Arabic intensive care delirium screening checklist's validity and reliability: A multicenter study

Nada S. Al-Qadheeb, PharmD, FCCM⁎,1, Lama H. Nazer, PharmD, FCCM b, Tharwat M. Aisa, MD c, Hassan O. Osman, MD a, Asia S. Rugaan, MD d, Ahmad S. Alzahrani, MD d, Iyad M. Ghonimat, RN e, Alaaldin M. Mohammed, MD a, Khalid Maghrabi, MD, FCCP f, Abdelullah A. Alrowaished, MD g, Najah H. Hussein, RN e, Yahya A. Maslamani, MD i, Sawsan Falatah, RN h, Yoanna Skrobik, MD, FRCP(c) MSc, FCCMi

⁎ Corresponding author at: Department of Critical Care, Hafer Albatin Central Hospital, Hafer Albatin, Saudi Arabia
1 Permanent address: King Fahad Specialist Hospital, P.O.Box 15215, Dammam 31444, KSA Phone: +966-505293415, Fax: +966138426251, nada.qadheeb@kfsh.med.sa

Purpose: To develop an Arabic version of Intensive Care Delirium Screening Checklist (ICDSC) and assess its validity and reliability among critically ill patients.

Materials and methods: Multicentered study of convenience sample of adult ICU patients. Arabic translation was performed with rigorous back-to-back translation methods. Concurrent validity was established by calculating the sensitivity and specificity of two examiner assessments compared to a psychiatric evaluation. Kappa coefficients describe interrater reliability, whereas Cronbach α and composite reliability depict internal consistency.

Results: Three hundred critically ill patients were enrolled. Of these, validity testing was assessed in 180 patients. ICDSC screening was positive for delirium in 11% of enrolled patients. The area under the receiver operator characteristic (ROC) curve is 0.9413, with predicted sensitivity 70% (95% confidence interval [CI]: 60–81%) and specificity 99% (95% CI: 98–100%). The Arabic ICDSC showed acceptable internal consistency (Cronbach α = 0.63 and composite reliability = 0.64). Interrater agreement was excellent (Kappa coefficient [κ] = 0.85).

Conclusions: Arabic ICDSC is a valid and reliable delirium-screening tool among Arabic-speaking ICU population. Future studies could address whether these findings are generalizable to a higher proportion of mechanically ventilated patients, and address acceptability and reliability in other Arabic language critical care settings.

© 2019 Elsevier Inc. All rights reserved.
monitor for delirium and 58% do not use a validated screening tool [19]. The most frequent barriers to delirium screening include lack of medical support, complexity of screening tools and difficulty in evaluating intubated and sedated patients [20]. Nonetheless, the exact reasons for failing to perform routine delirium monitoring remain unclear [21].

Among Arabic speaking population, a recently published psychometric study in a tertiary care hospital in Saudi Arabia described a delirium incidence of 63% [22]. Such findings suggest that delirium is similarly common in Arabic-speaking ICU patients and emphasizes the need for delirium assessment tools in Arabic. While CAM-ICU has been translated into several languages including Arabic, German, Greek, Swedish and many more [22-33], the ICDSC translation has been validated in seven languages including Turkish, Japanese, Portuguese, Italian, Dutch, Swedish and Danish [34-40]. The validity and reliability of the Arabic version of ICDSC has not been investigated.

This study aimed to develop an Arabic version of ICDSC tool and to test its validity and reliability in two Middle-Eastern Arabic-speaking countries.

2. Materials and methods

2.1. Setting and study design

This prospective, multicenter cross-sectional study was conducted in four intensive care units in three Saudi Arabian and one Jordanian hospitals. Site A is a 200-bed secondary hospital with 19 medical and surgical ICU (MSICU) beds. Site B is a 1331-bed tertiary hospital with 53 MSICU beds. Site C is a 1500-bed tertiary hospital with 28 MSICU beds. Site D is a 170-bed tertiary hospital with 18 MSICU beds. All four ICUs are closed units integrating a well-established pain and sedation practice. Institutional Review Board approval was obtained for each participating center.

2.2. Participants

Consecutive patients admitted to any of the four medical and surgical ICUs from April 2018 to October 2018 were screened for enrollment. Adult patients 18 years or older were eligible if they were hospitalized in the ICU for at least 24 h and were deemed conscious by Sedation Agitation Scale [SAS] ≥ 3 or a Richmond Agitation-Sedation scale [RASS] ≥ −2. Patients were excluded if they were comatose, non-Arabic speakers, had a condition that might preclude delirium evaluation [severe dementia, schizophrenia, Alzheimer’s disease, severe encephalopathy] or are unwilling to participate (Supplementary material 1, Table 1). Informed consent was obtained from patients or their legally authorized representatives prior to study enrollment. To ensure feasibility of recruitment at each site, recruitment was tracked through a site-specific consort diagram. An 85% screening and attempted recruitment threshold of eligible patients was considered acceptable for study feasibility at each site.

