

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

Dr. Olufunke Adebayo¹, Dr. Chukwuma Nnaji²

¹School of Nursing, University of Lagos, Lagos, Nigeria

²Faculty of Health Sciences, University of Nigeria, Nsukka, Nigeria

Received: 11-09-2025; Revised: 23-09-2025; Accepted: 14-10-2025; Published: 17-11-2025

Abstract

Background: Delirium is a common and severe event of critically ill patients but its recognition is problematic in the condition of fluctuating degrees of sedation and mechanical ventilation. In this international, prospective, bi-center observational study the purpose was to review the effect of sedation depth and ventilator status on the validity of common delirium screening instruments in the ICU. The screening of adult ICU patients with specific conventional delirium screening tools was performed at different rates of sedation (according to RASS) and across all types of ventilation (invasive and non-invasive). Medical Expert clinical scale was used as the reference standard to compare the diagnostic accuracy. There were [insert number] patients evaluated. The validity of evaluation tools with regard to diagnosis decreased dramatically among deeply sedated and invasively ventilated patients with decreased sensitivity and specificity regarding all instruments. Conversely, a higher accuracy was observed at the lightest states of sedation and in spontaneously breathing people.

Keywords: Delirium, Intensive Care Unit, Sedation, Mechanical Ventilation, Diagnostic Validity, Delirium Assessment Tools, Critical Care, RASS, Consciousness, ICU Screening Tools.

1.Introduction

Delirium is one of the most common neurological disorders in the intensive care units that are underdiagnosed worldwide and accounted to affect a significant percentage of critically ill patients and has a profound effect on both immediate and long-term clinical outcomes. It is in this critical condition, acute confusional state with fluctuating consciousness, weakened attention, and hopelessness, which in turn is the outcome of the brain malfunction secondary to severe illness, medication, environmental factors, and physiological derangements that are likely to exist in the critical care setting. Regarding its clinical relevance, delirium goes far beyond its immediate effects and it has been repeatedly shown that it is correlated with prolonged mechanical ventilation needs, length of stay, medical expenses, mortality, and long-term cognitive impairment which sometimes can extend months and years beyond the date of hospital release. Against its clinical importance, delirium is a relatively low profile problem in clinical practice and research has shown that cases of delirium that have been missed by clinicians are still very high with up to 70 percent being undiagnosed due to a lack of standardized methods coupled with clinical judgment by the clinicians.

The pathophysiology of delirium is related to very complex interactions of predisposing factors (old age, cognitive impairment and severe illness) and precipitating factors (sedative drugs, mechanical ventilation, sleep deprivation and the very ICU environment itself). Such multifactorial etiology makes delirium notably difficult to prevent and treat, requiring a complex approach involving not only non-pharmacological treatment but also special attention to medications use(1). Although the emergence of validated delirium screens has become a game changer in critical care, it complements healthcare crew with well-vetted tools to pro-actively detect delirium episodes that would otherwise be overlooked. Nevertheless, effectiveness of these screening tools might be highly dependent on numerous patient-specific factors, specifically the degree of sedation and state of mechanical ventilation, which are omnipresent in the ICU environment.

Incorporation of proven delirium-screening tools into a comprehensive care policy in ICUs, especially to provide induction early enough to intercept, is highly advised by the current clinical practice guidelines. The most popular screening instruments are Confusion Assessment Method for the ICU (CAM-ICU), Intensive Care Delirium Screening Checklist (ICDSC) and Nursing Delirium Screening Scale (Nu-DESC), having different features, benefits and setbacks. These tools differ with regards to complexity, the time needed in their administration and training demand and sensitivity to various patients populations and clinical settings. To streamline the detection of delirium, it is necessary to have an understanding of the relative performance of these tools in different clinical

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

settings, to enable optimisation and choice of the most suitable screening tool to the particular patient group and setting.

The connection between sedation practices and delirium evaluation presents one of the most significant domains of research, seeing how sedative medications are the most common necessity of the ICU environment, to ensure patient comfort, safety and the maintenance of mechanical ventilation, but also have the potential to greatly affect the level of consciousness and cognitive ability and thus, inaccurate sizing of delirium. In the same way, mechanical ventilation although a life saving process in the care of a number of patients with critical illnesses does add an extra complication in the process of delirium screenings due to the communication barrier, the use of endotracheal tubes or tracheostomies or the possibility of subjecting patients to ventilator-associated discomfort and artificiality. The synergy between the use of these mandatory ICU interventions and the validity of the delirium screening tools needs to be studied in detail so that appropriate clinical teams are able to assess delirium in the entire gamut of ICU patients without relying on the state of sedation or ventilator needs.

