCAM in the ED. Based on this new screening tool for the detection of delirium being implemented into our ED, the local Ethics Committee granted approval of the study without the requirement for written or verbal patient consent. However, all patients were informed about the procedure, and patients who refused to participate were subsequently not enrolled. The evaluation team informed the treating physician about the result of the gold standard diagnostic test only after all other components of the study had been completed; this therefore avoided any bias towards the physician's test result but also guaranteed the best possible treatment for the patient.

Study Setting and Population

The study was carried out in the ED of a large university teaching hospital, visited by approximately 50 000 patients each year. Screening for delirium was limited to the medical examination section of the ED, since surgical patients usually don't remain there for long enough to permit bCAM and gold standard assessments (see below). A convenience sample of patients was enrolled from May-August 2016, Monday through Friday from 8 a.m. to 4 p.m. The enrollment window was adapted to the routine availability of the participating specialist consultants, whose assessments served as the gold standard. If the number of potentially-eligible patients was not compatible with the availability of the physician on-duty, enrollment was performed consecutively in the order of patient admission to avoid bias. Patients fulfilled the following inclusion criteria: (i) minimum age of 70 years, and (ii) admission to the ED less than 12 hours before delirium screening started. This broad time

window hence allowed the inclusion of patients who were admitted to the ED in the late evening or during the night. Patients who were able to sit in the waiting room were only included if an examination room was available for conducting the tests; this measure ensured the patient's privacy and avoided false test results due to noisy or disruptive surroundings. Moribund patients were not enrolled for ethical reasons. Patients were excluded if: (i) they had previously been enrolled in the study, (ii) they had been placed in an isolation room (to avoid risk of infection), (iii) they were not able to complete the bCAM due to deafness, blindness, mutism, not being able to speak German, or, severe dementia (i.e. previous single digit Mini-Mental-Status-Examination (MMSE), if this was documented in the patient's medical report, or evidence from surrogate interviews or the medical record of the inability to perform basic personal care due to cognitive impairment), (iv) they were comatose or in a state of stupor.[23] Patients who met the enrollment criteria were then assessed for the presence of any exclusion criteria via medical examination as well as by perusing the electronic patient file and previous medical records.

Study Protocol

The bCAM was performed by emergency physicians (EPs) specializing in internal medicine. The medical section of the emergency medicine department has 20 full physician appointments. Of those, 11 are on a permanent basis and nine on rotation. The section is operated by three shifts, with three internal medicine residents, one internal medicine attending physician and one neurology resident present during day shifts and two internal medicine residents plus one attending physician in service during night shifts. To promote consistency in the generation of data, tests were mostly performed by the on-duty emergency medicine consultant. Each EP received a detailed introduction both to the evaluation process and the bCAM case report sheet, as detailed below.[24] The gold standard for delirium was

derived from the diagnosis made by one of four consultant physicians, i.e. consultant psychiatrists and neurologists, rotating on the emergency ward, each of whom had more than ten years of work experience and carried out delirium screening on a routine basis. Delirium assessment by a psychiatrist or neurologist is considered the gold standard as they are specialized on diseases of the central nervous system.[25] Nevertheless, in order to achieve concordant results and operationalize the diagnostic process, a diagnostic flowsheet requiring the neurologist or psychiatrist to review each item characteristic for delirium (as defined in the DSM-5[4]) was additionally used. Part of the gold standard diagnostic was the MMSE. With regard to the consultant physician's time constraints, it was conducted by a research assistant shortly before the consultant physician's personal approach to the patient. As the MMSE is an objective assessment it was considered valid even when performed by a research assistant. The bCAM and gold standard diagnostic were completed within three hours of enrollment in no fixed order. All assessors were blinded to each other's results. The research assistant also reviewed further patient data (see below) and entered them into a database, along with the test results. For this purpose the research assistant's blinding was lifted, but only after all tests for one patient had been completed.

Measurements

The bCAM was performed according to the official bCAM flowsheet.[24] All assessments were performed in German. The German translation of the bCAM was adapted from the validated German version of the CAM-ICU[26,27] and modified for Item 2, where the Attention Screening Examination from the CAM-ICU was replaced with the 'months of the year backwards' test (MOTYB) in the bCAM.[18] The bCAM flowsheet is shown in the supplementary section (Figure S1). It consists of four items testing for: (i) acute onset or fluctuating course of altered mental status (Item 1); (ii) inattention (Item 2); (iii) altered level of consciousness

(Item 3); (iv) disorganized thinking (Item 4). Delirium was deemed to be present in a patient if both Items 1 and 2 were positive, as well as either Item 3 or 4.[18]

In this study, the recognition of an altered mental status or a fluctuating course (Item 1) was based on information provided by the medical emergency team, ED staff, or surrogate interviews, whenever available. In inconclusive cases, telephone interviews were conducted whenever possible to receive further information about the patient. If there was not enough information to make a decision about Item 1, the EP had the option to mark "unsure", which, according to the study of Han et al., was considered positive if all other items were suggestive of delirium.[18] The presence of an altered level of consciousness (Item 3) was determined by the Richmond Agitation-Sedation Scale (RASS).[28] The official German version of the RASS scale was printed on the case report sheet used in this study to allow the EP to verify his/her judgement.[29] If a patient refused to answer a question or follow a command, the respective bCAM item was judged as positive.[18] The original bCAM version allows ending the testing at an early point of time, once delirium is ruled out according to the bCAM algorithm. In our study all bCAM assessments were conducted completely in order to obtain the necessary data for performing item subanalyses. At the end of the test, the EP noted the time needed to perform the bCAM and fill in the Flowsheet.

