

ORIGINAL

Pain assessment tools for non-verbal adult ICU patients: systematic review

Herramientas de evaluación del dolor no verbales para pacientes adultos en UCI: revisión sistemática

Merrou Soumia¹ , Khawla Hamdoun², Nouhad Yaych², Baslam Abdelmounaim³ , Zakariya Ouhaz³ 

1. Health Sciences Research Center, Faculty of Medicine and Pharmacy, Cadi Ayyad University, Marrakech

2. Higher Institute of Nursing and technical health sciences, Marrakech

3. Laboratory of Pharmacology, Neurobiology, Anthropobiology and Environment, Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakech, Morocco

Corresponding author

Baslam Abdelmounaim

E-mail: baslamounaim@gmail.com

Received: 2 - XI - 2023

Accepted: 28 - XI - 2023

doi: 10.3306/AJHS.2024.39.02.144

Abstract

Aim: determine valid and reliable pain assessment tools for ICU patients who are unable to self-report their pain and discomfort. Design: Systematic review based on PRISMA Checklist 2020 "Preferred Reporting Items for Systematic Reviews and Meta-Analyses". This study is still in the registration process in PROSPERO.

Methods: We have collected data from December 20th, 2022 to March 28th, 2023, in the following databases: PubMed, Google Scholar, NIH Library, Cochrane central register, and Medscape. We've collected 789 studies. They were all screened "title and abstract and full-text screening" by the two researchers independently. The total number of included studies is 25.

Results: Twenty-five (n=25) studies were included and seven hundred sixty-four studies (n=764) were excluded. Quality appraisal of included studies was done by the two researchers using the CASP checklist. Eighteen (n=18) studies discussed CPOT, sixteen (n=16) studies discussed about BPS, three (n=3) studies discussed BPS-NI, three (n=3) studies discussed NPAT, six (n=6) studies discussed ONVPS, nine (n=9) studies discussed RNVPS, four (n=4) studies discussed PAINAD and two (n=2) studies discussed DPS. CPOT, BPS, BPS-NI, RNVPS, and PAINAD were all valid and reliable instruments to be used among the non-verbal adult ICU population. The CPOT is to be preferred since it showed nearly perfect properties. ONVPS, DPS, and NPAT showed low evidence of psychometric properties..

Key words: Pain assessment tool, pain scale, non-verbal ICU patients, adult ICU population, valid and reliable pain assessment tools.

Resumen

Objetivo: Determinar herramientas de evaluación del dolor válidas y fiables para pacientes en UCI que no pueden comunicar su dolor e incomodidad. Diseño: Revisión sistemática de la literatura basada en la Lista de verificación PRISMA 2020 "Ítems de Reporte Preferidos para Revisiones Sistemáticas y Metaanálisis". Este estudio se encuentra aún en proceso de registro en PROSPERO.

Métodos: Recopilamos datos desde el 20 de diciembre de 2022 hasta el 28 de marzo de 2023, en las siguientes bases de datos: PubMed, Google Scholar, Biblioteca NIH, Registro central Cochrane y Medscape. Hemos recopilado 789 estudios. Todos fueron evaluados "por título y resumen y revisión de texto completo" por los dos investigadores de manera independiente. El número total de estudios incluidos es 25.

Resultados: Se incluyeron veinticinco (n=25) estudios y se excluyeron setecientos sesenta y cuatro (n=764) estudios. La evaluación de calidad de los estudios incluidos fue realizada por los dos investigadores utilizando la lista de verificación CASP. Dieciocho (n=18) estudios discutieron sobre CPOT, dieciséis (n=16) estudios discutieron sobre BPS, tres (n=3) estudios discutieron sobre BPS-NI, tres (n=3) estudios discutieron sobre NPAT, seis (n=6) estudios discutieron sobre ONVPS, nueve (n=9) estudios discutieron sobre RNVPS, cuatro (n=4) estudios discutieron sobre PAINAD y dos (n=2) estudios discutieron sobre DPS. CPOT, BPS, BPS-NI, RNVPS y PAINAD fueron todos instrumentos válidos y fiables para ser utilizados en la población de adultos no verbales en UCI. El CPOT es preferible ya que mostró propiedades casi perfectas. ONVPS, DPS y NPAT mostraron baja evidencia de propiedades psicométricas.

Palabras clave: Herramienta de evaluación del dolor, escala de dolor, pacientes no verbales en UCI, población de adultos en UCI, herramientas de evaluación del dolor válidas y fiables.

Cite as: Soumia M, Hamdoun K, Yaych N, Abdelmounaim B, Ouhaz Z. Pain assessment tools for non-verbal adult ICU patients: systematic review. *Academic Journal of Health Sciences* 2024; 39 (2):144-155 doi: 10.3306/AJHS.2024.39.02.144

Introduction

Intensive care unit (ICU) services are a source of many aggressions for patients, who must undergo the application of heavy and often invasive support techniques. And because of the technicality of the care and the imperatives of security and surveillance, they create an extremely difficult living environment for them¹. According to some authors, almost 30% of patients experience pain at rest and 50% during various nursing interventions. The majority of patients discharged from an ICU identify the pain experienced as a huge source of stress².

Pain takes on subjective dimensions; in other words, its description is based upon the patient's own description of pain, which is not always possible: the intensive care patients have a diminished ability to communicate, and it causes that the patient may experience pain without being able to reveal it in a way that intensive care personnel understand. Thus, it becomes necessary to take into consideration the case of patients who are unable to formulate and express their experience of discomfort.

A lack of the ability to communicate pain verbally does not eliminate patient's potential for experiencing pain. However, the inability to communicate pain verbally remains a significant ethical and medical challenge, to recognize this pain as well as to choose an appropriate treatment. For this reason, this systematic review conducted conforming to PRISMA 2020 checklist aims to determine valid and reliable pain assessment tools to be used among this ICU population.

Methods

Study design

This study is designed in accordance with systematic review criteria which is conducted conforming to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. The protocol of this systematic review is still in the registration process in PROSPERO (ID: 431148).

Eligibility criteria

Inclusion criteria

- Grey literature and published full text studies in English, Italian and French.
- Studies that address nonverbal pain assessment tools.
- Selection of ICU (every type of intensive care units) adult (aged 18 years or older) population who have impaired ability to communicate their pain experience.