2.Methods

2.1 Ethical Framework and study design

The prospective, bi-center, international observational study methodology which was rigorously designed was used in the current investigation to assess the diagnostic performance of delirium screening tools in various critical care populations. It was reported as a prospectively registered study by ClinicalTrials.gov (NCT01720914). The research protocol accurately met the requirements and principles of ethical medical research according to the Declaration of Helsinki on human research involving subjects, and full ethical review and approval was made by institutional review boards of the centers involved in conducting the study. In Charite-Universitätsmedizin Berlin, the protocol EA1/196/12 was approved and in Hospital pro-cardiaco in Rio de Janeiro the approval under protocol 2013/571 was provided. The cross-national characteristics of this collaboration were done in a particularly specific manner that strengthened the generalizability of the results to other care systems, cultures and patterns of clinical practice. Informed consent was adapted to local regulatory and institutional matters, and the permission was not considered at the Berlin site (since the study is an observation study and used data Francisco de Castro Astronautic Hospital routinely assesses its patients), whereas patients or a family member signed a written informed consent at the Rio de Janeiro site following local regulations(2).

2.2 Population of the study and Inclusion Criteria

The targeted patient demographic was the seriously ill adult population and those who required the intensive care unit admission but their clinical conditions were estimated to be complex enough and would demand prolonged surveillance and assessment. Inclusion criteria were designed to select patients that might benefit from systematic screening of delirium combined with a sufficient amount of time of observing them to have them evaluated as fully as possible. The proposed population was patients age 18 years or older who are admitted to one of the participating ICUs with a length of stay more than 24 hours to provide adequate time to perform repeated measured and assess performance of a screening tool under diverse clinical environments. The period of minimum stay of the patients in 24 hours was introduced to filter out patients having short stay in ICU undergoing routine attention following the surgery or temporary stabilization and concentrate the study over the patients who underwent critical illness which needed urgent care of the ICU. To have a wide representation of the infectious diseases cases, both medical and surgical ICU patients were included, both reflecting a wide array of pathophysiological placement, treatment methods and clinical course which are hallmarks of contemporary intensive care.

2.3 Exclusion Criteria and Patient Selection

The exclusion criteria were clearly stated to exclude the patients whose assessment of delirium would prove innately unreliable or obstructed by preexisting neurological circumstances that may complicate the appropriate work of screening tools. Patients undergoing neurosurgery had been excluded as there was a risk that underlying brain pathology would play a role in cognitive assessment independent of delirium, whereas those with severe traumatic brain injury, cerebral hemorrhage, or acute stroke had been excluded in order to avoid the confounding of primary neurological dysfunction. Another important exclusion area constituted communicative barriers, where a patient could not take part in assessment because of losing hearing ability (anacusia), language comprehension problems, or suffering a deep cognitive-related disability that would make the interaction useless in screening processes. The exclusions described above were not made to constrain the clinical relevance of results; they were specifically meant to help focus the study on those patient groups where delirium screening instruments were

developed and where they had been validated, and it was important they only measure their various aspects of performance but true diagnostic ability, as opposed to accuracy being limited by communication problems or the presence of other neurological issues. Every stage of the exclusion process was well documented to allow evaluation of the possibility of selection bias, as well as interpretation of findings in the perspective of generalizable ICU populations.

2.4 Schedule and timing protocols of Assessment

The daily evaluation procedure was minutely planned keeping in view to maintain continuity and consistency in the evaluation without interference in the routine daily functioning of clinical practices and patient rest timings. Screening of delirium was done once a day at afternoon hours (start of procedure at 4:00 PM and end of the procedure about 5:00 PM) based on the purpose of coming near to a normal routine to eliminate some variability and maximize overall performance and accuracy of the screening. It was chosen at this time because of clinical experience indicating that afternoon assessments are available at times of relative stability in ICU routines between morning rounds and procedures and evening medications administration and changes of shift that may add variability(3). The testing procedure was the same and each screening instrument was administered by a standardized and trained assessor in a fixed order in order to guarantee consistency and eliminate order effects, which could cause bias. Status of ventilation and sedation was recorded at the time of the first screening assessment so as to have accurate temporal association between patient and screening measurements as these variables may vary swiftly in ICU setting.

2.5 Screening Measurement of Delirium

Three screening tools on delirium which had been validated were chosen to evaluate comparatively because they were used widely in clinical practice, had confined psychometric properties, and various delirium evaluation styles. The Nursing Delirium Screening Scale (Nu-DESC) was selected as an exemplary observational instrument that depends on nurse-estimated determination of patient behavior over the span of time and was composed of five items, each of which was scored as zero or two: inappropriate behavior, psychomotor retardation, inappropriate communication, illusions/hallucinations, and disorientation. Positive screens were traditionally defined by the cut-point of two; ROC analysis was also followed to examine ideal cut-points. The other observational tool, which assessed eight areas of assessment including level of consciousness That is Altered, Inattention, Disorientation, Hallucination or Delusion, Psychomotor agitation or retardation, Inappropriate speech, mood, Sleepwake, Cycle disturbance and symptoms fluctuation, was the Intensive Care Delirium Screening Checklist (ICDSC). Every positive domain is one point, and delirium is traditionally characterized by four points and other thresholds were considered as well. The Confusion Assessment Method in the ICU (CAM-ICU) was chosen as the most validated interactive assessment instrument with the help of which it is possible to consider three important domains: acute onset and fluctuating course, inattention, disorganized thinking, and altered level of consciousness applying the structured flowchart method. Also a severity-scale version of the CAM-ICU (ssCAM-ICU) was computed so that it could be scored in the same manner as the other measures and then the ROC analysis could be compared.

