Scoring systems, expert assessment, and identification of risk factors for the emergence of delirium in paediatric patients: Prospective cohort study

Djayanti Sari, PhD[1], Yunita Widyastuti, PhD[1], Haniifa Herly Hendy, MD[2], Ignasius Adi Dharma, MD[2], Anindita Pancarani, MD[2], Andre Krislee, MD[2]

[1]Anesthesiology and Intensive Therapy Department, Faculty of Medicine, Public Health, and Nursing, Gadjah Mada University/Dr. Sardjito General Hospital, Yogyakarta, Indonesia, [2]Faculty of Medicine, Public Health, and Nursing, Gadjah Mada University, Yogyakarta, Indonesia

ABSTRACT

Introduction: Emergence delirium (ED) is a transient irritative and dissociative state that arises after the cessation of anaesthesia in patients who do not respond to calming measures. There are many risk factors for ED, but the exact cause and underlying mechanism have not been determined because the definition of ED is still unclear in consensus. This study aims to determine ED incidence, identify ED risk factors and external validation of Watcha, Cravero and expert assessment to Pediatric Anesthesia Emergence Delirium (PAED) scoring system in ED prediction.

Materials and Methods: This study is a prospective cohort study on 79 paediatrics who underwent elective surgery with general anaesthesia. Parameter measures include the incidence of ED, ED risk factors, and the relationship between PAED, Watcha, Cravero score and expert assessment. The ED risk factor was analysed using univariate and multivariate analysis. The relationship between PAED, Watcha, Cravero score, and expert assessment was determined using Receiver Operating Characteristic (ROC) curve analysis.

Results: The incidence of ED was 22.8%. All parameters examined in this study showed $p < 0.05$. Watcha's scoring correlates with the PAED scoring and shows the highest discrimination ability with AUC 0.741 and $p < 0.05$.

Conclusion: The incidence of ED in paediatrics is relatively high. Compared to others, Watcha score are more reliable for ED prediction. However, some demographic and perioperative factors are not the risk factor of ED.

KEYWORDS:
Emergence delirium, general anaesthesia, paediatric, risk factor, scoring system

INTRODUCTION

Emergence delirium (ED) is a mental disturbance during recovery from general anaesthesia, which consists of hallucination, delusion and confusion that manifests with moaning, anxiety, involuntary physical activity, and thrashing in bed. ED is a term often used to describe changes in behaviour after anaesthesia. Nevertheless, until now, there is no clear consensus for ED.[1] The incidence of ED is still unclear, ranging from 10 to 80%.[2,3] In some studies, it is mentioned that in conditions where pain and other confounding factors can be controlled, the incidence of ED may drop from 2 to 80% to 20 to 30%. Although ED can occur in adults, the highest incidence of ED is in children aged 2 to 7 years.[4]

It is a self-limited disorder (lasting 5 to 25 minutes after cessation of anaesthesia). It can repeatedly occur for up to 2 days and can be dangerous to the patient self. It may cause physical damage by removing intravenous lines, drainage tubes, patient monitoring devices, dressings or wound protectors.[1] A large prospective cohort study states that morbidity and mortality of postoperative delirium is around 7.2 to 8.7%.[5]

Controlling ED in children is a challenge. It requires more time, a longer length of stay in post anaesthesia care unit (PACU), and the addition of staff to treat patients after anesthesia.[3] During this decade, ED became one of the concerns whether it happened after surgery or not. Children with ED have a 1.43 times greater risk of experiencing maladaptive behaviour with long-term effects.[6] In different studies reported in patients who are preoperatively experiencing anxiety, the incidence of ED increases and even persists for up to 14 days.[7]

There are many risk factors for ED, but the exact cause and underlying mechanism have not been determined because the definition of ED is still unclear in consensus. Some of the factors that are considered to cause ED are age (children around 74%), previous ED (65%), mental condition of the patient (severe anxiety around 57%), untreated postoperative pain (60%), aesthetic method (with gas anaesthesia, rapid emergence in about 55%), and surgical procedures (surgery of the ear, eye or tooth in about 50%).[8,9]