Exclusion criteria

- Transcultural adaptation and summarized articles.
- Repetitive papers

Information sources and search strategy

the study was performed from December 20th 2022 to March 28th 2023, in the following databases: Pub

med, Google scholar, NIH library, Cochrane central register, Medscape, using the following keywords: pain assessment tool, pain scale, non-verbal ICU patients, adult ICU population, valid and reliable pain assessment tools. These keywords addressed the different elements of the search question in order to select all possible studies. Multiple combinations of search keywords were done using Boolean connectors including: AND, OR, such as: pain scale OR pain assessment tool – pain scale AND non-verbal ICU patient. This strategy generated 789 studies.

Selection and data collection process

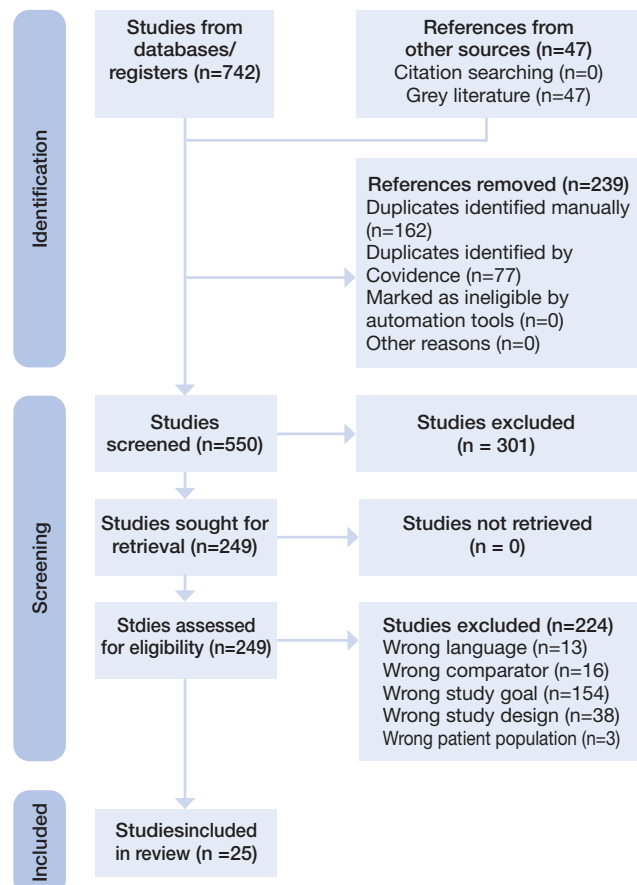
The articles that were found in the search were screened by two independent authors, who assessed whether titles, abstracts and full-text studies respected the inclusion criteria. the articles that respected inclusion criteria were selected for data analysis. Conflicts were resolved by at least two researchers.

Data collection concerned: title, author/authors, year, country in which the study was conducted, aim of study, study design, start date/end date, population description, age, outcome, pain scale, tool description, reliability, validity, clinical utility and feasibility, sensitivity and specificity, limitations.

Quality appraisal

To assess the quality of the included studies we used CASP check lists.

Figure 1: Flow Diagram.



Results

Summary of studies in accordance with variable properties

Table I: Summary of studies in accordance with variable properties.

Studies	TOOL Psychometric properties	CPOT	BPS	BPS-NI	ONVPS	RNVPS	NPAT	PAINAD	DPS
(Al Darwish et al., 2016)	Reliability Validity	X X	X X			X X			
(Azevedo-Santos & DeSantana, 2018)	Reliability Validity	X X	X X		X X		X X		
(Birkedal et al., 2021)	Reliability Validity	X X	X X						
(Cade, 2008)	Reliability Validity	X X	X X		X X				
(Chanques et al., 2009)	Reliability Validity		X X	X X					
(Chanques et al., 2014)	Reliability Validity	X X	X X	X X		X X			
(Chatelle et al., 2008)	Reliability Validity		X					X	X
(Creek 2019)	Reliability Validity	X X	X X	X X	X X	X X	X X	X X	
(Dorji 2019)	Reliability Validity	X X							
(Favaretto et Carraro 2015)	Reliability Validity	X X	X X			X X			
(Goe, 2018)	Reliability Validity								
(Juarez et al., 2010)	Reliability Validity		X X			X X			
(Kabes et al., 2009)	Reliability Validity				X X	X X			
(Keane, 2013)	Reliability Validity	X X							
(Li et al., 2008)	Reliability Validity	X X	X X		X X				
(Paulson-Conger et al., 2011)	Reliability Validity	X						X	
(Pereira, 2016)	Reliability Validity	X X				X X			
(Ross et al., 2016)	Reliability Validity	X X							
(Salamat et al., 2021)	Reliability Validity	X X	X X			X X			
(Severgnini et al., 2016)	Reliability Validity	X X	X X						
(Stites, 2013)	Reliability Validity	X X	X X		X X	X X	X X		
(Tapp et al., 2019)	Reliability Validity	X X						X	
(Val et al., 2009)	Reliability Validity								X X
(Weldon, 2017)	Reliability Validity	X	X						
(Young et al., 2006)	Reliability Validity		X X						

This table showed that from the 25 included studies:

- 68% discussed reliability of CPOT and 64% its validity.
- 56% treated reliability of BPS and 60% its validity.
- 12% focused on reliability of BPS-NI and 12% its validity.
- 24% examined reliability of ONVPS and 24% its validity.
- 62% reviewed reliability of RNVPS and 36% its validity.
- 12% discussed reliability of NPAT and 12% its validity.
- 8% referred to reliability of PAINAD and 12% referred to its validity.
- 4% of studies brought up DPS's reliability and 8% discussed its validity.

Summary of included studies

Table II: Summary of included studies.