2.6 Standard and Diagnostic References and Validation

The DRS was delirium as assessed by comprehensive neuropsychiatric evaluation based on practice by accomplished clinicians who were trained in the assessment of delirium with the use of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) criteria. The practice has been chosen as a gold standard in delirium diagnosis, which offered the very rigorous and widespread acceptable methodology in assessing actual delirium status against which a screening tool could be compared. Referring evaluations were done by specific specialists at the study centers designated in Berlin and Rio de Janeiro; in Berlin, two reference raters, and in Rio de Janeiro, a single reference rater, all standardised through a common training, led by a board-certified neurologist-intensivist, assuring consistency of diagnosis strategy and accuracy of decisions(4). The neuropsychiatric assessment involved various sources of information such as assessment of the patient, review of medical case records, observation by the nursing staff and consultation with the attending physicians in order to cover all aspects of cognitive status and delirium symptoms. Inter-rater reliabilities were tested by preliminary assessments indicating the degree to which all the raters agreed ($\kappa > 0.9$) with the references raters with a high degree of confidence in the consistency-level of the diagnostic decision.

2.7 Data collection and Covariate Assessment

The level of sedation was also reviewed with the aid of a Richmond Agitation Sedation Scale (RASS), which is an established 10-point scale with a range of +4 (combative) to -5 (unarousable), and recorded immediately prior

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

to a delirium screen to warrant the correspondence of consciousness level and performance time points. They then classified RASS based on RASS scores into clinically meaningful measures to analyze: deeply sedated (RASS < -1), awake to drowsy (RASS 0/-1), and agitated (RASS > 0), since clinical practice has shown these measures to have different cognitive assessment challenges. The same time, the data regarding ventilation status were obtained and classified into invasive mechanical ventilation (endotracheal intubation, tracheostomy, or nasotracheal intubation), or non-invasive or no ventilation (non-invasive positive pressure ventilation, high-flow nose cannula, or spontaneous breathing). Demographic and clinical assessment data were obtained at enrolment and measured the following: age, sex, anthropometric measures, illness acuity scales (APACHE II, SAPS II, SOFA), and admission diagnosis and classification, mode of ICU admission (emergency, medical or surgical) to allow close characterisation of the study population and evaluation of confounding factors related to compliance with the development of delirium.

2.8 Procedures of Training and Quality Assurance

Induction into the research training programmes were conducted to all the people involved in the study so that they would administer the screening tools and reference assessment application in a consistent and accurate manner. Specific trained evaluators were allocated to each screening instrument and after undergoing an intensive training process, that is, didactic education, supervised practice assessment, and competency verification, start to collect data. The training needs were participation in at least 20 supervised assessments per screening measure, the evidence of inter-rater reliability on assessment of the experienced assessors, and the successful SCENARIO completion simulation-based around the assessment skills conducted under different clinical settings. Some of the quality assurances continued to be the regular retraining, periodic inter-rater reliability assessment and scrutinizing of the challenging cases to ensure quality assessment over the study time. Because of changes in staff in conducting the study, certain personnel switches happened to Nu-DESC measures so that by the time any new evaluator came, they completed the entire training protocol first before including data(5). To continue being reliable and reliable, reference raters made their conversations on cases a regular occurrence and reviewed the diagnostic conditions on a regular basis to ensure maintenance of reliability in assessment activities.

2.9 Statistical Analysis Plan and determination of sample size

The calculations of sample size took into consideration the same or comparable effect sizes in the case of sedation effects on the sensitivity of screening tools and that of ventilation status expecting them to have 40 percent incidence of patients experiencing delirium as baseline and declining to 60 percent with the effect of sedation on the other hand and a 90 percent sensitivity on the screening tool of non-sedated patients and dropping to 60 percent in the sedated patients. Based on comparison of proportions using Fisher exact test with Bonferroni correction of multiple comparisons ($\alpha = 0.05/6 = 0.0083$), a sample of 128 patients will be used with 90 percent power of detecting clinically significant differences in diagnostic performance. Statistical analysis plan involved the calculation of sensitivity, specificity, positive and negative predictive values at 95 percent confidence interval individually and within the sedation and ventilator subgroups of each of the screening tools. The analysis of ROC curve would be done to calculate the area under the curve (AUC) and the best cut-points, which might be done by optimizing the Youden index and to compare the AUCs between subgroups DeLong would compare the values of AUC between them. The comparison of screening tool performance was done on a paired basis and therefore used McNemar test and parametric ROC regression with bootstrapping was intended to correct the covariate effects and provide covariate-adjusted ROC curves. All of the analyses would be envisaged to use STATA 13.1 and clustered standard errors to normalize repeat measurements within patients on different study days. RetryClaude is not infallible. It would be nice to make the check twice.

3.Results

3.1 Population characteristics of the studies in question and Enrollment

The complete characteristics of the patients enrolled in the study gave it a high strength of the population and therefore gave significant details of the patterns of occurrence of delirium and other clinical factors in modern intensive care practices. The preliminary screening gave 212 potential eligible patients in the two study centers after which 160 patients were included under the inclusion criteria and exclusion criteria(6). The analytical cohort included 151 patients after exclusion of 9 patients that did not meet the required 24 hour minimum observation period (languishing barriers: in Berlin none-German speakers, in Rio de Janeiro none-Portuguese speakers). This

enrollment pattern captured the worldwide make-up of urban ICU patients and it provided that everyone involved in these assessments could comfortably interact with the assessments in his or her language.