The gold standard consisted of a diagnostic report sheet that was developed prior to the study in order to minimize inter-observer variability (Figure 1). Delirium was deemed present if all five items were positive. Any item evaluated as negative immediately ruled out delirium and ended the assessment. In addition to their personal assessment, the consultant physicians based their assessments on the patient's performance on the MMSE. If a patient was not able

to complete all components of the MMSE, due, for example, to visual or motoric impairment, the final score was calculated by linear transformation of the actual score reached.[30] Finally, when the presence of delirium was confirmed by the consultant physician's assessment, he/she was asked to define the subtype of each delirium case in order to allow bCAM performance analysis for each subtype.

All test results and patient data were subsequently entered into a database, including any information about dementia mentioned in the patient's medical records. To screen for a possible selection bias, the diagnosis documented in the ED discharge report was compared between enrolled vs. excluded study patients. In cases of variable diagnoses amongst non-delirious patients, documentation was limited to that which best explained the patient's main complaints upon admission. For patients with delirium, the diagnosis that best explained its cause was recorded.

Additionally, we collected the following patient variables for all patients: Information about the severity of illness was taken from the electronic patient records by documenting the Emergency Severity Index (ESI). This triage tool is used in the ED for the stratification of patients from Level 5 (least urgent) to 1 (most urgent), according to acuity and resource needs.[31] The ESI of enrolled vs. excluded study patients was compared to identify a possible selection or spectrum bias. In addition, the Acute Physiology Score (APS) was calculated for each patient to examine whether the severity of illness influenced the diagnostic accuracy of the bCAM. The APS is part of the Acute Physiology and Chronic Health Evaluation II, which results from twelve routine physiological measurements in the form of a continuous variable, where higher scores indicate a higher severity of illness.[32]

Missing data was documented as such in the database and indicated as missing information in the respective tables and figures. If data was missing for a particular analysis, the respective patients were excluded from this calculation and the number of included patients was reported explicitly.

Data Analysis

Measures of central tendency and range of dispersion are indicated as mean values with 95% confidence intervals (CI) for continuous variables with normal distribution, or as median values with interquartile ranges (IQRs) for non-normally distributed data. Categorical variables are reported as absolute numbers and proportions. Statistical significance was reached if $p < 0.05$, as calculated by the t-test (continuous variables with normal distribution) Mann-Whitney-U-Test (continuous variables with non-normal distribution), or Fisher's Exact Test (categorical variables). The presence of normal distribution was determined by the Kolmogorov-Smirnov-Test. Based on an estimated delirium prevalence of 14%,[9] a sample-size calculation determined that a minimum of 200 patients should be enrolled in order to obtain reliable results for the diagnostic accuracy of the bCAM. Sensitivities, specificities, positive likelihood ratios (LR+), negative likelihood ratios (LR-), positive predictive values (PPV), negative predictive values (NPV) and odds ratios (ORs) were calculated with their 95% CIs for the overall bCAM result and individual bCAM items compared to the gold standard. The LR reveals the probability of obtaining a positive/negative bCAM test result in a patient with this disease, divided by the possibility for a patient without this disease to obtain the respective test result. The PPV/NPV reveals the probability of patients with a positive/negative bCAM test result that truly have/don't have delirium defined by the gold standard. In difference to the sensitivity, the PPV/NPV also takes the prevalence of the condition into account. The OR was used to calculate the increase in the risk of delirium,

according to the gold standard, with a positive bCAM test result. To evaluate inter-observer reliability between bCAM raters and gold standard raters, kappa statistics were calculated, to compare their respective results for both RASS scores and the assessment of acute onset.

A multivariable logistic regression was performed to determine whether bCAM sensitivity or specificity were influenced by the following covariates: age, the presence of dementia, the APS (indicating the patient's severity of illness) or by the MMSE-score (indicating the degree of the patient's cognitive impairment). Patients were only included if values for all covariates were available. bCAM sensitivity was calculated as the predicted probability of a positive bCAM test result for patients in whom delirium was detected according to gold standard diagnosis. The bCAM specificity was calculated as the predicted probability of a negative bCAM test result for patients in whom delirium was ruled out according to gold standard diagnosis.