Reference	title	country	design	RESULTS
(Al Darwish et al., 2016)	Evaluation of Pain Assessment Tools in Patients Receiving Mechanical Ventilation	KSA	Descriptive cohort study	In this study, all 3 nonverbal pain scales were reliable and valid: <ul style="list-style-type: none"> • BPS <ul style="list-style-type: none"> - Reliability: (r)=0.90 - Validity: Cronbach (α)= 0.95 • CPOT <ul style="list-style-type: none"> - Reliability: (r)=0.93 - Validity: Cronbach (α)= 0.95 • RNVPS <ul style="list-style-type: none"> - Reliability (r)=0.86 - Validity: Cronbach (α)= 0.86
(Azevedo-Santos & DeSantana, 2018)	Pain measurement techniques: spotlight on mechanically ventilated patients	Brazil	Systematic review of literature	<ul style="list-style-type: none"> • CPOT: this tool is valid in the following languages “French, English, Turkish, Chinese, Polish, Korean, Swedish, Finnish, Dutch, Italian, Danish and Spanish”. It was validated in the following simples “post operative heart surgery, neurosurgery, surgical wards, patients with brain injury, burned and delirium patients. The inter-rater reliability in both the English and Swedish version was greater than 0,80 (ICC). As for the French and Spanish versions weighted Kappa varied from [0,52-1]. For the internal consistency Cronbach (α) coefficients vary from 0.31 to 0.81. • BPS: it’s valid in different languages “French, English, Chinese, Swedish, Polish, Finnish, Dutch, Portuguese and Spanish”. This tool was found acceptable in the “Non-intubated Traumatic brain injury”. The internal consistency Cronbach (α) coefficient ranged from 0.63 to 0.72. The inter-rater reliability Kappa coefficient ranged from 0.67 to 0.83. • NPAT: many limitations were reported in the validation study of this tool. Samples were from Cardiothoracic surgery, cardiology, medical and surgical patients. The studies were in English. • ONVPS: It was validated in English, Iranian and finish in post cardiac surgery, trauma, burned and neurological ICU. The NVPS has good inter-rater agreement and discriminant validity (Cronbach α coefficients ranged from 0.36 to 0.75 and ICC ranged from 0.60 to 0.76)
(Birkedal et al., 2021)	Comparison of two behavioral pain scales for the assessment of procedural pain: A systematic review	Norway	Systematic review	<ul style="list-style-type: none"> • Inter-rater reliability: was validated in 8 studies. 4 of these 8 studies used weighted kappa. BPS kappa varied from [0,73 to 0,94] as for CPOT it changes from [0,70 to 0,96]. • Internal consistency: 7 studies used Cronbach (α) BPS [0,74-0,90] and CPOT [0,70-0,93]. • Discriminant validity :6 studies used different analyses method • Criterion Validity: it was cited in 3 studies. In 1 of them BPS and CPOT were associate to self-reported pain and (p < 0,5)
(Cade, 2008)	Clinical tools for the assessment of pain in sedated critically ill adults	UK	Systematic review	<ul style="list-style-type: none"> • BPS: reliability and validity were discussed in 3 studies They showed that this tool’s construct validity was higher during painful procedures. For internal consistency 2 studies calculated Cronbach(α) and it was greater than 0,64. • CPOT: In one study, construct validity (paired t-test) demonstrates this tool scores increased during positioning. For the inter-rater reliability, weighted Kappa coefficient changed from 0,52 to 0,88 during various procedures. • ONVPS: one study showed that inter-rater reliability Coefficient (α) =0,78. There was no information about internal consistency. For criterion validity, a correlation between FLACC and NVPS was calculated “Pearson’s correlation = 0,86”) (p <0,05).
(Chanques et al., 2009)	Assessing pain in non-intubated critically ill patients unable to self-report: an adaptation of the Behavioral Pain Scale	France	Observational psychometric study	<ul style="list-style-type: none"> • BPS-NI: <ul style="list-style-type: none"> - Internal consistency Cronbach (α) = 0.79 - Inter-rater reliability weighted Kappa coefficient varies from 0,82 to 0,89 - Discriminant validity: high scores were observed during painful procedures. There wasn’t any change in scores in non-painful procedures.

(Chanques et al., 2014)	Psychometric comparison of three behavioural scales for the assessment of pain in critically ill patients unable to self-report	USA	Descriptive cohort study	<ul style="list-style-type: none"> • BPS and BPS-NI: <ul style="list-style-type: none"> - Internal consistency: Cronbach (α)=0,80 - Discriminant validation: the score was higher during procedures. • CPOT: <ul style="list-style-type: none"> - Internal consistency: Cronbach (α)=0,81 - Discriminant validation: the score was higher during procedures. • ONVPS: <ul style="list-style-type: none"> - Internal consistency: Cronbach (α)=0,76 - Discriminant validation: the score was higher during procedures.
(Chatelle et al., 2008)	Mesurer la douleur chez le patient non communicant	Belgium	Systematic review	<ul style="list-style-type: none"> • PAINAD: was developed to assess pain in patients with advanced dementia. • BPS: is a validated pain assessment tool. It includes the following items body movements and respiratory rate. • DPS: still in the validation stage. This pain scale evaluates face, respiratory rate and tears.
(Creek, 2019)	Pain assessment tools for the non-verbal critical care adult : an integrative review of the literature	USA	Integrative review of the literature	<ul style="list-style-type: none"> • BPS: <ul style="list-style-type: none"> - Discriminant validity: remarkable increase during nociceptive periods - Construct validity: 4 studies showed that there was a big difference between scores in rest and painful procedures - Criterion validity: 2 studies showed that there was a correlation between BPS and self-report - Internal consistency: In 8 studies, Cronbach(α) varies from 0,70 to 0,95 - Inter-rater reliability: 4 studies showed good agreement • BPS-NI: <ul style="list-style-type: none"> - Internal consistency: 1 study showed internal consistency >0,70 - Inter-rater reliability: Kappa > 0,60 - Validity: low • CPOT: <ul style="list-style-type: none"> - Construct validity: high during painful procedures - Discriminant validity: 5 studies showed that it was high during nociceptive periods - Criterion validity: average correlation to self-report ($r=0,419$, $p<0,01$) - Internal consistency: Cronbach (α) changes from 0,70 to 0,95 - Inter-rater reliability: it was found to be moderate to good • ONVPS: <ul style="list-style-type: none"> - Discriminant validity: it was supported by 6 studies - Criterion validity: it was low in physiology item - Inter-rater reliability: 2 studies showed that Kappa varies between 0,71 and 0,80. • RNVPS: <ul style="list-style-type: none"> - Criterion validity: these items were high during self-reporting pain "physiology 1: (19,8%)" and "physiology2: (40%)" - Discriminant validity: high in painful procedures - Internal consistency: Cronbach (α)> 0,70 - Inter-rater reliability: good • PAINAD: <ul style="list-style-type: none"> - Inter-rater reliability: ICC=0,80 • NPAT: <ul style="list-style-type: none"> - Validity: moderate - Internal consistency: Cronbach (α)>0,70
(Dorji, 2019)	Implementation and Evaluation of Critical Care Pain Observation Implementation and Evaluation of Critical Care Pain Observation Tool (CPOT)	USA	descriptive study	<ul style="list-style-type: none"> • Inter-rater reliability: agreement of 93,3%. • Validity: 97,95% of nurses agreed that CPOT is a valid tool.