Some limitations of existing studies on ED are that the numbers of samples are relatively small, done in a limited scope, and do not get much attention from the world of medicine, especially in the field of anaesthesia. In addition, There are many multifactorial aetiologies of ED and no gold standard for establishing ED diagnosis.[4]

This article was accepted: 22 February 2024
Corresponding Author: Djayanti Sari
Email: jayantisari@ugm.ac.id

Many scores for assessing ED are available, usually having three to five categories. Three most widely used scales are the Pediatric Anesthesia Emergence Delirium Scale (PAEDs), Watcha and Cravero. Each diagnostic tool available has advantages, disadvantages and methods.[1,2,10] However, scoring with PAEDs also has some limitations; some items have objective criteria, the responses are not well defined, and there is a possibility of false-positive values, even though each item has been validated from previous studies. Some items in PAED are not specific to ED, and there is no consensus on whether cut-off values should be included for ED.[1]

The existing problems, such as limited ED diagnosis and scoring, a large number of cases, variety of complications, can affect the safety and quality of surgery/anaesthesia services. Meanwhile, ED treatment and prevention can be done without difficulty by recognising risk factors, diagnosing ED, and providing the right drugs. Therefore, it is necessary to conduct a study to determine the incidence of ED, identify the risk factors, and external validation of the ED scoring system such as PAED, Watcha, and Cravero score systems as predictors of ED.

MATERIALS AND METHODS

This research was approved by the Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada research ethics commission. The approval number is KE/KF/0024/EC.

Research Design

This study uses a prospective cohort study design. The aim is to carry out external validation of PAED, Watcha and Cravero as predictors of ED in patients at Dr. Sardjito General Hospital. Sampling data is carried out using a non-probability sampling method in a consecutive manner where all subjects who meet the criteria are recruited into the study sequentially.[11] The inclusion criteria are: (1) Patients undergoing elective surgery, (2) paediatric age between 2 to 10 years and (3) ASA I-II. Exclusion criteria in this study are: (1) Diffable/communication interference, (2) history of neurological disorders and (3) history of cardiovascular disorders. Samples were taken prospectively and sequentially for two months after receiving ethical approval. With a data collection period of 2 months, the number of samples obtained is about 79 subjects.

Variable Dependent and Independent

The dependent variable in this study is the incidence of ED, assessed by experienced paediatric anaesthesiologists. There were two paediatric anaesthesiologists involved in the study. There were different thresholds for ED. The PAED score was >12, the Watcha score was 3-4 and the Cravero score was 4-5. The output variable in this study on the categorical scale is ED (yes or no). The independent variable is the variable that exists on the PAED, Watcha and Cravero scores (Supplementary material 1, 2 and 3).[1,4]

Procedure

Patients' age, body weight, gender, premedication, type of surgery and intraoperative anaesthesia and postoperative analgesia were also recorded. All children were observed by two experienced paediatric anaesthesiologists in a video recording. Observations are recorded at three-time points: (1) at the patient's initial arrival at the PACU, (2) the worst score for the initial 10 minutes, and (3) the worst score for the next 10 minutes by another researcher. At the same time, paediatric anaesthesiologists will observe whether an ED happened to the patient and decide on the therapy administration according to the protocol.

Data Analysis

To assess the inter-observer agreement, we used the kappa test. The kappa score between 0.81 to 0.99 shows almost perfect agreement, 0.61 to 0.80 shows good agreement, 0.41 to 0.60 shows moderate agreement, 0.21 to 0.40 shows fair agreement, and 0.01 to 0.20 shows slight agreement.12 Data analysis in this study focuses on the ability of discrimination and calibration. Discrimination ability was evaluated with a receiver operating characteristic (ROC) and the area under the curve (AUC).11 AUC values ranged from 0 to 1. The wider the AUC, the better the scoring ability to detect ED events. When an AUC area is found between 0.9 to 1, the scoring ability to detect ED could be interpreted as remarkable. Meanwhile, the scoring ability is good when the AUC area is found between 0.8 to 0.9. A mediocre ability is when an AUC is found between 0.7-0.8, and a poor ability is when the area found is between 0.6 to 0.7. While if the area is between 0.5 to 0.6, then the scoring is considered failed to detect ED.[11] The univariate were determined using independent t-test (if numerical) or chi-square (if categorical). After that, if $p < 0.05$, we continue to multivariate analysis to determine the risk factor ($p < 0.05$).