Han et al. suggested a two-step-screening with the Delirium Triage Screen (DTS) as a first highly sensitive rule-out test and the bCAM as a second step to rule in delirium for all patients who had a positive result on the DTS.[18] In a secondary analysis we recreated the DTS using the RASS evaluation of the bCAM and the 100-7 subtraction test of the MMSE, which is an alternative test for the backwards spell test, which was used for the DTS. Patients were considered positive on the DTS if they had a RASS-score $\neq 0$ and/or made > 1 error in the subtraction test or if they were not able to perform the test. A sensitivity analysis of the bCAM was subsequently performed for all patients with a positive DTS. A descriptive comparison of sensitivity values for each delirium subtype was also performed to assess any differences in detection rates. We also performed an exploratory analysis of the bCAM

sensitivity and specificity when only Item 1 and 2 were considered. Delirium was consequently considered present when patients were positive for both of these two bCAM items. If one of these items was negative, delirium was ruled out. All data analyses were performed with IBM SPSS statistics 23 software (IBM Corp., IBM SPSS Statistics for Windows, Version 23.0, Armonk, NY) and MedCalc for Windows, version 17.9.2 (MedCalc Software, Ostend, Belgium).

Results

A total of 673 patients were screened during the enrollment period. Of these patients, 385 were subsequently excluded for various reasons (Figure 2).

A total of 288 patients were ultimately enrolled in the study (Table 1). The median age was 78 years (IQR 74 – 82) and 55.2% were female. Enrolled and excluded patients were similar in age and sex (Table 1). However, enrolled patients were significantly more likely to be categorized as ESI 2 ($p = 0.03$), whereas ESI 4 patients were more often excluded ($p < 0.01$). Enrolled vs. excluded patients were generally similar in terms of their diagnoses, except for significant differences in the neurological, cardiovascular and gastrointestinal inflammation group. Delirium was diagnosed by the specialist consultants for the gold standard in 46 out of the 288 enrolled non-surgical patients (16.0%). Hypoactive delirium was present in 26 patients, hyperactive delirium in 3 patients and mixed-type in 14 patients. There were 3 cases in which delirium was not classified by the specialist consultant ("no motor subtype").[14]

Eight different EPs acting as emergency medicine consultants performed 88.2% of the bCAMs. The testing procedure had a median duration of 3 minutes (IQR = 2-5) and the median time interval between the bCAM and gold standard assessment was 59 minutes (IQR = 20-97).

The diagnostic performance of the bCAM, and of each item, are shown in Table 2. While excellent results were obtained for specificity (93.8%, 95% CI 90.0 – 96.5), sensitivity was only 65.2% (95% CI 49.8 – 78.7). In comparison with the gold standard diagnosis, 30 bCAM results were true positive, 15 were false positive, 227 were true negative and 16 were false negative. Cohen's kappa indicated acceptable to moderate concordance between bCAM raters and gold standard raters for the following shared items: acute onset $\kappa = 0.50$ (95% CI 0.24 – 0.77), fluctuating course $\kappa = 0.25$ (95% CI 0.00 – 0.25) and altered level of consciousness $\kappa = 0.49$ (95% CI 0.33 – 0.64). The high level of sensitivity for Item 1 (acute onset of cognitive impairment) was due to the opportunity for the EPs to mark the answer "unsure"; this was made use of in 88 cases, in which 21 patients had delirium according to the gold standard assessment. The details for the bCAM results of the 16 false-negative patients are shown in the supplementary section (Figure S2). The secondary analysis for a two-step delirium screening showed a positive DTS for 102 patients. In this subgroup the bCAM had a sensitivity of 69.8% (95% CI 53.87 – 82.8). In the exploratory analysis of a bCAM version where only Item 1 and 2 of the bCAM were considered, the sensitivity was 82.6% (95% CI 68.6 – 92.2) and the specificity was 86.8% (95% CI 81.9 – 90.8).

The multivariable logistic regression showed no effect of age, APS or the presence of dementia on the bCAM sensitivity or specificity (Table 3). However, the bCAM sensitivity significantly increased when the patient scored less on the MMSE ($p = 0.03$), whereas the

specificity was not significantly influenced (p = 0.09). Comparison of delirium subtypes showed a trend towards better bCAM detection rates for hypoactive (18 out of 26 patients) and hyperactive delirium (two out of three patients), compared to mixed-type delirium (eight out of 14 patients).

Discussion

Diagnosing delirium in older ED patients is important in order to provide adequate therapy and improve patient outcomes.[1,2] Despite evidence for the benefits of an early diagnosis, such as a reduction of ED revisits within one month after discharge,[33] the majority of delirium cases remain undetected by EPs.[16] Several delirium screening instruments have been validated for the ED.[34] Among the short screening instruments (< 5 minutes), the mCAM-ED, which had been validated very recently after completion of our study, reached the best sensitivity (90%) and specificity (98%),[35] followed by the bCAM (84.0% sensitivity and 95.8% specificity), which – to the best of our knowledge - has only been validated for the ED in one previous study.[18]

The bCAM was developed as a delirium screening tool that meets the specific requirements for implementation in an ED.[18] The present study validated the bCAM in a German-language ED setting and thus provided reliable data for the validity of the German translation of the bCAM.[21]