(Favaretto & Carraro, 2015)	Gli strumenti di valutazione del dolore per il paziente critico non comunicante in terapia intensiva. Revisione di letteratura	Italy	systematic review	<ul style="list-style-type: none"> • BPS: <ul style="list-style-type: none"> - Inter-rater reliability: good - Discriminant validity: valid • CPOT: <ul style="list-style-type: none"> - Inter-rater reliability: good - Discriminant validity: most valid • RNVPS: <ul style="list-style-type: none"> - Inter-rater reliability: low - Discriminant validity: valid
(Goe, 2018)	Evaluation of a Critical-Care Pain Observation Tool Quality Evaluation of a Critical-Care Pain Observation Tool Quality Initiative	USA	Retrospective pre and post design	<ul style="list-style-type: none"> • The frequency of pain assessment documentation by nurses was higher after the implementation of CPOT. • No difference was found in the use of analgesic and sedatives in the pre and post implementation groups. • There wasn't any remarkable change in patient's outcome.
(Juarez et al., 2010)	Comparison of Two Pain Scales for the Assessment of Pain in the Ventilated Adult Patient	USA	Study expanded upon a pilot study	<ul style="list-style-type: none"> • BPS: <ul style="list-style-type: none"> - Reliability: Cronbach (α)=0,70 • RNVPS: <ul style="list-style-type: none"> - Reliability: Cronbach (α)=0,75 • Scores were higher in surgical patients than medical patients. • Scores were higher during turning compared to rest in both pain scales.
(Kabas et al., 2009)	Further Validation of the Nonverbal Pain Scale in Intensive Care Patients	USA	Nonexperimental and methodological. Observational	<ul style="list-style-type: none"> • ONVPS: <ul style="list-style-type: none"> - Inter-rater reliability: 94.7% agreement. - Internal consistency: Cronbach (α) was high during procedures and after 0,62. Before procedures (α)= 0,36 - Validity: Valid • RNVPS: <ul style="list-style-type: none"> - Inter-rater reliability: 90.8% agreement. - Internal consistency: Cronbach (α)=0,36 during (α)=0,72 and after (α)=0,71 - Validity: valid. Mean rank is high during procedures.
(Keane, 2013)	Validity and reliability of the Critical Care Pain Observation Tool: A Validity and reliability of the Critical Care Pain Observation Tool: A replication study	USA	Cohort study	<ul style="list-style-type: none"> • Inter-rater reliability: weighted Kappa varies from 0,34 to 1,0 • Discriminant validity: The score was higher during nociceptive periods
(Li et al., 2008)	A Review of Objective Pain Measures for Use with Critical Care Adult Patients Unable to Self-Report	USA	Critical review	<ul style="list-style-type: none"> • BPS: <ul style="list-style-type: none"> - Construct validity: the score was high during painful procedure P .01 - Internal consistency: Cronbach's 0.6444; 0.721) - Inter-rater reliability: It was not consistent between 3 studies • CPOT: <ul style="list-style-type: none"> - Criterion validity: not confirmed - Internal consistency: no data - Inter-rater reliability: good agreement, • ONVPS: <ul style="list-style-type: none"> - Construct validity: - Internal consistency: good coefficient of 0.78. - Inter-rater reliability: no information.
(Paulson-Conger et al., 2011)	Comparison of Two Pain Assessment Tools in Nonverbal Critical Care Patients	USA	Descriptive, comparative, prospective design	<ul style="list-style-type: none"> • Internal consistency: <ul style="list-style-type: none"> - CPOT: 0,76. - PAINAD: 0,80. • The correlation between the PAINAD and CPOT was 0.86 ($p < 0,01$). • The correlation between the average of the two scores and their difference was 0.27 ($p < 0 .05$).
(Pereira, 2016)	Acute Pain Symptom Assessment and Management in Nonverbal Puerto Rican Patients in the Early Postoperative Period	USA	Cohort study	<ul style="list-style-type: none"> • CPOT and RNVPS are positively correlated: <ul style="list-style-type: none"> - Upon arrival to PACU: correlation coefficient $r=0,88$. - After 120 minutes $r=0,89$.
(Ross et al., 2016)	Validation of the Critical-Care Pain Observation Tool with Seriously Ill Patients	Canada	Prospective cohort study	<ul style="list-style-type: none"> • Inter-rater reliability: ICC > 0,75 "all procedures nociceptive and non-nociceptive". • Discriminant validity: non difference in scores before, during and after gentle touch, it was different during turning procedures "Friedman test".
(Salamat et al., 2021)	A Systematic Review of Pain Assessment in	Iran	Systematic Review	<ul style="list-style-type: none"> • CPOT