RESULTS

Before conducting the ED assessment, socialisation and training were done. The assessment is done through video recording and is assessed separately. The inter-observer agreement test shows a result of 0.76, which means that the agreement between the assessors is good. During the data collection period, 79 data were obtained. Demographic baseline data and those associated with ED incidence assessed using PAED can be seen in Table I and Table II. The mean age of the subjects in this study was 5.19 ± 2.6 years, body mass index (BMI) 15.7 ± 3.9 and surgery duration 113.4 ± 84.8 minutes. In general, research subjects were primarily male, ASA 2, underwent non-major surgery type, calm when separated from their parents, general anaesthesia techniques with endotracheal intubation, premedication with more than one drug, induction with a non-inhalant agent (intravenous) anaesthesia, maintenance via inhalation, without using muscle relaxants, without regional anaesthesia, use of analgesics during and post-surgery other than fentanyl.

According to the PAED score, the mean age was 5.05 ± 2.9 years for the subjects with ED, slightly younger than those without ED. In addition, the average BMI was 15.8 ± 4.8, and the operating time was 123.5 ± 97.7 minutes. In subjects who experienced ED, it was more common in female, ASA 2 physical status, non-major surgery, calm when separated from their parent, general anaesthesia intubation techniques, given more than one drug of premedication,

Table I: Baseline demographic data

Variables	Total
Age (years), mean±SD	5.19 ± 2.6
Body mass index (BMI), mean±SD	15.7 ± 3.9
Gender	
• Male, n (%)	46 (58.2%)
• Female, n (%)	33 (41.8%)
American Society of Anesthesiologists (ASA)	
• I, n (%)	38 (48.1%)
• II, n (%)	41 (51.9%)
Type of operation	
• Major, n (%)	15 (19.5%)
• Non-Major, n (%)	62 (80.5%)
Operation duration, mean±SD	113.4 ± 84.8
Child's behaviour on parental separation	
• Calm, n (%)	58 (74.4%)
• Restless, n (%)	20 (25.6%)
Anaesthesia technique	
• Total intravenous anaesthesia (TIVA), n (%)	14 (17.7%)
• Laryngeal mask airway (LMA), n (%)	26 (32.9%)
• Intubation, n (%)	39 (49.4%)
Premedication:	
• One drug, n (%)	22 (27.8%)
• More than one drug, n (%)	57 (72.2%)
Induction:	
• Inhalation, n (%)	19 (24.1%)
• Non-inhalation, n (%)	60 (75.9%)
Use of muscle relaxants	
• Yes, n (%)	37 (46.8%)
• No, n (%)	42 (53.2%)
Maintenance:	
• Inhalation, n (%)	77 (97.5%)
• Non-inhalation, n (%)	2 (2.5%)
Use of regional anaesthesia	
• Yes, n (%)	7 (8.9%)
• No, n (%)	72 (91.1%)
Analgesic uses during surgery:	
• Fentanyl, n (%)	9 (11.4%)
• Non-fentanyl, n (%)	70 (88.6%)
Postoperative analgesic usage:	
• Fentanyl, n (%)	5 (8.8%)
• Non-fentanyl, n (%)	52 (91.2%)

induced with non-inhalation, without using muscle relaxant, maintenance anaesthesia with inhalation, without using regional anaesthesia techniques, using non-fentanyl analgesics during and postoperatively.

In this study, none of the patients had communication, neurological, cardiological and eye contact disorders, nor were there any patients with decreased consciousness. All operations are elective operations. All variables studied did not make a significant difference to the incidence of ED.

The number of subjects who experienced ED as measured by PAED, Watcha, Cravero score, and expert assessment can be seen in Table III. From Table III, the highest number of patients diagnosed with ED were based on the Watcha score with 63.3%, and the lowest was based on expert's assessment with 10.1%. Details of the number of patients diagnosed with ED (based on PAED >12) either by Watcha, Cravero scoring or expert assessment can be seen in Figure 1. Only two patients (11.1%) agreed positively on ED incidence using Watcha, Cravero scores and expert assessments.