In our study, the bCAM reached excellent specificity (93.8%) but only moderate sensitivity (65.2%). Although we applied DSM-5 criteria, the compatibility of our test results with those of previous studies using DSM-IV criteria - including the previous bCAM validation study[18] - is ensured. This is because the core content was maintained and we applied a less-strict interpretation of the DSM-5 criteria that rendered it more conformant with the DSM-IV.[36] Specifically, Criterion A was considered positive if the rater observed a disturbance in either attention or orientation, and Criterion B was applicable to the patient if the disturbance was either of acute onset, or took a fluctuating course.[36] The bCAM took a median time of 3 minutes to perform. This exceeds the time indicated in the literature[18] but is still acceptable for implementation in an ED. Of note, the assessment duration will be shorter when the bCAM is routinely performed, as it can be ended at an early point of time, once delirium is ruled out according to the bCAM algorithm, whereas in this study as well as in the study by Han et al.[18] the bCAM raters always performed the complete test. Compared to a reported delirium detection rate of 31.6% in the absence of a screening tool,[9] our study shows that the proportion of recognized delirium cases strongly improves when the bCAM is used. However, this screening tool needs to become more sensitive in order to reduce the number of missed cases, given the potential clinical implications.[5]

We found possible explanations for some of the false negatives, such as a strongly-fluctuating or rapidly-improving course of delirium, very subtle manifestations, or inconsistent information obtained through surrogate interviews about the patient's previous mental health situation. This does not, however, fully explain the poor sensitivity of the bCAM.
Low sensitivity might also result from milder manifestations of delirium that are missed during bCAM assessment.[18] Future studies should evaluate the bCAM diagnostic accuracy stratified by delirium severity, which can be determined using the delirium index.[37]

We did not observe any influence of age, the presence of dementia or severity of illness (indicated by the APS) on bCAM test accuracy. However, in line with the clinical observation that patients with a low MMSE score are more prone to the development of delirium, we found an increasing probability of true positives in those with lower MMSE scores. Of note, the MMSE itself is influenced by delirium severity[38] and it is not possible to distinguish between pre-existing cognitive impairment and the detrimental effect of delirium on cognition. In an additional analysis, we investigated bCAM performance according to the different types of delirium. In accordance with the literature, our study demonstrated that hypoactive delirium was the most frequent form observed in these elderly patients.[4,5] We found moderate bCAM sensitivity for all motoric subgroups, with the lowest sensitivity for the mixed type, which might be explained by its rapid fluctuation and hence potentially varying profile observed by each of the assessors. Subgroup analysis was not reported by Han et al,[18] and also in our study, the low number of cases in the hyperactive and mixed group limits conclusions.

Our reconstruction of the two-step delirium screening proposed by Han et al.[18] showed that it does not considerably improve the sensitivity of delirium screening (69.8% vs. 65.2%). However, it reduces the number of bCAM screenings. The development of a more sensitive first step delirium screening tool, which can also be conducted by non-physicians, would therefore reduce the EPs' work load.

We also performed sub-analyses of single bCAM items to further investigate the reason for the low sensitivity. Single-item performances are not mentioned by Han et al.[18] Previous studies have observed that the MOTYB is best suited for screening of inattention.[39–41] However, these studies were conducted in English-speaking countries and screening

instruments run the risk of losing validity when they are translated into a foreign language.[21] Therefore it is crucial, that our study confirms the good level of sensitivity (87%) of the MOTYB for delirium screening in German EDs (Table 2). On the other hand, most false-negative bCAM results were due to the negative test results for altered level of consciousness (RASS) and disorganized thinking (Figure 1). In concordance with another study[42] we observed a poor RASS sensitivity for delirium of 57.8%, even when referring to the consultant physicians' assessment. There was an indication that Han et al. had assessed more patients with altered level of consciousness. They found, that the RASS was 82 – 84% sensitive for delirium for older ED patients,[43] which probably contributed to the high bCAM sensitivity compared to our study. Considering the results of our single item analysis, we performed an exploratory analysis which revealed a considerable increase of the bCAM sensitivity if only Item 1 and 2 were considered (65.2% for the original bCAM vs. 82.6% in the exploratory analysis). Our results are affirmed by the good validity of the MOTYB for delirium screening found by O'Regan et al.[39] A condensed delirium screening assessment consisting only of the evaluation of an acute onset of cognitive impairment and of the MOTYB should therefore be considered and validated in a prospective study.

We also found that Cohen's kappa for Item 1 and 3 only showed moderate concordance between bCAM raters and gold standard raters (κ=0.25 – 0.5). This has probably contributed to the low bCAM sensitivity, which might therefore be improved by providing training for the bCAM assessment for all EPs. Of note, the bCAM could generally also be performed by trained nurses or nurse assistants.[12]

Limitations

Some limitations of the study warrant consideration. Due to the limited availability of the EPs and psychiatry/ neurology consultants, we were unable to include all potentially-eligible patients. However, enrolling the patients in order of admission most likely kept the selection bias to a minimum. Furthermore, we were not able to screen patients at night, the time at which first onset of delirium is frequently observed.[4] Future studies should therefore aim for complete inclusion of all eligible patients presenting to the ED.