	Mechanically Ventilated Patients			<ul style="list-style-type: none"> - The internal consistency: Cronbach (α) varies between 0.56 and 0.94 - Inter-rater reliability: (Intraclass Correlation Coefficient) ICC = 0.56 in one study, but it was good in three studies (ICC \geq 0.91) • BPS <ul style="list-style-type: none"> - The internal consistency: Cronbach (α) > 0.7 in six studies and < 0.70 in two studies, also it was cited in two studies that (α)=0.950 - Inter-rater reliability: (Intraclass Correlation Coefficient) ICC > 0.86 • RNVPS <ul style="list-style-type: none"> - The internal consistency: Cronbach (α) > 0.70 - Inter-rater reliability: (Intraclass Correlation Coefficient) ICC > 0.68 in two studies.
(Severgnini et al., 2016)	Accuracy of Critical Care Pain Observation Tool and Behavioral Pain Scale to assess pain in critically ill conscious and unconscious patients: prospective, observational study	Italy	Prospective, observational study	<ul style="list-style-type: none"> • CPOT and BPS showed a good criterion and discriminant validity ($p < 0.0001$). • BPS is more specific (91.7 %) than CPOT (70.8 %), but less sensitive (BPS 62.7 %, CPOT 76.5%). • COPT and BPS scores were significantly correlated with VAS ($p < 0.0001$).
(Stites, 2013)	Observational Pain Scales in Critically Ill Adults	USA	Review of the Literature	<ul style="list-style-type: none"> • CPOT: <ul style="list-style-type: none"> - Criterion validity: the score was low in patients that denies pain - Discriminate validity: the scores were high during stimulus. - Inter-rater reliability: In two studies Kappa varies between 0,52 and 0,88. • BPS: <ul style="list-style-type: none"> - Discriminant validity: the score was high during painful stimulation - Construct validity: there was a correlation between BPS and NRS - Inter-rater reliability: Cronbach (α)=0,94 - Internal consistency: Cronbach (α)> 0,63 • NVPS: <ul style="list-style-type: none"> - Discriminant validity: no data - Construct validity: correlation of NVPS and NRS poor, 0.38. - Inter-rater reliability: in both the ONVPS and RNVPS, the score levels were high during painful periods. - Internal consistency: Cronbach (α)= 0.80 • NPAT: <ul style="list-style-type: none"> - Validity: weak - Inter-rater reliability: concordance coefficient=0.72; 95% confidence interval
(Tapp et al., 2019)	Observational Pain Assessment Instruments for Use with Nonverbal Patients at the End-of-life: A Systematic Review	Canada	Systematic Review	<ul style="list-style-type: none"> • CPOT: It was invented by Gelinas et al in 2007. • PAINAD: <ul style="list-style-type: none"> - Good agreement in the following domains: facial expression, vocalization and body language. - Good indicator: physical item - Less good indicators: were the following categories: breathing, consolability, behavioural and physiological changes.
(Val et al., 2009)	Evaluation de la douleur chez le patient non communicant en Unité de Soins Intensifs	Belgium	Cohort	<ul style="list-style-type: none"> • Internal consistency : [0,559 - 0,637]
(Weldon, 2017)	Comparison of the Behavioral Pain Scale and the Critical-Care Pain Observation Tool in Assessing Pain in Ventilated Critical Care Patients	USA	Non-experimental, correlational, comparative design	<ul style="list-style-type: none"> • CPOT: Cronbach (α) =0,921 • BPS: Cronbach (α) =0.95 • CPOT and BPS scores were higher during turning.
(Young et al., 2006)	Use of a Behavioural Pain Scale to assess pain in ventilated, unconscious and/or sedated patients	Australia	Prospective, descriptive repeated measures study design	<ul style="list-style-type: none"> • Internal consistency: The Cronbach's alpha value for the BPS was 0.64, indicating moderate to good reliability • Inter-rater reliability: <ul style="list-style-type: none"> - Before procedures: good agreement (82% to 91%) - After procedures: low agreement "eye care range between 64% and 73%" and "after repositioning: ranging between 36% and 46%."

Quality assessment

Table III: Quality appraisal of included studies.

Reference	1	2	3	4	5	6	7	8	9	10	11	12	total
(Al Darwish et al., 2016)	Y	Y	Y	Y b-Y	a-Y b-C	a-Y	Y	Y	Y	Y	Y	Y C:1 N:0	Y:13
(Azevedo-Santos & DeSantana, 2018)	Y	C	Y	N	Y	Y	Y	Y	Y	Y		C:1 N:1	Y:8
(Birkedal et al., 2021)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		C:0 N:0	Y:10
(Cade, 2008)	Y	C	Y	C	Y	Y	Y	Y	Y	Y		C:2 N:0	Y:8
(Chanques et al., 2009)	Y	Y	Y	Y b-Y	a-Y b-C	a-Y	Y	Y	Y	Y	Y	Y C:1 N:0	Y:13
(Chanques et al., 2014)	Y	Y	Y	Y b-Y	a-Y b-C	a-C	Y	Y	Y	Y	Y	Y C:2 N:0	Y:12
(Chatelle et al.2008)	Y	C	N	C	Y	Y	Y	Y	Y	Y		C:2 N:1	Y:7
(Creek,2019)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y		C:0 N:1	Y:9
(Dorji, s. 2019)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		C:0 N:0	Y:10
(Favaretto & Carraro, 2015)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y		C:0 N:1	Y:9
(Goe, 2018)	Y	C	Y	Y	Y b-Y	a-Y	C	Y	Y	Y	Y	C:2 N:0	Y:10
(Juarez et al., 2010)	Y	Y	Y	Y b-Y	a-Y b-Y	a-Y	Y	Y	Y	Y	Y	Y C:0 N:0	Y:14
(Kabes et al., 2009)	Y	C	Y	Y b-C	a-C b-C	a-Y	Y	Y	Y	Y	Y	Y C:4 N:0	Y:10
(Keane, 2013)	Y	Y	Y	Y b-Y	a-Y b-Y	a-Y	Y	Y	Y	Y	Y	Y C:0 N:0	Y:14
(Li et al., 2008)	Y	C	Y	C	Y	Y	Y	Y	Y	Y		C:2 N:0	Y:8
(Paulson-Conger et al., 2011)	Y	Y	Y	Y b-Y	a-Y b-Y	a-Y	Y	Y	Y	Y	Y	Y C:0 N:0	Y:14
(Pereira,2016)	Y	Y	Y	Y b-Y	a-Y b-C	a-Y	Y	Y	Y	Y	Y	Y C:1 N:0	Y:13
(Ross et al., 2016)	Y	Y	Y	Y b-Y	a-Y b-C	a-Y	Y	Y	Y	Y	Y	Y C:1 N:0	Y:13
(Salamat et al., 2021)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		C:0 N:0	Y:10
(Severgnini et al., 2016)	Y	Y	Y	Y b-Y	a-Y b-C	a-Y	Y	Y	Y	Y	Y	Y C:1 N:0	Y:13
(Stites, 2013)	Y	N	C	N	Y	Y	Y	Y	Y	Y		C:1 N:2	Y:7
(Tapp et al., 2019)	Y	Y	N	Y	Y	Y	Y	Y	C	Y		C:1 N:1	Y:8
(Val et al., 2009)	Y	Y	Y	C a-Y	a-Y b-Y	a-C	Y	Y	Y	Y	C	Y C:3 N:0	Y:11
(Weldon,2017)	Y	Y	Y	Y	Y b-C	a-Y	C	Y	Y	Y		C:2 N:0	Y:9
(Young et al., 2006)	Y	Y	Y	Y b-Y	a-Y b-N	a-C	Y	Y	Y	Y	Y	Y C:1 N:1	Y:12

Y: yes, C: can't tell, N: no CASP checklists were used to assess the quality of included studies. Studies scored good quality are the studies that don't contain any "no" answers. Studies scored average quality, are the studies which contains 1 "no" answer. Studies scored low quality are the studies which contains 2 or more "no" answers.

Table IV: Selected assessment tools and their domains.