A correlation test was carried out between the PAED score and Watcha, Cravero score and expert assessment before the ROC analysis. From Table III, only Watcha's scoring correlates with PAED scoring ($p < 0.05$).

The three scores analysed by the ROC curve, only Watcha and Cravero scores showed discrimination in the incidence of ED (Figure 2). The ROC curve from the expert assessment did not show discrimination incidence of ED ($p = 0.935$).

The size of the AUC area shows how much the accuracy of a test is. In this study, the accuracy of scoring and assessments made by experts (can be seen in Table IV). From Table IV, both Cravero and expert assessment's have AUC between 0.5 to 0.6, which means that only Watcha's score shows the highest discrimination ability compared to Cravero's and expert assessment, although only classified as mediocre.

From the data that have been collected, there were no variables with p-value <0.05, so we stopped at the univariate analysis.

Table II: The incidence of emergence delirium with Pediatric Anesthesia Emergence Delirium

Variables	With emergence delirium	Without emergence delirium	*p*-value
Age (years), mean±SD	5.05 ±2.9	5.6 ± 2.8	0.795
BMI, mean±SD	15.8 ± 4.8	15.4 ± 3.8	0.919
Gender			
• Male, n (%)	7 (38.9%)	39 (63.9%)	0.101
• Female, n (%)	11 (61.1%)	22 (36.1%)	
ASA			
• I, n (%)	6 (33.3%)	32 (52.5%)	0.186#
• II, n (%)	12 (66.7%)	29 (47.5%)	
Type of operation			
• Major, n (%)	4 (22.2%)	11 (18.6%)	0.741#
• Non-major, n (%)	14 (77.8%)	48 (81.4%)	
Operation duration, mean±SD	123.5 ± 97.7	110.8 ± 84.5	0.837#
Child's behaviour on parental separation			
• Calm, n (%)	11 (61.1%)	47 (78.3%)	0.216#
• Restless, n (%)	7 (38.9%)	13 (21.7%)	
Anaesthesia technique			
• TIVA, n (%)	2 (11.1%)	12 (19.7%)	0.493#
• LMA, n (%)	5 (27.8%)	21 (34.4%)	
• Intubation, n (%)	11 (61.1%)	28 (45.9%)	
Premedication:			
• One drug, n (%)	6 (33.3%)	16 (26.2%)	0.561#
• More than one drug, n (%)	12 (66.7%)	45 (73.8%)	
Induction:			
• Inhalation, n (%)	2 (11.1%)	17 (27.9%)	0.212#
• Non-inhalation, n (%)	16 (88.9%)	44 (72.1%)	
Use of muscle relaxants			
• Yes, n (%)	8 (44.4%)	29 (47.5%)	0.999#
• No, n (%)	10 (55.6%)	32 (52.5%)	
Maintenance:			
• Inhalation, n (%)	17 (94.4%)	60 (98.4%)	0.406#
• Non-inhalation, n (%)	1 (5.6%)	1 (1.6%)	
Use of regional anaesthesia			
• Yes, n (%)	2 (11.1%)	5 (8.2%)	0.655#
• No, n (%)	16 (88.9%)	56 (91.8%)	
Analgesic uses during surgery:			
• Fentanyl, n (%)	3 (16.7%)	6 (9.8%)	0.999#
• Non-fentanyl, n (%)	15 (83.3%)	55 (90.2%)	
Postoperative analgesic usage:			
• Fentanyl, n (%)	1 (11.1%)	4 (8.3%)	0.418#
• Non-fentanyl, n (%)	8 (88.9%)	44 (91.7%)	

Statistical analysis with t-test and chi-square#.

Table III: The correlation between Pediatric Anesthesia Emergence Delirium cores and Watcha, Cravero and expert assessment score

Scoring	Number of patients with ED (N total=79)	*p*-value
PAED >12, n (%)	18 (22.8%)	ref
Cravero >3, n (%)	32 (40.5%)	0.057
Watcha >2, n (%)	50 (63.3%)	0.002*
Expert assessment, n (%)	8 (10.1%)	0.99

*$p < 0.05$.