2.3. Translation and cross-cultural adaptation process

The tool was originally developed and validated for delirium screening in mechanically and non-mechanically ventilated ICU patients and designed to evaluate level of consciousness, inattention, disorientation, hallucinations, psychomotor activity, speech or mood disturbance, sleep disturbance and fluctuation of symptoms. A cut-off score of ≥ 3 present items corroborated with a psychiatrist’s clinical diagnosis of delirium [13].

Arabic translation of the original ICDSC tool was performed with a rigorous back-to-back translation method according to the recommendations of the International Society for Pharmacoeconomics and Outcomes Research task force for translation and cultural adaptation [41].

Two authors (TA and HO) independently translated the ICDSC scale from English to Arabic. The translated version was then given to a bilingual ICU nurse blinded to the original English version of the ICDSC, to perform the back translation. To examine unresolved translation issues, these two authors identified and resolved the inadequate expressions/concepts of the translation, as well as any discrepancies between the forward translation and the existing. Experts then reviewed the back-translation with an emphasis on conceptual and cultural equivalence rather than linguistic equivalence.

The translated scale was reviewed by 15 Arabic native speaker ICU nurses to ensure understandability and cultural relevance of the translation. Proofreading and final adjustments completed the final version of Arabic ICDSC’s production.

2.4. Delirium assessment

Trained ICU nurses, physicians, and critical care pharmacists screened patients for delirium using the Arabic version of ICDSC (Supplementary material 2). A multi-faceted approach was utilized to train clinicians. Each investigator was instructed and trained on how to use the Arabic version of ICDSC through one-on-one teaching provided by the primary investigator at each site and through online materials to educate about delirium. An independent delirium expert was available for questions. Assessments at each of the four participating sites were carried out by two trained examiners independently. The paired ratings required that the two assessments are performed within one hour. A psychiatrist validated the presence/absence of delirium within two

---

Table 1

Baseline characteristics.

<table>
<thead>
<tr>
<th></th>
<th>No. of patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site A</td>
<td>60 (20)</td>
</tr>
<tr>
<td>Site B</td>
<td>90 (30)</td>
</tr>
<tr>
<td>Site C</td>
<td>90 (30)</td>
</tr>
<tr>
<td>Site D</td>
<td>60 (20)</td>
</tr>
<tr>
<td>Age, yrs (mean ± SD)</td>
<td>52.2 ± 16.7</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>148 (49.3)</td>
</tr>
<tr>
<td>Mechanical ventilation, n (%)</td>
<td>22 (7.3)</td>
</tr>
<tr>
<td>Days in ICU before enrollment (median)</td>
<td>2 (1–8)</td>
</tr>
<tr>
<td>Sedation Agitation Scale at the time of assessment RASSa (median)</td>
<td>0 (−1–0)</td>
</tr>
<tr>
<td>Riker-SASb (median)</td>
<td>4 (3–4)</td>
</tr>
<tr>
<td>Location before hospitalization, n (%)</td>
<td></td>
</tr>
<tr>
<td>Home with spouse</td>
<td>77 (25.7)</td>
</tr>
<tr>
<td>Home with other family members</td>
<td>217 (72.3)</td>
</tr>
<tr>
<td>Other</td>
<td>6 (2.0)</td>
</tr>
<tr>
<td>Location before ICU admission, n (%)</td>
<td></td>
</tr>
<tr>
<td>Emergency department</td>
<td>135 (45)</td>
</tr>
<tr>
<td>Hospital ward</td>
<td>97 (32.3)</td>
</tr>
<tr>
<td>OR following elective surgery</td>
<td>42 (14)</td>
</tr>
<tr>
<td>OR following emergency surgery</td>
<td>6 (2.0)</td>
</tr>
<tr>
<td>ICU at outside hospital</td>
<td>135 (45)</td>
</tr>
<tr>
<td>Ward at outside hospital</td>
<td>3 (1.0)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (0.3)</td>
</tr>
</tbody>
</table>

ICU = intensive care unit; OR = operating room; RASS = Richmond Agitation Sedation Scale; SAS = Sedation Agitation Scale.