Pain scales	Domains	Scoring	Total scoring
BPS	Facial expressions Movements of upper limbs Compliance with ventilation	1-abr 1-abr 1-abr	3-dic
BPS-NI	Facial expressions Movements of upper limbs Vocalization	1-abr 1-abr 1-abr	3-dic
ONVPS	Face Activity Guarding Physiologic 1 Physiologic 2	0-2 0-2 0-2 0-2 0-2	0-10
RNVPS	Face Activity Guarding Physiologic Respiratory	0-2 0-2 0-2 0-2 0-2	0-10
CPOT	Facial expression Body movements Compliance with the ventilator or vocalization Muscle tension	0-2 0-2 0-2 0-2	0-8
NPAT	An affective response to a situation Change in placement and positioning of the body and extremities when not engaged in any care activities Sound cues or vocalizations other than speech Expressions on face Body responses that imply a protection of the body from contact with external touch	0-2 0-3 0-2 0-2 0-3	0-12
PAINAD	Breathing Negative vocalization Facial expression Body language Consolability	0-2 0-2 0-2 0-2 0-2	0-10
DPS	Adaptation au respirateur Expression du visage Comportement moteur Larmes	1-abr 1-abr 1-may 1-feb	abr-15

Behavioral pain scale (BPS):

According to², BPS pain scores increased during the painful repositioning procedure comparing to the non-painful eyecare procedure, they also made a reference to the fact that changes in BPS score during eye care is a result of autonomic responses to touch not pain since patients had not any facial trauma or facial surgery. These findings are consistent with those reported by³⁻⁸researches who found also that BPS scores increases significantly between painful and non-painful situations, which supports BPS discriminant validity. The research of² supported also the validity of BPS with logistic regression analysis, which showed that the odds of an increase in BPS score was 25 times higher for repositioning as a painful procedure compared with eye care as a non-painful procedure.

Regarding the reliability of BPS² revealed that due to small numbers of data, it was not realizable to conduct formal tests of reliability. Inter-rater reliability results showed that it was easier for two assessors to agree on the level of low pain but when assessing increased pain level some variance was found when using the BPS. The research of⁹ showed that the BPS met the homogeneity criterion for reliability indicator. And it has shown an evident Inter-rater reliability. In the study of⁴, it was found that the reliability of BPS was nearly perfect (kappa coefficient: 0.81 ± 0.03).

Behavioral pain scale non intubated (BPS-NI):

The discriminative validation was supported in the study of⁴ by the significant increasing of the median BPS-NI value from rest to nociceptive procedure. The good internal consistency was supported by a Cronbach value (standardized Cronbach $\alpha = 0.79$), which means that BPS-NI items were well correlated between them. The discriminant validity of this tool was supported in the study of⁴ by the increasing of its median score from baseline to nociceptive procedure ($p < 0.001$) and a significant decrease 10 minutes after the nociceptive procedure ($p < 0.001$).

In the study of⁴ the inter-rater reliability was evaluated by weighted kappa coefficients and it was nearly perfect for BPS. The weighted kappa coefficients for each domain were as following, for the facial domain, it was 0.75 which is considered as an excellent kappa, for the vocalization domain, it was 0.78 which is considered as an excellent kappa, and finally the upper limb domain with a good kappa of 0.61. In the study of⁷, it was found that inter-rater agreement was 96% for the BPS-NI scores for both types of procedures (nociceptive and non-nociceptive) and 90% for the BPS-NI scores for painful procedures only, which is considered as a good inter-rater reliability.

Original non verbal pain scale (ONVPS):

⁸found that this ONVPS tool was not adequately tested. The developers of this tool compared the ONVPS with the FLACC as a 'gold standard', while the FLACC was not a validated tool among adult population. Which is not rational in principle. The authors¹⁰ found that the physiologic II item was not discriminant at an acceptable level because it showed a small variation between before, during, after a nociceptive procedure. Regarding the internal consistency, it was found in the study of⁴ that NVPS internal consistency was low in non-intubated non-verbal patients. Therefore, this tool was not validated for non-intubated non-verbal ICU patients.

In the study of⁴, the domains of ONVPS were evaluated by weighted kappa coefficients to examine the inter-rater reliability of this tool. The essential results were as following: for the respiration and activity domain, it was 0.54 and 0.52 respectively, which is considered a moderate kappa. For the guarding domain, it was a poor kappa of 0.32, and finally the physiologic II domain which had a very poor kappa of 0.02. Generally, the ONVPS demonstrated poor to just moderate inter-rater reliability.

Revised non-verbal pain scale (RNVPS):

In the study of¹¹, it was found that the RNVPS showed satisfactory validity, with a Cronbach $\alpha = 0.86$. The Physiology and Respiratory domains showed low psychometric properties evidence, they were not detecting pain. The studies of^{4,6} showed that this tool is discriminant valid, since there was a significant difference in median scores during two similar situations which differs by the intensity and the length of the procedures.

The RNVPS was reliable according to¹², with a Cronbach α coefficient of 0.89. And this result is consistent with the findings of¹¹, who found a Cronbach α coefficient of 0.86. In the study of⁶, it was shown that this instrument had also a good inter-rater reliability with an ICC=0.92 and ICC = 0.68 in two different surveys.

Critical care pain observation tool (CPOT):

The CPOT is the most discussed instrument in the 25 included studies. Its evidence of reliability and validity were clearly represented in the studies of^{4,13}. In the study of¹³, it was found that the validity of this tool was supported by the Friedman test results, which showed in the first hand that CPOT's score had negligible differences before, during and after a non-nociceptive procedure (gentle touch). While on the other hand, it showed a very significant difference in the same three times during turning procedure considered as painful situation. The study of⁴ showed that CPOT's validity was supported by the significant increasing of its score from rest time to nociceptive procedure ($P < 0.001$) and a decreasing 10 min after the procedure. And regarding the responsiveness, it was tested by the effect size coefficient which was large (> 0.80). Reference must be made to the fact that in spite of all CPOT advantages,

it has some limitations, essentially: its use among traumatic brain injured patients who present different facial expressions in comparison with the other patients when experiencing pain³. Therefore, more studies are necessary to generalize the use of this tool in ICUs.

Inter-rater reliability of CPOT was tested in the research of⁴ using the weighted kappa coefficients, and results were as following: for the facial category the coefficient was 0.81 which represents an excellent kappa, for the breathing domains, the kappa was 0.71 which is a good kappa, and finally the body movements and muscle tension domains with a moderate kappa of 0.42 and 0.43 respectively. In the study of¹³, the reliability was tested with the interclass coefficient correlation (ICC), the results were as following: the highest ICC was detected before the nociceptive turning procedure, whereas the lowest ICC was observed with assessments related to gentle touch as a non-nociceptive procedure.