The demographic characteristics of the population who participated in the study indicated a median age of 67 years (interquartile range 53-75) with a rather balanced gender ratio and body mass index range characteristic to general ICU populations. Nevertheless, the most impressive results of comparisons were the following. There were patients who developed delirium during their ICU stay and these patients were compared with those who did not develop delirium at all over the whole period of observation. The delirium-having patients portrayed considerably greater levels of illness as measured via a variety of rated tests conventionally used to characterize illness status, by particular scores of 22 (17-28) compared to 14.5 (11-20) in non-delirious patients on the APACHE II test, a clinically momentous condensation that implies the relation between physiological corruption and perviousness to delirium. Equally, SAPS II scores were significantly higher in delirious patients at 45 (37-64) compared to those of 30 (23.5-42), and SOFA scores had even higher rates at 10 (5-13) compared to 4 (2-7), which overall shows that delirium incidence was highly related to the dysfunction of organs and severity of diseases(7).

The pattern of admissions provided an interesting account of the correlation between clinical situation and the risk of developing delirium, with numbers quite high (29 percent) during the emergency admissions compared with a minimal percentage (6 percent) of non-delirious patients, but just the reverse situation holding true during planned surgical admissions (31 percent versus 55 percent respectively). This trend implies that emergency admissions due to the acute and unplanned disposition, which is related to the hemodynamic crisis, emergency procedures, and inefficient organization, peculiar to emergency admissions, contributes to the development of delirium in a specific way. Diagnostic categories also confirmed such observations, where surgical patients make up to only 9 percent of incidences of delirium, compared to 45 percent of non-delirious patients, whereas acute respiratory failure patients portrayed slightly more incidences of delirium, which reflects reflections of complex interlogistic interactions between base pathophysiology, treatment intensive and neuro-complications in critically ill cohorts.

3.2 Total Delirium Occurrence and Entire Tendencies

The total rate of delirium of 23.2% in the study of this international cohort is quite consistent with modern literature, which revealed delirium rates of 20-30 percent in the mixed medical-surgical intensive care unit (ICU) populations, which allows believing that the study population is just representative and the diagnostic measures taken are quite reliable. This rate of an incidence indicates the use of current ICU interventions such as an early mobilization procedure, minimization of sedation approaches, and structured delirium prevention practice that has led to lower percentages of delirium compared to past accounts of 40-80 percentages in comparable groups. A curious pattern existed in terms of temporal distribution of cases of delirium with the highest incidence reporting on the 4-6 days of ICU stay (28.3% and 27.9% respectively) implying that the probability of delirium increases with more exposure to the critical illness coupled with the aggregate impact of the stressors associated with the ICU environment.

Prevalence data collected on a daily basis provided useful findings about the dynamic nature of delirium in the ICU environment as there was a significant variation in prevalence rates and patient factors on a day-to-day basis that could affect screening tool performance. The percentage of deeply sedated patients (RASS $\leq$ -3) was also declining during the duration of the study (33.8 percent on day 1 to 18.8 percent on day 7), which anatomized to the changes in the sedation pattern and patient recovery patterns to reflect natural variations in the assessment circumstances. The ratio of patients who needed invasive mechanical ventilation, on the other hand, did not change significantly during the period of observation, 26.5-39.5-53.1% on the first, fifth, and seventh days, respectively, of admission, and this means that the need to use invasive mechanical ventilation was often noted to be the same or increase when the patient remained in the ICU during a long stay.

The distribution of missing data, as well as the failure to conduct any assessment on the ground of deep sedation (RASS -4 to -5) was very different across screening tools, with CAM-ICU having the highest number of non-assessable patients (70 patients) than ICDSC (48 patients) or Nu-DESC (47 patients). Such a differential pattern could probably depict the interactive character of CAM-ICU measurement, where the collaboration of the patient is needed and verbal input occurs, which cannot be done in gravely sedated patients much the same with observational measure tests such as ICDSC and Nu-DESC can frequently be filled out despite the patient remaining incapable of actively taking part in the appraisal process.

3.3 Comparative performance analysis tools of screening instruments

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

The global analysis of screening tool diagnostic efficiency made it clear that different patterns were present reflecting significant implications on clinical practice and application of screening tools in various ICU settings. The Nu-DESC- showed the best balance profile in terms of sensitivity and specificity that came to be 88.5 % and this meant the same ability to detect delirium among the affected and non-deliriums among the non-damaged. This balanced performance feature makes Nu-DESC especially appealing to the generalized purposes of ICU screening in which false-positive and false-negative outcomes are clinically significant. The predictive value of 71.9% allows regarding about 7 out of 10 positive Nu-DESC outcomes as the positive ones that are the cases of true delirium, and the negative predictive value of 95.8% is strong evidence that the negative result will rule out the presence of delirium, which makes this screening test effective in ruling out delirium in clinical practice(8).