We predefined a 3-hour time window for the two assessments to ensure the feasibility of the study. This might have resulted in discordant observations in some cases, due to the highly fluctuating course of delirium[4] and possible influences of interim interventions, such as pain management or reorientation measures. However, the comparison of true positive and false negative bCAM results showed no significant difference concerning the varying time intervals between bCAM and gold-standard assessment (data not shown). There is an unequal distribution of some symptoms in enrolled vs. excluded patients. Patients with neurological health problems were more likely to meet the exclusion criteria for testing suitability. In contrast, cardiovascular patients were more likely to be included because the ED treatment algorithm for this patient group allowed more time to perform the test. Moreover, the exclusion of moribund patients has probably caused a skew away from hypoactive delirium and might consequently have affected the bCAM sensitivity and specificity.

Furthermore, severely-ill patients (as indicated by a higher ESI-Score) were more likely to be included (see Table 1). This was most probably due to the limited availability of examination rooms for ESI 4 patients, who were usually asked to remain in the waiting room. This

approach might have caused spectrum bias towards an overestimation of the bCAM sensitivity.[44] However, another study that declared the same problem was nevertheless able to demonstrate the limited influence of this aspect by analyzing bCAM validity in a subgroup of hospitalized patients and obtaining results similar to those for the total study sample.[18] Our study was conducted at a single ED in patients who were 70 or older, as the incidence of delirium in elderly people is especially high.[1] Therefore, the present results might not pertain to other settings or populations.

We computed the inter-rater reliability neither of the bCAM, nor of the gold standard, given that it would have exceeded the patients' capacity and could have induced training effects in the patients. Video or audio-taping was deemed inappropriate for privacy reasons. Instead, we limited the number of gold standard assessors to clinical experts.

When comparing our results to those of Han et al.,[18] it has to be considered that the two study populations slightly differ. As our age cut-off was 70 years vs. 65 years at Han's study, our median age was higher (78 vs. 73.5 years), furthermore Han et al. only included one patient per day and their enrolled patients tended to have a higher severity of illness than in our study (for example ESI 2: 65% vs. 33.3%).[18]

Conclusion

The bCAM is a short delirium screening tool that shows good specificity but modest sensitivity in a German ED. Since delirium may sometimes be the only indicator of a life-threatening health problem in elderly patients and has a dramatic effect on patient outcome, its detection should hold top priority. Given the lack of validated German-language delirium

screening tools for EDs, the bCAM should be applied with caution, and its sensitivity may be improved by providing training for all examiners. The bCAM still requires further modification to boost sensitivity.

References

1. Inouye SK, Westendorp RGJ, Saczynski JS. Delirium in elderly people. The Lancet 2014;383:911–22.

2. Kennedy M, Enander RA, Tadiri SP, Wolfe RE, Shapiro NI, Marcantonio ER. Delirium risk prediction, healthcare use and mortality of elderly adults in the emergency department. J Am Geriatr Soc 2014;62:462–9.

3. Fick DM, Hodo DM, Lawrence F, Inouye SK. Recognizing delirium superimposed on dementia:, Assessing nurses' knowledge using case vignettes. J Gerontol Nurs 2007;33:40–49.

4. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. American Psychiatric Association. Arlington, VA: American Psychiatric Association, 2013. 992 p.

5. Han JH, Wilber ST. Altered mental status in older patients in the emergency department. Clin Geriatr Med 2013;29:101–36.

6. Ahmed S, Leurent B, Sampson EL. Risk factors for incident delirium among older people in acute hospital medical units, A systematic review and meta-analysis. Age Ageing 2014;43:326–33.

7. Han JH, Zimmerman EE, Cutler N, Morandi A, Schnelle J, Dittus RS, et al. Delirium in older emergency department patients: recognition, risk factors, and psychomotor subtypes. Acad Emerg Med 2009;16:193–200.

8. Han JH, Shintani A, Eden S, Morandi A, Solberg LM, Schnelle J, et al. Delirium in the emergency department, An independent predictor of death within 6 months. Ann Emerg Med 2010;56:244-252.e1.

9. Singler K, Thiem U, Christ M, Zenk P, Biber R, Sieber CC, et al. Aspects and assessment of delirium in old age. First data from a German interdisciplinary emergency department. Z Gerontol Geriatr 2014;47:680–5.

10. Gower LEJ, Gatewood MO, Kang CS. Emergency department management of delirium in the elderly. West J Emerg Med 2012;13:194–201.

11. Park JM, Stern TA. Delirious Patients. In: Stern TA, Fricchione GL, Cassem Ned H., Jellinek M, Rosenbaum JF, eds. Massachusetts General Hospital handbook of general hospital psychiatry. 6th ed. Philadelphia: Saunders/Elsevier, 2010:93–104.

12. LaMantia MA, Messina FC, Hobgood CD, Miller DK. Screening for delirium in the emergency department, A systematic review. Ann Emerg Med 2014;63:551-560.e2.

13. Barron EA, Holmes J. Delirium within the emergency care setting, occurrence and detection: a systematic review. Emerg Med J 2013;30:263–8.

14. Meagher D. Motor subtypes of delirium, Past, present and future. Int Rev Psychiatry 2009;21:59–73.

15. Han JH, Eden S, Shintani A, Morandi A, Schnelle J, Dittus RS, et al. Delirium in older emergency department patients is an independent predictor of hospital length of stay. Acad Emerg Med 2011;18:451–7.