---

n = 240.  
b = 60.

hours of the two examiners performing independent Arabic ICDS evaluations in two Saudi hospitals based on DSM-IV (Diagnostic and Statistical Manual of mental disorders, fourth edition) [42] criteria. The raters’ examinations were administered with each result blinded from the other and the obtained results, in turn, were blinded to psychiatric assessment. All assessment records were kept separately to further ensure blinding.

2.5. Statistical analysis

All statistical analyses were performed using SPSS version 19 (SPSS Inc., Chicago, IL., USA) and Statistical Analysis Solutions (SAS) version 9.4 (SAS Institute Inc., Cary, NC). ICDS validity was measured by calculating the sensitivity, specificity, positive and negative predictive values of examiners in two Saudi Arabian tertiary hospitals in comparison with the psychiatrist clinical evaluation using the DSM-IV-TR. Interrater reliability was evaluated as the agreement between raters with Kappa coefficient. Moreover, internal consistency was described using the Cronbach’s statistic and composite reliability. Based on sensitivity and specificity of 90%, an assumed delirium prevalence of 21%, precision of 8% and a 95% confidence interval (CI), the minimum sample size of this study was calculated to be 262 patients.

3. Results

A total of 651 patients were evaluated and 300 patients were enrolled in the study (Fig. 1). Descriptive statistics of study participants are displayed in Table 1. The mean age was 52 ± 17 years; 49% were males. Reasons for admission to the ICU were primarily post-operative cases, respiratory disease and sepsis.

3.1. Validity

Validation of the Arabic version of ICDS was performed at two Saudi Arabian tertiary hospitals. Two independent examiners and a psychiatrist evaluated a total of 180 patients; 37 (21%) patients were diagnosed with delirium based on DSM-IV criteria and 30 (17%) patients screened positive for delirium based on the Arabic ICDS assessment (Table 2). The remaining two sites were not able to perform validation as no psychiatrist was available. Overall, 32 (11%) patients from the four sites screened positive for delirium based on the Arabic ICDS assessment.

The ROC curve with its area under the curve of 0.9413 is shown in Fig. 2. In keeping with the original scale validation, an overall sensitivity of 70% (95% CI: 60–81%) and a specificity of 99% (95% CI: 98–100%) was demonstrated at ICDS cutoff score of ≥4. Furthermore, subgroup analysis of the examiners assessments at the two sites using the Arabic ICDS showed sensitivity ranging from 45% to 100% and specificity ranging from 97% to 100% (Supplementary material 1, Table 2).

3.2. Reliability

We assessed the level of agreement between assessments with the Arabic ICDS by calculating interrater reliability (Kappa coefficient [κ]). A total of 300 paired assessments were performed at the four participating sites. Significant and substantial agreement between raters was determined with an overall κ of 0.85. The individual ICDS item was also completed demonstrating good to excellent concordances (Table 3). Moreover, interrater reliability of ICDS scores between examiners at the four participating centers was strong with Kappa coefficient ranging from 0.66 to 1.0 (Supplementary material 1, Table 3). Finally, the Arabic ICDS demonstrated acceptable internal consistency with Cronbach α = 0.63 and composite reliability of 0.64. However, Cronbach α improved to 0.70, if item 7 (sleep/wake cycle disturbance) was removed. Item 2 (inattention) contributed the most to the reliability of the Arabic ICDS.

4. Discussion

In this multicenter, cross-sectional study, we successfully developed an Arabic version of ICDS and demonstrated an acceptable validity and
Table 3: Inter-rater reliability of the Arabic intensive care delirium screening checklist.

<table>
<thead>
<tr>
<th>ICDSC domain</th>
<th>k</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered level of consciousness</td>
<td>0.566</td>
<td>0.125–1.000</td>
</tr>
<tr>
<td>Inattention</td>
<td>0.834</td>
<td>0.735–0.933</td>
</tr>
<tr>
<td>Disorientation</td>
<td>0.823</td>
<td>0.725–0.921</td>
</tr>
<tr>
<td>Hallucination/delusion</td>
<td>0.805</td>
<td>0.694–0.916</td>
</tr>
<tr>
<td>Psychomotor agitation</td>
<td>0.869</td>
<td>0.768–0.970</td>
</tr>
<tr>
<td>Inappropriate speech or mood</td>
<td>0.819</td>
<td>0.698–0.940</td>
</tr>
<tr>
<td>Sleep/wake cycle disturbance</td>
<td>0.755</td>
<td>0.679–0.831</td>
</tr>
<tr>
<td>Symptom fluctuation</td>
<td>0.745</td>
<td>0.636–0.854</td>
</tr>
<tr>
<td>Overall (n = 300)</td>
<td>0.852</td>
<td>0.753–0.951</td>
</tr>
</tbody>
</table>

ICDSC = intensive care delirium screening checklist, k = kappa coefficient.
CI = confidence interval.

reliability in detecting delirium among critically ill Arabic speaking population.