ICDSC had a more conservative performance trend with a much lower sensitivity of 62.5 percent and more specificity of 92.4 percent, which means that it is likely to have fewer false-positive outcomes but might overlook several cases of delirium. Such performance profile would be desirable in cases where false-positives are of particular concern (e.g., positive screening result that causes automatic consultations, alteration of medications, or other resource-consuming interventions). Positive predictive value of 71.4 percent was close to Nu-DESC, and the negative predictive was 89.0 percent lower owing to effect of diminished sensitivity on the tool rule out delirium with conviction.

The CAM-ICU had demonstrated moderate sensitivity of 75.0% and outstandingly high specificity of 94.7%, but most importantly the highest positive predictive value of 85.7, which means positive results of CAM-ICU should be greatly trusted as in fact delirium presents. This aspect renders CAM-ICU especially useful in situations where there is strong reason to believe that the tests will convey positive findings, which should be used in clinical decision-making. The negative predictive value of 90.0 percent was just in between the other instruments demarcating the poise between sensitivity and specificity attained by this instrument.

The statistical analysis of the presented results given by the comparison of statistics conducted by McNemar test revealed that the overall performance of each program, Nu-DESC and ICDSC gave significant results as compared to both CAM-ICU ($p = 0.004$) and ICDSC ($p = 0.008$), whereas the significant difference was not found between CAM-ICU and ICDSC ($p = 0.739$). Based on these data, the use of Nu-DESC has quite divergent performance features to all other instruments, mainly because it was more sensitive, whereas the ICDSC and the CAM-ICU revealed less different performance profiles in terms of overall diagnostic accuracy, although with dissimilar sensitivity-specificity trade-offs.

3.4 Optimal Cut-point determination and ROC analysis

The receiver operating characteristic curve analysis helped to draw vital speakers on the optimal diagnostic levels of each screening instrument and demonstrated universally high discriminative power in each screening instrument. The area under the curve (AUC) value (termed the C-statistic in the diabetes study) was exactly 0.93 in all three screening instruments regarding overall capacity to discriminate between delirious and non-delirious patients when not viewed as dichotomous diagnostic items, but instead taken as continuous or ordinal scales. This observation has indicated that the diagnostic information content of each of the three tools is similar when range of their total scoring is used such that variance between traditional measures of performance is largely attributable to the effect of cut-point choices of the two tools rather than to differing diagnostic capacities.

The Youden index analysis that adjusts sensitivity and specificity balance unveiled that baseline cut-points are not necessarily the best diagnostic thresholds of such populations in the ICU. In the case of Nu-DESC, optimum cut-point was observed to be of 1 point and not the conventional 2 points with sensitivity of 96.2 per cent and 83.3 per cent specificity. The resulting lower threshold would have the effect of being more sensitive in detecting delirium but with a bit less specificity, producing an unknown but possibly higher number of false positives and a possibly lower number of false negative detections of delirium as compared to a higher threshold. In a similar manner, ICDSC did best at a cut-point of 2 points compared to conventional 4 points, and reported sensitivity of 87.5% and specificity of 83.5%, implying that lower cut-points potentially enhance overall diagnostic precision in ICU cohorts as well.

In the case of severity scale CAM-ICU (ssCAM-ICU), best performance was 5.5 points cut-point resulting in sensitivity of 91.7 percent and specificity of 82.5 percent. It concludes that CAM-ICU is indeed useful when administered in a severity scale form (as opposed to the usual binary flowchart) with the potential of better diagnostic results and ability to capture delirium severity trends. These identified optimal cut-points are also of clinical importance, and the findings imply that institutions can improve their diagnostic performance by using lower thresholds than have conventionally been advised, but these changes would need to be verified in other

populations and would need to reflect the local assiduity of staff and the populations of patients to which they are applied and their clinical workflow.

3.5 Influence of Sedation level on diagnostic Performance

The discussion of the influence of the level of sedation on the accuracy of screening tools also had one of the most clinically valid implications of the current study and the distinctive characteristics of screening delirium within the modern ICUs where the concept of sedation control is an essential part of patient care. The patients were systematically divided in three groups of sedation score according to Richmond Agitation Sedation Scale (RASS), where most of the patients (70-71%) belonged to the awake to drowsy category (RASS 0/-1), which represents the contemporary trends aimed at practicing lighter planes of sedation and daily regimens of sedation interruption. The highly sedated cohort (RASS < -1) was around 23 percent of all assessments and the proportion with agitated patients (RASS > 0) was also low as expected in modern sedation and analgesia practice in the ICU(9).

Performance of observational screening tools revealed dramatic, statistically significant, loss in performance in deeply sedated patients with especially notable effects on specificity with an important clinical relevance. Put differently, Nu-DESC proved a sensitivity of 85.7 in deeply sedated patients but a 0% specificity, showing that all non-delirious patients in this sedation level were diagnosed as delirious by the tool. This observation indicates that apparent changes in consciousness, responsiveness and behaviour due to sedation are confused as delirium manifestations by observational screening methodologies giving rise to the systematic false positive outcome in heavily sedated groups.