Table IV: Area under curve on all three scores in diagnosing ED (based on PAED >12)

Scoring	AUC	95% CI	*p*-value
Cravero >3	0.630	0.476-0.784	0.104
Watcha >2	0.741	0.615-0.867	0.002*
Expert assessment	0.503	0.353-0.660	0.935

*$p < 0.05$

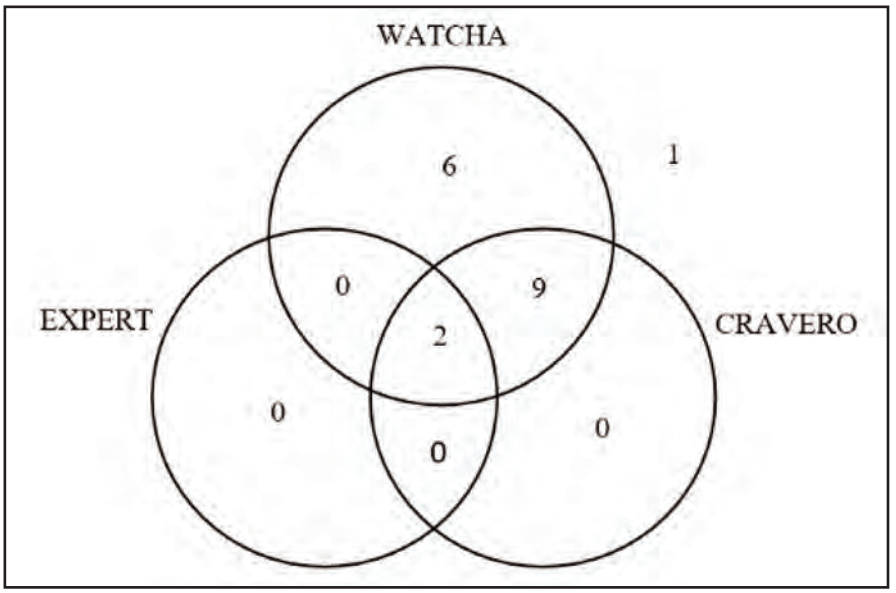

Fig. 1: Venn diagram between Cravero score, Watcha score and expert assessment in diagnosing emergence delirium (based on Pediatric Anesthesia Emergence Delirium >12)