ICU physicians, nurses and pharmacists assessed patients for delirium using the Arabic ICDSC. Of these assessments, 50% were compared with psychiatrist clinical evaluation using DSM-IV-TR to determine the concurrent validity of the tool. Previous studies have reported that delirium occurs in up to 80% of patients in acute care setting and 10% to 80% in critically ill patients [43-46]. In this study, the psychiatrist clinical evaluation identified delirium in 21% of critically ill patients. This suggests that delirium is similarly common among critically ill Arabic-speaking population and highlights the continuous need for a reliable and valid tool to assess delirium in this population.

The overall sensitivity and specificity of the Arabic ICDSC, 70% and 99%, respectively is consistent with previous studies suggesting sensitivity and specificity of the ICDSC ranging from 64 to 99% and 51–88%, respectively (Supplementary material 1, Table 4) [13,35,37,38]. Our results show inter-rater variability. High specificity values (97%, 98%, 100% and 100%) were achieved from all four examiners, while remarkable difference in sensitivity were noted, 45%, 59%, 93% and 100%, respectively. These findings are similar to those by Van Eijk et al. where ICU physicians showed higher specificity (96%) than sensitivity (29%) [38]. Such results might be due to lack of experience in detecting delirium which reinforce the need for a more effective structured training program and extensive training methods to demonstrate higher agreement with expert raters; such programs improve ICU physicians’ and nurse’s ability to detect delirium with the ICDSC [47].

The interrater reliability of Arabic ICDSC between examiners in all four ICU settings was good to excellent for each individual component of the tool, as well as the overall score x 0.85. These results are consistent with earlier studies, where the ICDSC also demonstrated moderate to high interrater reliability with x of 0.7 [13,34,37]. This might be related to the fact that all examiners had a comprehensive understanding of ICU delirium. As for the item correlation, Cronbach’s alpha of 0.63 is considered an acceptable internal consistency. Each of the eight ICDSC items is highly distinct and specific for the diagnosis of delirium [48]. Subgroup analysis for examiner’s assessment from all participating centers showed excellent to excellent interrater reliability with kappa coefficient ranging from 0.66 to 1.0. However, the correlation for four ICDSC items in one of the sites, namely inattention, disorientation, psychomotor agitation and sleep/wake cycle disturbance was noted to be slightly lower. This variation may be related to differences in individual clinicians’ level of expertise.

Our study has several strengths. It is the first to assess the validity and reliability of the Arabic ICDSC as a delirium assessment tool for Arabic critically ill patients, and was performed as a multicenter evaluation in two Middle Eastern countries. The original ICDSC Arabic translation considered the local cultural and linguistic context. Cross-cultural adaptation is an essential step in ensuring the equivalence in the meaning of the instrument at a conceptual level across cultures [49]. The study evaluated a mixed population of critically ill patients, supporting the generalizability and the external validity of our findings. In addition, delirium assessments were carried out within two hours to reduce the risk of variability in terms of diagnosis due to the fluctuating nature of delirium. Furthermore, subgroup analysis was conducted to measure the psychometric features of ICDSC according to examiners assessment at each site.

The main limitations of our study are that evaluations occurred once during ICU stay; in this case the overall delirium prevalence might have been underestimated. A relatively small number of patients with neurological diagnosis were enrolled limiting the generalizability to this specific patient group. While there may have been differences in the patients’ characteristics between the institutions, we cannot provide a full comparison of some data that were not captured such as the APACHE scores.

Moreover, while ventilated patients were not excluded, few ventilated patients were enrolled, limiting the generalizability of our findings.

5. Conclusions

The Arabic ICDSC is a valid and reliable tool for the detection of delirium among Arabic speaking critically ill patients and should therefore be incorporated in the daily ICU practice in Arabic ICU settings. Future studies could address whether these findings are generalizable to a higher proportion of mechanically ventilated patients while documenting their validity across a broad range of illness severity markers (e.g. APACHE scores), and address acceptability and reliability in other Arabic language critical care settings.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Declaration of Competing Interest

None.

Acknowledgement

The authors would like to thank all the ICU clinicians and staff at study sites for making this study possible. We are grateful to Ms. Khawla Anmar and Ms. Dalia Alrimawi for their help in data analysis and statistics.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jccr.2019.08.025.

References