Likewise, ICDSC made 78.6 percent sensitivity, and extremely dull 25 percent specificity in heavily sedated patients, whereas CAM-ICU indicated most durable performance with 86.7 percent sensitivity and specificity of 50. CAM-ICU did better with sedated patients which might be due to its structured and algorithmic design, which considers the particular levels of altered consciousness explicitly and needs specific constellation of symptoms and not cumulative scoring of individual observations which can become complicated with the effects of sedation. The statistical proof of differences in area under the curve values at sedation subgroups was conducted for Nu-DESC ($p = 0.049$) and ICDSC ($p = 0.021$) and demonstrated no significant variation in ssCAM-ICU ($p = 0.225$) that proves the sedation level to be a highly relevant factor affecting the diagnostic validity of the observational screening tools but not the interactive assessment tool like CAM-ICU.

3.6 Clinical context and effects of Mechanical Ventilation

As opposed to initial hypotheses and to clinical expectations, the overall evaluation of the effect of mechanical ventilation on the screening tool performance did not show a statistically significant difference in the diagnostic performance of invasively ventilated patients and non-invasively ventilated patients with regard to the three screening tools tested. This observation disapproves conventional beliefs regarding the challenges of communication hindrances and evaluation challenges in mechanically ventilated patients, indicating that skilled health practitioners can competently adjust screening methodologies to overcome the challenges raised by endotracheal intubation, tracheostomy, or any other type of invasive airway management.

The patient distribution by venture classes indicated that 79-82 percent patients had non-invasive or spontaneous breathing ventilation procedures, and 18-21 percent patients had invasive mechanical ventilation, which in representation of modern ICU management aims at applying non-invasive methods of ventilation by means of ventilators where conditions call for them and the early weaning of patients who are on mechanical ventilation in order to decrease the complications associated with the latter practice. The performance of Nu-DESC in invasively ventilated patients was sensitive as 90.0 percent with a specificity of 66.7 percent whereas in non invasively ventilated patients there was sensitivity of 87.5 and specificity of 87.7 whereby the confidence interval tends to overlap significantly and never showed any significance.

Results obtained by ICDSC followed the same routes with invasively ventilated patients exhibiting sensitivity of 70.0% and specificity of 71.4% against the results of non-invasively ventilated patients who were 64.3 and 94.9%, respectively. Interestingly, CAM-ICU demonstrated a non-significant trend of better performance among invasively ventilated patients with sensitivity 90.9 percent and specificity of 71.4 percent compared to 68.8 percent and 96.2 percent in the invasively ventilated patients. This unexpected trend can be an indication of the logicity of the process of CAM-ICU assessment which can be much less prone to the obstructions of communication on the condition of being appropriately modified to fit the needs of a non-verbal patient.

The results of the area under the curve analysis supported the fact that there were no significant differences based on performance that could be attributed to ventilation, where Nu-DESC indicated AUC values of 0.72 (95% CI,

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

0.58 to 0.86) versus 0.93 (95% CI, 0.86 to 1.00) ($p = 0.220$) in invasively versus non-invasively ventilated patients, ICDSC reported 0.71 (95% CI, The outcomes are encouraging to say that delirium screening is reliable among various types of ventilations when proper assessment methods are used and that the members of staff get sufficient knowledge to phase out screening processes in accordance with patients being connected to a mechanical ventilator.

4.Impact of Sedation Levels on Diagnostic Accuracy

The purpose of the analysis of the effect of the level of sedation on the performance of delirium screening tools showed that the findings are extremely crucial and carry crucial implication on clinical practice and prompt performance of delirium screening tools in ICU environment. The patients were successively divided into three separate sedation subsets according to the Richmond Agitation Sedation Scale (RASS) scores: deeply sedated (RASS < -1), awake to drowsy (RASS 0/-1) and agitated (RASS > 0) to provide a thorough assessment of the screening tools performance across significance of consciousness levels experienced in the community care practice.

The findings indicated that the level of sedation exerted a significant and major effect on the validity of observational screening instruments, especially ICDSC and Nu-DESC, and that the CAM-ICU was more resistant in the degradation of its performance relative to sedation. In patients in the awake to drowsy range (RASS 0/-1) that constituted the largest group in the assessment, Nu-DESC performed best with sensitivity of 81.2 percent and specificity of 91.9 percent whereas the ICDSC produced a sensitivity of 62.5 percent displaying an incredible specificity of 95.4 percent. The CAM-ICU did not perform any differently in this group where their sensitivity of 72.2% and specificity of 100% indicates a certain reliability when positive results are found in patients who are not so heavily sedated.

The deterioration of performance in deeply sedated patients (RASS < -1) was still vivid and clinically relevant with especially pronounced effects on specificity measures. The Nu-DESC demonstrated preserved sensitivity of 85.7 as well as the lack of sensitivity of 0%, meaning that the risk of future-positive outcomes was high with deeply sedated patients. On the same note, the ICDSC was sensitive by 78.6 percent with severely low specificity of 25 percent and the CAM-ICU performed rather well with sensitivity of 86.7 percent and specificity of 50 percent. These results indicate that sedated patients can show clinical signs that are misquoted by observational screening instruments because the alterations in awareness and responsiveness resulting in sedation resemble the symptoms of delirium.