16. Hustey FM, Meldon SW, Smith MD, Lex CK. The effect of mental status screening on the care of elderly emergency department patients. Ann Emerg Med 2003;41:678–84.

17. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in Older Persons, Advances in Diagnosis and Treatment. JAMA 2017;318:1161–74.

18. Han JH, Wilson A, Vasilevskis EE, Shintani A, Schnelle JF, Dittus RS, 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;62:457–65.

19. Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). JAMA 2001;286:2703–10.

20. Inouye SK. Clarifying Confusion, The Confusion Assessment Method. Ann Intern Med 1990;113:941–8.

21. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures, Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8:94–104.

22. Cohen JF, Korevaar DA, Daniel A, Altman DG, Bruns DE, Gatsonis CA, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies, Explanation and elaboration. BMJ Open 2016;6:e012799.

23. Boehm L, Ely EW, Pun BT, Stollings J. Confusion assessment method for the ICU (CAM-ICU), The complete training manual, 2014. (Accessed May 24, 2017, at http://www.icudelirium.org/docs/CAM_ICU_training.pdf).

24. Han JH. Brief confusion assessment method (bCAM) Flow Sheet. (Accessed May 24, 2017, at http://www.icudelirium.org/docs/bCAM_Flowsheet.pdf).

25. Su JA, Tsai CS, Hung TH, Chou SY. Change in accuracy of recognizing psychiatric disorders by non-psychiatric physicians, Five-year data from a psychiatric consultation-liaison service. Psychiatry Clin Neurosci 2011;65:618–23.

26. Guenther U. Die "Confusion Assessment Method (CAM)" für Intensivstationen (CAM-ICU), Übungsleitfaden. (Accessed May 24, 2017, at http://www.icudelirium.org/docs/CAM_ICU_training_German.pdf).

27. Guenther U, Popp J, Koecher L, Muders T, Wrigge H, Ely EW, et al. Validity and reliability of the CAM-ICU Flowsheet to diagnose delirium in surgical ICU patients. J Crit Care 2010;25:144–51.

28. Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. Am J Respir Crit Care Med 2002;166:1338–44.

29. Schäfer UB, Massarotto P, Lehmann A, Wehrmuller C, Spirig R, Marsch S.

Übersetzungsverfahren eines klinischen Assessmentinstrumentes am Beispiel der Richmond Agitation-Sedation Scale (RASS) [Translation of the Richmond Agitation-Sedation Scale (RASS) for use in German Swiss intensive care units]. Pflege 2009;22:7–17.

30. Busse A, Sonntag A, Bischkopf J, Matschinger H, Angermeyer MC. Adaptation of dementia screening for vision-impaired older persons Administration of the Mini-Mental State Examination (MMSE). Journal of Clinical Epidemiology 2002;55:909–15.

31. Gilboy N, Tanabe T, Travers D, Rosenau AM. Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4. Implementation Handbook 2012 Edition. Agency for Healthcare Research and Quality. Rockville, MD: Agency for Healthcare Research and Quality, 2011. 114 p. (AHRQ Publication; vol. 12-0014).

32. Knaus WA, Draper, Elizabeth A. Wagner, Douglas P., Zimmerman JE. APACHE II: A severity of disease classification system. Crit Care Med 1985;13:818–29.

33. Delaney M, Pepin J, Somes J. Emergency Department Delirium Screening Improves Care and Reduces Revisits for the Older Adult Patient. J Emerg Nurs 2015;41:521–4.

34. Tamune H, Yasugi D. How can we identify patients with delirium in the emergency department? A review of available screening and diagnostic tools. Am J Emerg Med 2017;35:1332–4.

35. Hasemann W, Grossmann FF, Stadler R, Bingisser R, Breil D., Hafner M, et al. Screening and detection of delirium in older ED patients, Performance of the modified Confusion Assessment Method for the Emergency Department (mCAM-ED). A two-step tool. Intern Emerg Med 2017.

36. Meagher DJ, Morandi A, Inouye SK, Ely EW, Adamis D, Maclullich AJ, et al. Concordance between DSM-IV and DSM-5 criteria for delirium diagnosis in a pooled database of 768 prospectively evaluated patients using the delirium rating scale-revised-98. BMC Med 2014;12:164.

37. McCusker J, Cole MG, Dendukuri N, Belzile E. The delirium index, a measure of the severity of delirium, New findings on reliability, validity, and responsiveness. J Am Geriatr Soc 2004;52:1744–9.

38. Cole MG, Bailey R, Bonnycastle M, et al. Partial and No Recovery from Delirium in Older Hospitalized Adults, Frequency and Baseline Risk Factors. J Am Geriatr Soc 2015;63:2340–8.

39. O'Regan NA, Ryan DJ, Boland E, Connolly W, McGlade C, Leonard M, et al. Attention! A good bedside test for delirium? J Neurol Neurosurg Psychiatry 2014;85:1122–31.