Nonverbal pain assessment tool (NPAT):

The authors of³ and ¹⁴ found many limitations in their search essentially there was no information provided about the situations when patients pain were assessed by this tool. Therefore, they suggested that more studies are paramount to test NPAT validity and applicability among non-verbal ICU population. In the study of¹², it was revealed that in spite of the fact that NPAT was developed for non-verbal patients, it has never been validated in this population. Moreover, the NPAT consist of a domain of "verbal cues" which is contradictory with the principle of a nonverbal pain assessment tool.

Pain assessment in advanced dementia scale (PAINAD):

The authors¹⁵ found that interrater reliability of the PAINAD indicated good reliability across many searches with a Pearson r range from 0.75 to 0.97, reference must be made to the fact that most Pearson r ranges over 0.80. They found also that this instrument is discriminant valid since its scores are significantly higher during painful procedures comparing with non-painful procedures. The Validity was also demonstrated by using quality improvement data. The results showed that PAINAD scores before pain medication with dosage according to this tool were 6.7 ± 1.8 and after 30 minutes of pain medication administration were 1.8 ± 2.2 . Therefore, PAINAD was considered valid. For¹⁴, this tool needs more assessing of the construct validity, discriminant validity, and internal consistency in the adult critical care setting since PAINAD was developed for a superior reason adults with cognitive impairment.

SOS Doulousi pain scale (DPS) :

This pain assessment tool was not tested enough to examine its psychometric properties evidence, it was validated only by its developer¹⁰, using Wilcoxon test for non-parametric data. The results showed that the pain scores change in three situations differing by the intensity

of the procedure: at rest, non-nociceptive procedure, and during a nociceptive procedure. The scores seemed to increase in accordance with growing discomfort. The authors, also tested the DPS's internal consistency using Cronbach's test which varied from 0.559 to 0.637 depending on the situation¹⁰.

Discussion

In this review, non-verbal pain assessment tools were tested for two paramount psychometric properties: validity and reliability. Some tools showed good evidence of these properties, while some did not. And some showed their evidence among a particular population.

CPOT, the most utilized instrument in the included studies, showed nearly perfect reliability and validity properties. Its validity is supported essentially by the significant increasing of its score from resting time to nociceptive procedure. Its inter-rater reliability showed also good evidence with a moderate to excellent weighted kappa coefficients of its domains. It revealed some limitations among ICU patients with traumatic brain injury, however, for the general ICU non-verbal population, it is recommended to be the first choice.

Concerning BPS, BPS-NI, RNVPS, and PAINAD, all these tools showed good validity and reliability. The BPS score was invented by Payen in 2001. This tool showed good discriminant and construct validity (scores were high during painful procedures compared to non-painful procedures like rest or eye care. For criterion validity, Creek, n.d. showed a correlation between BPS and self-report. In most studies, inter-rater reliability Kappa was higher than 0,67 and internal consistency higher than 0,64. The BPS can be used for patients with traumatic brain injury. The BPS-NI is devoted only to non-intubated patients unable to express their pain. The "compliance with ventilation" domain from BPS score was changed into "vocalization" domain. ⁷ study showed that BPS-ni's reliability is good (Internal consistency Cronbach (α) = 0.79 and Inter-rater reliability weighted Kappa coefficient varies between 0,82 and 0,89), also the scores were high during painful procedures and there wasn't any change in scores in non-painful procedures which indicates good validity. In the RNCVPS, the physiologic 2 item in the ONVPS was changed to respiratory item. Many authors

support the use of the Revised rather than the original version of the NVPS because it did demonstrate good evidence of psychometric properties. ¹⁰ support this score's validity because the mean rank was high during painful procedures. For the reliability, two studies showed an internal consistency Cronbach (α) > 0,70¹⁶ and (Creek, n.d.). PAINAD was adapted from the FLACC and DS-DAT scores. It was developed to assess pain in Non-verbal patients with advanced dementia. This tool has good agreement in the following domains: facial expression, vocalization and body language and good. Inter-rater reliability of this tool was good (ICC=0,80)¹⁷. Each of these instruments is specific to a certain population in ICU and to generalize them among all adult non-verbal ICU population, more researches and reconstruction of these tools are certainly necessary.

Low evidence of psychometric properties was revealed in each of DPS, NPAT and ONVPS. The first was only validated by its developer, so more studies are necessary to test the validity and reliability among our population. The second tool, despite the fact that it was developed originally for this population, it was never been validated among them, therefore this tool is not recommended to be used. In the developing process of ONVPS, this tool was compared to FLACC as a gold standard, while the FLACC was never validated among our population, we can't consider this instrument valid and reliable. Also³⁴ study showed that Criterion validity was low in physiology item, which mean this instrument need more testing.

Conclusion

In a nut shell, this systematic review revealed that. In the one hand, each of DPS, NPAT and ONVPS are considered neither, reliable nor valid since they showed weak evidence of psychometric properties. In the other hand, the CPOT is to be preferred among general non-verbal ICU population since it showed nearly good properties. Concerning, BPS, BPS-NI, RNVPS and PAINAD, it is recommended to use each for its specific population.

Competing interests

The authors certify that there is no conflict of interest.