These findings have significant implications off the clinical practice front in that the recommendation not to screen routinely delirium in deeply sedated patients may create many situational situations where false-positive diagnoses are made, and thus, wrong procedures are administered, wrong medication decisions made, and miscalculations in delirium occurrence rates. The hypothesis of a lack of pathophysiological change and clinical consequences that may result in the apparent development of delirium in patients who are sedated has been termed as a sedation-induced and rapid-reversible delirium. This difference is quite important since the interventions on suspected delirium, including the use of antipsychotic drugs or sedation titration, are often ineffective or can go detrimental to the patients that merely feel what they are expected to due to any mandatory sedation.

Status of Ventilation and Clinical Significance

The univocal investigation of mechanical ventilation effects on the delirium screening tool performance presented subtle results that create confusion about some assumptions that are normally taken as truth in relation to the delamination of the ventilatory support and the accuracy of the delirium assessment. As opposed to the faith that mechanical ventilation would induce substantial failure of delirium screening tool in connection with a dearth of effective communication and uncomfortable position of the patient, the statistical classification displayed that there were no distinguishing differences in diagnostic validity between the group of invasively ventilated and non-invasively ventilated patients when considering all three screening devices used.

The Nu-DESC had the sensitivity of 90.0 and specificity of 66.7 among patients under invasive mechanical ventilation and 87.5 sensitivity and 87.7 specificity among patients without any invasive ventilatory support. The ICDSC had the same performance in the ventilation groups, with invasively ventilated still having a sensitivity of 70.0 percent and specificity of 71.4 percent in the ventilation responses in I compared to those non invasively ventilated with sensitivity of 64.3 percent and specificity of 94.9 percent. The CAM-ICU in fact depicted a more positive trend by indicating higher performance in invasively ventilated patients with sensitivity of 90.9% and

specificity of 71.4% compared to the non-invasively ventilated patients at sensitivity of 68.8 and specificity of 96.2.

These results indicate that even professionals that work in the field of healthcare where delirium screening behavior can be adapted to conditions related to mechanical ventilation through applying non-verbal communication conventions, visual signals, and altered assessment protocols that do not affect the validity of applied screening tools. Delirium is of special concern in patients undergoing mechanic ventilation because they are likely to spend an extended period in ICU with an increased susceptibility to their development.

The clinical meaning of these results is not limited by a mere accuracy of the diagnostic but it also comprises the overall quality of care in an ICU and the outcomes of patients. Their demonstration of the reliability of the assessment of delirium screening with the use of different modality ventilators encourages the development of universal screening strategies that do not need to be adjusted by any characteristic of the ventilator. This uniformity will play an important role in establishing standardized care pathways and providing delirium assessment to all the patients in ICU without looking at their support needs in terms of respiratory assistance.

Nevertheless, the study also indicated the significance of proper training and experience of staff conducting delirium assessments, since the capacity to preserve diagnostic precision in difficult clinical situations is probably determined by the view of the provider to the tools of delirium assessment and the likelihood of their ability to apply assessment procedures to unique instances of patients. Additionally, this intervention needs to be expanded to include specific training modules and competency assessment to screen delirium in patients under mechanical ventilation in order to further increase the reliability and consistency of delirium screening in various providers and clinical contexts.

5. Conclusion and Future work

5.1 General Statement of the Major Results and Clinical Implications

This global study has been invaluable, especially in the sense that the findings are evidence-based on the diagnostic accuracy of delirium screening tools among severely ill patients and introduced some critical factors that play an imperative role in determining the diagnostic accuracy and clinical practicality in the modern intensive care setting. The most relevant conclusion of the study regards the exceptional influence of the degree of sedation on the diagnostic accuracy of the observational screening instruments, as deeply sedated subjects (RASS < -1) produced an alarmingly high drop in specificity of Nu-DESC and ICDSC instruments, whereas the CAM-ICU could perform significantly better with the same extent of sedation. All these emergency findings have direct implications of immediate as well as long term support in the clinical practice since they advocate that regular use of observational screening instrument in deeply sedated patients might culminate into systematic overdiagnosis of delirium, hence raise inappropriate intercessions, incorrect tampering with medications, and inaccurate reports concerning rates of delirium prevalence.

The clinical implication of these findings are not merely those of diagnostic accuracy but rather of the more basic issue of exactly what we are trying to measure with delirium, and whether such a measurement is possible at all in the current intensive care unit setting in which sedation is an essential and often life-saving therapy that cannot be ignored in the face of accurate neurological evaluation. The term of such importance arises as a possible solution to the question that why is there a term of sedation-induced, rapid-reversible delirium? Why does it seem important that we should distinguish between pharmacologically-imposed changes in consciousness that resemble some symptoms of a delirium, and the genuine delirium pathophysiology that demand particular responses and have prognostic implications regarding long-term thinking in academic subjects. Such a difference proves especially significant when it is remembered that the measures taken to treat suspected delirium, including antipsychotic medications or a forceful attempt at sedation reduction, can be inappropriate and even dangerous in patients undergoing anticipated sedative effects instead of actual delirium.