40. Voyer P, Champoux N, Desrosiers J, Landreville P, Monette J, Savoie M, et al. Assessment of inattention in the context of delirium screening: one size does not fit all! Int Psychogeriatr 2016;28:1293–301.

41. Adamis D, Meagher D, Murray O, O'Neill D, O'Mahony E, Mulligan O, et al. Evaluating attention in delirium: A comparison of bedside tests of attention.

42. Grossmann FF, Hasemann W, Kressig RW, Bingisser R, Nickel CH. Performance of the modified Richmond Agitation Sedation Scale in identifying delirium in older ED patients. Am J Emerg Med;35:1324–6.

43. Han JH, Vasilevskis EE, Schnelle JF, Shintani A, Dittus RS, Wilson A, et al. The Diagnostic Performance of the Richmond Agitation Sedation Scale for Detecting Delirium in Older Emergency Department Patients. Acad Emerg Med 2015;22:878–82.

44. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Acad Emerg Med 2013;20:1194–206.

Supporting Information:

The following supporting information is available in the online version of this paper:

Figure S1. Brief Confusion Assessment Method (bCAM) Flowsheet.

Figure S2. Flow-diagram of false-negative bCAM results.

Table 1
Patient Characteristics

Characteristic	**Enrolled Patients (n=288)**	**Excluded Patients (n=385)**	**p-value[1]**
Sex, absolute (%)			
Male	129 (44.8)	186 (48.3)	0.39
Female	159 (55.2)	199 (51.7)	
Age (yr), median (IQR)	78 (74 - 82)	78 (74 - 84)	0.46
ESI, median	3	3	0.07
ESI, absolute (%)			
1	3 (1.0)	12 (3.1)	0.11
2	124 (43.1)	133 (34.5)	0.03*
3	158 (54.9)	211 (54.8)	1
4	2 (0.7)	19 (4.9)	< 0.01*
5	1 (0.3)	1 (0.3)	1
Missing Information		9 (2.4)	
Diagnosis, absolute (%)			
Neurological	51 (17.7)	94 (24.4)	0.04*
Cardiovascular	105 (36.5)	106 (27.5)	0.02*
Respiratory	14 (4.9)	22 (5.7)	0.73
Infection/urinary passage/sepsis	42 (14.6)	54 (14.0)	0.91
Tumor-related problems	8 (2.8)	3 (0.8)	0.06
Bleeding/anaemia	15 (5.2)	26 (6.8)	0.42
Exsiccosis	6 (2.1)	7 (1.8)	1
Other types of organ failure	6 (2.1)	13 (3.4)	0.36
Intoxication or substance withdrawal	9 (3.1)	8 (2.1)	0.46
Metabolic/ endocrine	11 (3.8)	11 (2.9)	0.52
Gastrointestinal inflammation	7 (2.4)	0 (0.0)	< 0.01*
Other	14 (4.9)	40 (10.4)	0.01*
Missing Information		1 (0.3)	

APS (n = 287)[2]	
Median (IQR)	7 (5 - 9)
Documented dementia (excluding end-stage dementia), absolute (%)	
Yes	14 (4.9)
No	274 (95.1)
Procedure, absolute (%)	
Hospital discharge	52 (18.1)
Outpatient treatment	12 (4.2)
Admission to hospital	191 (66.3)
Intensive Care Unit	12 (4.2)
Stroke-Unit	12 (4.2)
Patient signed out against medical advice	7 (2.4)
Acute intervention (i.e. coronary angiography)	2 (0.7)

* Statistically-significant difference

[1] T-Test for continuous variables with normal distribution; Mann-Whitney-U-Test for continuous variables with non-normal distribution; Fisher's Exact Test for categorical variables

[2] One patient was excluded from this calculation, because the required parameters were missing.

Table 2
Diagnostic Performance of the bCAM

Items	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	LR+ (95% CI)	LR- (95% CI)	OR (95% CI)
bCAM result, overall	65.2 (49.8-78.7)	93.8 (90.0-96.5)	66.7 (54.0-77.3)	93.4 (90.5-95.5)	10.5 (6.2-17.9)	0.37 (0.25-0.55)	28.4 (12.7-63.2)
Item 1a	91.3 (79.2-97.6)	69.8 (63.6-75.6)	36.5 (31.8-41.6)	97.7 (94.3-99.1)	3.0 (2.5-3.7)	0.12 (0.05-0.32)	24.3 (8.4-70.3)
Item 1b	68.9 (53.4-81.8)	69.7 (63.5-75.4)	29.8 (24.4-35.8)	92.3 (88.5-94.9)	2.3 (1.7-3.0)	0.45 (0.29-0.69)	5.1 (2.6-10.1)
Item 2	87.0 (73.7-95.1)	68.6 (62.3-74.4)	34.5 (29.8-39.5)	96.5 (92.9-98.3)	2.8 (2.2-3.4)	0.19 (0.09-0.40)	14.6 (5.9-35.8)
Item 3	47.8 (32.9-63.1)	95.9 (92.5-98.0)	68.8 (52.8-81.3)	90.6 (88.0-92.7)	11.6 (5.9-22.8)	0.54 (0.41-0.72)	21.3 (9.0-50.1)
Item 4	73.9 (58.9-85.7)	86.4 (81.4-90.4)	50.8 (41.8-59.6)	94.6 (91.4-96.6)	5.4 (3.8-7.8)	0.30 (0.19-0.49)	17.9 (8.5-38.1)