References

1. Kotfis K, Zegan-Barańska M, Szydłowski Ł, Żukowski M, Ely WE. Metody oceny natężenia bólu u dorosłych pacjentów oddziałów intensywnej terapii – polska wersja językowa narzędzia CPOT (Critical Care Pain Observation Tool) i BPS (Behavioral Pain Scale). *Anaesthesiol Intensive Ther.* 2017 Mar;49(1):66-72. doi: 10.5603/AIT.2017.0010
2. Young J, Siffleet J, Nikolett S, Shaw T. Use of a Behavioral Pain Scale to assess pain in ventilated, unconscious and/or sedated patients. *Intensive and Critical Care Nursing.* 2006 Feb;22(1):32-9. doi: 10.1016/j.iccn.2005.04.004.
3. Al Darwish ZQ, Hamdi R, Fallatah S. Evaluation of Pain Assessment Tools in Patients Receiving Mechanical Ventilation. *AACN Advanced Critical Care.* 2016 Apr;27(2):162-72. doi: 10.4037/aacnacc2016287.
4. Azevedo-Santos I, DeSantana J. Pain measurement techniques: spotlight on mechanically ventilated patients. *JPR.* 2018 Nov;11:2969-80. doi: 10.2147/JPR.S151169.
5. Birkedal HC, Larsen MH, Steindal SA, Solberg MT. Comparison of two behavioural pain scales for the assessment of procedural pain: A systematic review. *Nurs Open.* 2021 Sep;8(5):2050-60. doi: 10.1002/nop2.714.
6. Cade CH. Clinical tools for the assessment of pain in sedated critically ill adults. *Nursing in Critical Care.* 2008 Oct;13(6):288-97. doi: 10.1111/j.1478-5153.2008.00294.x.
7. Chanques G, Payen JF, Mercier G, De Lattre S, Viel E, Jung B, et al. Assessing pain in non-intubated critically ill patients unable to self-report: an adaptation of the Behavioral Pain Scale. *Intensive Care Med.* 2009 Dec;35(12):2060. doi: 10.1007/s00134-009-1590-5.
8. Chanques G, Pohlman A, Kress JP, Molinari N, De Jong A, Jaber S, et al. Psychometric comparison of three behavioral scales for the assessment of pain in critically ill patients unable to self-report. *Crit Care.* 2014;18(5):R160. doi: 10.1186/cc14000.
9. Chatelle C, Vanhauzenhuysse A, Mergam AN, De Val M, Majerus S, Boly M, et al. mesurer la douleur chez le patient non communicant. *Rev Med Liege.*
10. Creek T. Pain Assessment Tools for the Nonverbal Critical Care Adult: An Integrative Review of the Literature.
11. Dorji KD. Implementation and Evaluation of Critical Care Pain Observation Tool (CPOT).
12. Favaretto DA, Carraro F. Gli strumenti di valutazione del dolore per il paziente critico non comunicante in terapia intensiva. *Revisione di letteratura.*
13. Goe CA. Evaluation of a Critical-Care Pain Observation Tool Quality Initiative.
14. Juarez P, Patricia MS, Deborah MS, Michelle MS, Bonnie A, Cheryl. Comparison of Two Pain Scales for the Assessment of Pain in the Ventilated Adult Patient. *Dimensions of Critical Care Nursing.* 2010 Nov;29(6):307-15. doi: 10.1097/DCC.0b013e3181f0c48f.
15. Kabes AM, Graves JK, Norris J. Further Validation of the Nonverbal Pain Scale in Intensive Care Patients. *Critical Care Nurse.* 2009 Feb;29(1):59-66. doi: 10.4037/ccn2009992.
16. Keane KM. Validity and Reliability of the Critical Care Pain Observation Tool: A Replication Study. *Pain Management Nursing.* 2013 Dec;14(4):e216–e225. doi: 10.1016/j.pmn.2012.01.002.
17. Li D, Puntillo K, Miaskowski C. A Review of Objective Pain Measures for Use With Critical Care Adult Patients Unable to Self-Report. *The Journal of Pain.* 2008 Jan;9(1):2-10. doi: 10.1016/j.jpain.2007.08.009.
18. Paulson-Conger M, Leske J, Maidl C, Hanson A, Dziadulewicz L. Comparison of Two Pain Assessment Tools in Nonverbal Critical Care Patients. *Pain Management Nursing.* 2011 Dec;12(4):218-24. doi: 10.1016/j.pmn.2010.05.008.
19. Pereira S. Acute Pain Symptom Assessment and Management in Nonverbal Puerto Rican Patients in the Early Postoperative Period. *University of Massachusetts Amherst.* doi: 10.7275/8994977.0.
20. Ross M, Boitor M, Gélinas C. Validation of the Critical-Care Pain Observation Tool With Seriously Ill Patients. *J Hosp Palliat Nurs.* 2016 Oct;18(5):413-20. doi: 10.1097/NJH.0000000000000266.
21. Salamat E, Sharifi F, Hasanloei MAV, Bahramnezhad F. A Systematic Review of Pain Assessment in Mechanically Ventilated Patients. *TOPAINJ.* 2021 Nov;14(1):22-31. doi: 10.2174/1876386302114010022.
22. Severgnini P, Pelosi P, Contino E, Serafinelli E, Novario R, Chiaranda M. Accuracy of Critical Care Pain Observation Tool and Behavioral Pain Scale to assess pain in critically ill conscious and unconscious patients: prospective, observational study. *J intensive care.* 2016 Dec;4(1):68. doi: 10.1186/s40560-016-0192-x.
23. Stites M. Observational Pain Scales in Critically Ill Adults. *Critical Care Nurse.* 2013 Jun;33(3):68-78. doi: 10.4037/ccn2013804.
24. Tapp D, Chenacher S, Gerard NPA, Berube-Mercier P, Gelinac C, Douville F, et al. Observational Pain Assessment Instruments for Use With Nonverbal Patients at the End-of-Life: A Systematic Review. *J Palliat Care.* 2019 Oct;34(4):255-66. doi: 10.1177/0825859718816073.
25. Val MDD, Pachioli M, Lorent C, Nguyen T. Evaluation de la douleur. 2009.
26. Weldon J. Comparison of the Behavioral Pain Scale and the Critical-Care Pain Observation Tool in Assessing Pain in Ventilated Critical Care Patients.
27. Al Darwish ZQ, Hamdi R, Fallatah S. Evaluation of Pain Assessment Tools in Patients Receiving Mechanical Ventilation. *AACN Advanced Critical Care.* 2016 Apr;27(2):162-72. doi: 10.4037/aacnacc2016287.
28. Azevedo-Santos I, DeSantana J. Pain measurement techniques: spotlight on mechanically ventilated patients. *JPR.* 2018 Nov;11:2969-80. doi: 10.2147/JPR.S151169.
29. Birkedal HC, Larsen MH, Steindal SA, Solberg MT. Comparison of two behavioural pain scales for the assessment of procedural pain: A systematic review. *Nurs Open.* 2021 Sep;8(5):2050-60. doi: 10.1002/nop2.714.
30. Chanques G, Payen JF, Mercier G, De Lattre S, Viel E, Jung B, et al. Assessing pain in non-intubated critically ill patients unable to self-report: an adaptation of the Behavioral Pain Scale. *Intensive Care Med.* 2009 Dec;35(12):2060. doi: 10.1007/s00134-009-1590-5.
31. Chanques GG, Pohlman A, Kress JP, Molinari N, De Jong A, Jaber S, et al. Psychometric comparison of three behavioral scales for the assessment of pain in critically ill patients unable to self-report. *Crit Care.* 2014;18(5):R160. doi: 10.1186/cc14000.
32. Chatelle C, Vanhauzenhuysse A, Mergam AN, De Val M, Majerus S, Boly M, et al. mesurer la douleur chez le patient non communicant. *Rev Med Liege.*