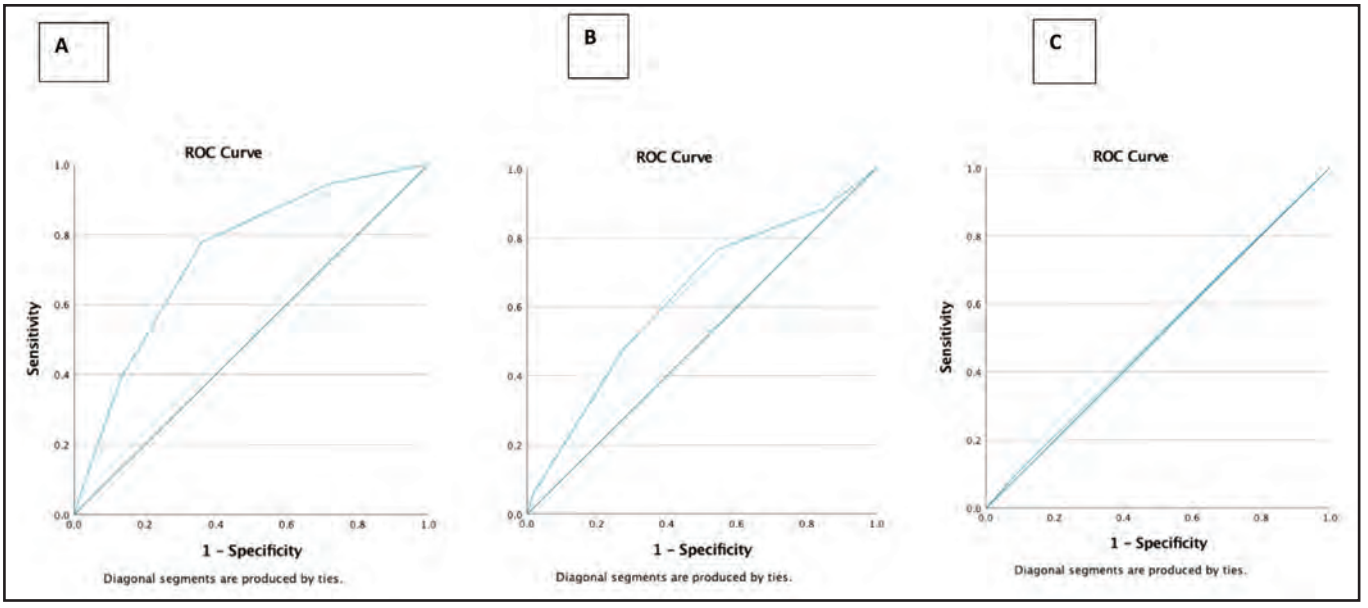

Fig. 2: Graph of receiver operating characteristic (ROC) on all three scores in diagnosing emergence delirium (based on PAED >12). A. Watcha, B. Cravero and C. Expert assessment

DISCUSSION

Post general anaesthesia ED is still considered a clinical problem for anaesthesiologists. Although there is no explicit agreement on the definition and measurement tool of ED, it is recognised by the patient's period of restlessness, agitation, inconsolable crying, disorientation, delusions, hallucinations, impaired cognitive function and memory.[13] In addition, there is no clear definition of ED and ED assessment instruments have various performance.

This study, which compares three scoring system, differs from the investigation conducted by Bajwa et al.[1] This study was focused on children aged 2 to 10 years, in contrast to previous studies involving children up to 18 years old. Additionally, the incorporation of expert assessment from two paediatric anaesthesiologist, rather than a single anaesthesiologist (as in the previous research) is expected to increase the validity of these findings.[1]

In assessing ED, PAED, Watcha and Cravero scores are widely used.[1] Based on existing knowledge and experience, the assessment carried out by paediatric anaesthesiologists is one way of detecting and measuring the incidence of ED and has been widely applied in diagnosing ED at Dr. Sardjito General Hospital. This study used PAED scoring as a standard. PAED score has been widely used because it has high sensitivity and specificity to detect ED incidence.[1] The PAED score with a cut-off >12 has a sensitivity of 100% and a specificity of 94.5%,[1,4]

so in this study, the PAED score is used as the standard to detect ED. Despite the high sensitivity and specificity, PAED scoring is not practical and fast to use.

The data obtained during this study were 79. One data was dropped out because the video recording did not exist. The data was taken by researchers who had been trained to use three ED assessment scores (PAED, Watcha and Cravero). An inter-observer agreement assessment was also carried out with good results. The evaluator assessed ED in the PACU room by all three scores without knowing the case.

As assessed by the PAED score, the incidence of ED is 22.8%, still within the range of ED incidence reported from various publications, around 20 to 80%.[1,2] In addition, the ED and non-ED groups were not significantly different, meaning that the two groups could be compared (Table II).

The incidence of ED was most diagnosed by Watcha scoring. The highest scoring able to detect for ED was Watcha scoring, followed by Cravero scoring, and lastly by expert assessment. Only two patients (11.1%) agreed positively on ED incidence using Watcha, Cravero scores and expert assessments. By looking at the numbers, this study illustrates that the agreement on ED assessment is still relatively low. However, the Watcha score being the only scoring correlated with the PAED score, has the highest discrimination compared to other scoring or expert assessments, even though the ability to detect ED incidents is classified as mediocre. From a previous study that compares PAED and Watcha, PAED has high sensitivity and specificity compared to Watcha. However, Watcha is easier to use in PACU.[1]

Many factors influence the incidence of ED, which makes the results of each study vary and differ from the results of this study. From previous studies, several risk factors that possibly can increase the incidence of ED are rapid emergence for anaesthesia, use of short-acting volatile aesthetic agents, postoperative pain, type of operation, age, preoperative anxiety, and child temperament.[14] However, we found that the variables we studied were unrelated to the incidence of ED.