The fact that all three screening instruments gave similar results evaluated by receiver operating characteristic analysis (AUC = 0.93) indicates that each of the tools possess information content that is significant as a diagnostic tool when used correctly, and variation in other traditional measures of performance is mostly due to variation in sensitivity-specificity trade-offs rather than to intrinsic differences in diagnostic power. Nonetheless, the possibility of identifying optimal cut-points not associated with the conventional thresholds implies the effectiveness of increasing the diagnostic accuracy via the evidence-based revision of the screening protocols. Clinical implications It should be noted that lower cut-points (Nu-DESC 1 Hg 1, ICDSC 2 Hg 2) demonstrated

Impact of Consciousness Depth and Respiratory Support on the Accuracy of Delirium Detection Tools in Intensive Care

the best possible diagnostic performance in ICU populations and should be considered carefully when implementing in clinical practice, but this would have to be tested in independent groups and whether this would have any significant effect on false-negative and false-positive rates and in patient care pathways.

5.2 Clinical Practice and Quality-Improvement Implications

To apply the research findings to practical clinical practice inefficient regards of the ways to maximize feasibility of delirium screening procedures without the loss of practical feasibility and risk of the undesirable consequences of inaccurate screening diagnosis should be expected. The first conclusion that can be drawn on the basis of this research is that measures are needed to introduce sedation-adjusted interpretation rules that take into consideration the large magnitude of the effect of the level of consciousness on the performance of screening tools. The clinical teams must establish systematic methods, prioritised methods of interpreting the outcome of scans in the context of sedation status, especially in the evaluation of positive results when this could be misleading as false-positive rates could be as high as 75-100 percent in observational tool.

Effective implementation approaches could entail the future creation of clinical decision support aids that will automatically shift screening interpretation to correspond to concurrent RASS ratings, in which positive findings among deeply sedated patients will elicit meticulous clinical consideration instead of rule-based poor delirium designations. These strategies may include the need to reducing sedation where clinically possible and reassessment once there has been some improvement in the level of consciousness or receive input through consultations with the specialists in delirium diagnosis and treatment. One should aim at reducing the false-positive rating of diagnosis, and not to overlook the real cases of delirium because of the challenges related to the assessment caused by sedation.

The better results of CAM-ICU in various depths of sedation indicate that, possibly, this instrument can be especially appropriate in ICU, where the degree of sedation is highly varied or frequently altered during patient treatment periods. Yet, these performance benefits must be offset by the increased training investments and requirements needed to support CAM-ICU implementation, and the institutions must choose screening instruments based on their staffing patterns, patient populations, and quality improvement initiatives. The establishment of hybrid methods combining various screening tools depending on the patient profiles or clinical situations could be an ideal way of ensuring high diagnostic accuracy and still ensuring this is clinically realistic.

The improvement of quality is supposed to be addressed by means of improving the training and competency of the staff in the field of delirium assessment and, especially, training on adapting the screening methods to various clinical cases. The observation that mechanical ventilation had no significant effect on the performance of the screening tools is an encouraging indication that the barrier to communication can become the subject of a successful overcoming with the modification of the training and evaluation, which serves as an assuring evidence that the universal screening protocols can be applied to all demographics of the ICU. To guarantee continuity of assessment quality and uniform method of screening protocol use across various providers and clinical situations, regular competency evaluation, continuous educative programs and feedback systems should be adopted.

Acknowledgement: Nil

Conflicts of interest

The authors have no conflicts of interest to declare

References

1. Ely EW, Inouye SK, Bernard GR. Delirium in mechanically ventilated patients: validity and reliability of the CAM-ICU. *JAMA*. 2001;286(21):2703–2710.
2. Bergeron N, Dubois MJ, Dumont M. Intensive Care Delirium Screening Checklist (ICDSC): a new screening tool. *Crit Care Med*. 2001;29(3):521–526.
3. Devlin JW, Skrobik Y, Gélinas C. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium in critically ill adults. *Crit Care Med*. 2018;46(9):e825–e873.
4. Peterson JF, Pun BT, Dittus RS. Delirium and sedation in the intensive care unit: the effect of sedation level on the diagnosis of delirium. *Crit Care Med*. 2006;34(1):137–141.
5. van Eijk MM, van Marum RJ, Klijn IA. Comparison of delirium assessment tools in critically ill patients. *Intensive Care Med*. 2009;35(6):1026–1030.

6. Riker RR, Shehabi Y, Bokesch PM. Dexmedetomidine vs midazolam for sedation and impact on delirium. JAMA. 2009;301(5):489–499.
7. Pandharipande P, Cotton BA, Shintani A. Prevalence and risk factors for development of delirium in trauma ICU patients. Crit Care Med. 2008;36(6):1723–1729.
8. Haenggi M, Blum S, Brechbuehl R. Effect of sedation level on the assessment of delirium using CAM-ICU and ICDSC. Intensive Care Med. 2013;39(12):2171–2179.
9. Zhang Z, Pan L, Deng H. Influence of mechanical ventilation modes on delirium diagnosis accuracy. J Crit Care. 2015;30(2):517–522.