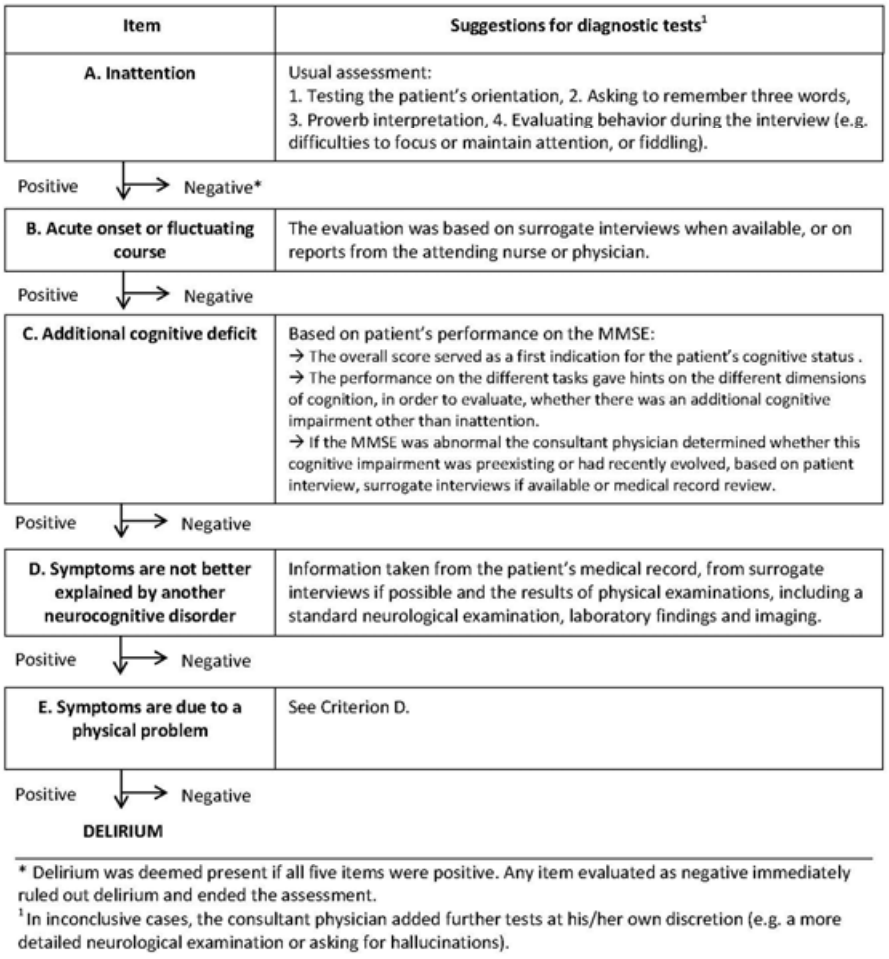
Item 1a = acute onset of altered mental status, Item 1b = fluctuating course, Item 2 = inattention, Item 3 = altered level of consciousness, Item 4 = disorganized thinking, PPV = positive predictive value, NPV = negative predictive value, LR+ = positive likelihood ratio, LR- = negative likelihood ratio, OR = odds ratio, CI = confidence interval

Table 3
Multivariable logistic regression: influence of patient characteristics on the bCAM validity

Covariate[1]	Sensitivity		Specificity	
	Odds ratio[2] (95% CI)	p-Value	Odds ratio[2] (95% CI)	p-Value
Age	1.03 (0.92 – 1.16)	0.57	1.03 (0.94 – 1.12)	0.31
APS	0.97 (0.78 – 1.21)	0.76	0.94 (0.79 – 1.13)	0.52
MMSE	0.88 (0.79 – 0.99)	0.03*	0.90 (0.79 – 1.02)	0.09
Documented dementia	0.74 (0.11 – 4.78)	0.75	1.29 (0.14 – 12.21)	0.82

[1] 285 patients were included in the multivariable logistic regression. 3 Patients were excluded due to missing data for the APS (1 patient) and the MMSE (2 patients).
[2] The odds ratio gives the relative amount by which the odds of a positive bCAM test result (compared to the gold standard diagnostic) increases (for odds ratios >1) or decreases (for odds ratios <1), when the value of the covariate is increased by 1 unit/ is considered present (for the categorical covariate "documented dementia").
* Statistically significant
CI = Confidence Interval, APS = Acute Physiology Score, MMSE = Mini-Mental Status Examination

Figure 1
Flowsheet for gold standard assessment



* Delirium was deemed present if all five items were positive. Any item evaluated as negative immediately ruled out delirium and ended the assessment.
[1] In inconclusive cases, the consultant physician added further tests at his/her own discretion (e.g. a more detailed neurological examination or asking for hallucinations).

Figure 2
Flow-diagram of the enrollment procedure