Variations in the identification of risk factors between this study and previous studies could occur due to several reasons. One potential reason is the homogeneity of the subject ages, ranging from 2 to 10 years, which likely resulted in the age variable not demonstrating significant differences in this investigation. One of a study in Indonesia that studied ED in children aged 1 to 12 years showed that preoperative anxiety and pain were risk factors for ED.[15] In our study, preoperative anxiety as represented by child behaviour on parental separation was not associated with the incidence of ED. This may be because behavioural variables are assessed subjectively so this shows different results. Despite these differences, this study showed the same results that age, ASA physical status, type of surgery, and induction score were not the risk factors of ED.[15]

There were some limitations in our study. The scoring system used as gold standard for ED incidence in this study is the PAED scoring, with a cut-off value of more than 12. This is because there are study that reported the cut-off value of 12 have the highest sensitivity and specificity compared to another cut-off. But, one meta-analysis showed that the cut-off 10 or 12 was not significantly different.[16] In addition, one of the variables that examines preoperative anxiety was subjective, so we recommend to using scoring that has been proven to be valid and reliable. This study does not rule out other variables that may be significant but have not been studied as risk factors for ED.

CONCLUSION

All the parameters studied are not risk factors of emergence delirium (ED) incidence in paediatric patients. Only Watcha's score correlates with the Pediatric Anesthesia Emergence Delirium (PAED) score and shows the highest discrimination ability compared to Cravero's score and expert assessment, although the ability to discriminate is relatively good.

CONFLICT OF INTEREST

There is no conflict of interest.

REFERENCES

1. Bajwa SA, Costi D, Cyna AM. A comparison of emergence delirium scales following general anesthesia in children. Paediatr Anaesth 2010; 20(8): 704-11.
2. Vlajkovic GP, Sindjelic RP. Emergence delirium in children: Many questions, few answers. Anesth Analg 2007; 104(1): 84-91.
3. Dahmani S, Delivet H, Hilly J. Emergence delirium in children: An update. Curr Opin Anaesthesiol 2014; 27(3): 309-15.
4. Lee CA. Paediatric emergence delirium: an approach to diagnosis and management in the postanaesthesia care unit. J Perioper Crit Intensive Care Nurs 2018; 04(01): 2-4.
5. Street JT, Lenehan BJ, Dipaola CP, Boyd MD, Kwon BK, Paquette SJ, et al. Morbidity and mortality of major adult spinal surgery. A prospective cohort analysis of 942 consecutive patients. Spine J 2012; 12(1): 22-34.
6. Kain ZN, Caldwell-Andrews AA, Weinberg ME, Mayes LC, Wang SM, Gaal D, et al. Sevoflurane versus halothane: Postoperative maladaptive behavioral changes: a randomized, controlled trial. Anesthesiology 2005; 102(4): 720-6.
7. Kain ZN, Mayes LC, Caldwell-Andrews AA, Karas DE, mcclain BC. Preoperative anxiety, postoperative pain, and behavioral recovery in young children undergoing surgery. Pediatrics 2006; 118(2): 651-8.
8. Cravero J, Surgenor S, Whalen K. Emergence agitation in paediatric patients after sevoflurane anaesthesia and no surgery: a comparison with halothane. Paediatr Anaesth 2000; 10(4): 419-24.
9. Kanaya A. Emergence agitation in children: risk factors, prevention, and treatment. J Anesth 2016; 30(2): 261-7.
10. Wilson TA, Graves SA. Pediatric considerations in a general postanesthesia care unit. J Post Anesth Nurs 1990; 5(1): 16-24.
11. Dahlan M. 2011. Statistik Untuk Kedokteran dan Kesehatan. 5th Edition. Salemba Medika.
12. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005; 37(5): 360-3.
13. Sikich N, Lerman J. Development and Psychometric Evaluation of the Pediatric Anesthesia Emergence Delirium Scale. Anesthesiol 2004; 100(5): 1138-45.
14. Reduque LL, Verghese ST. Paediatric emergence delirium. Contin Educ Anaesthesia, Crit Care Pain 2013; 13(2): 39-41.
15. Andriyanto L, Utariani A, Hanindito E, Santoso KH, Hamzah H, Puspita EA. Incidence of emergence agitation in pediatric patient after general anesthesia. Folia Medica Indones 2019; 55(1): 25.
16. Russell PSS, Mammen PM, Shankar SR, Viswanathan SA, Rebekah G, Russell S, et al. Pediatric anesthesia emergence delirium scale: a diagnostic meta-analysis. World J Clin Pediatr 2022; 11(2): 196-205.