

# BMJ Open Risk prediction models for emergence delirium in paediatric general anaesthesia: a systematic review

Maria-Alexandra Petre,<sup>1</sup> Bibek Saha,<sup>2</sup> Shugo Kasuya,<sup>3</sup> Marina Englesakis,<sup>4</sup> Nan Gai,<sup>5</sup> Arie Peliowski,<sup>5</sup> Kazuyoshi Aoyama <sup>5,6</sup>

**To cite:** Petre M-A, Saha B, Kasuya S, *et al.* Risk prediction models for emergence delirium in paediatric general anaesthesia: a systematic review. *BMJ Open* 2021;**11**:e043968. doi:10.1136/bmjopen-2020-043968

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-043968>).

Received 18 August 2020  
Revised 06 December 2020  
Accepted 08 December 2020

## ABSTRACT

**Objectives** Emergence delirium (ED) occurs in approximately 25% of paediatric general anaesthetics and has significant adverse effects. The goal of the current systematic review was to identify the existing literature investigating performance of predictive models for the development of paediatric ED following general anaesthesia and to determine their usability.

**Design** Systematic review using the Prediction model study Risk Of Bias Assessment Tool (PROBAST) framework.

**Data sources** Medline (Ovid), PubMed, Embase (Ovid), Cochrane Database of Systematic Reviews (Ovid), Cochrane CENTRAL (Ovid), PsycINFO (Ovid), Scopus (Elsevier) and Web of Science (Clarivate Analytics), ClinicalTrials.gov, International Clinical Trials Registry Platform and ProQuest Digital Dissertations and Theses International through 17 November 2020.

**Eligibility criteria for selecting studies** All randomised controlled trials and cohort studies investigating predictive models for the development of ED in children undergoing general anaesthesia.

**Data extraction and synthesis** Following title, abstract and full-text screening by two reviewers, data were extracted from all eligible studies, including demographic parameters, details of anaesthetics and performance characteristics of the predictive scores for ED. Evidence quality and predictive score usability were assessed according to the PROBAST framework.

**Results** The current systematic review yielded 9242 abstracts, of which only one study detailing the development and validation of the Emergence Agitation Risk Scale (EARS) met the inclusion criteria. EARS had good discrimination with c-index of 0.81 (95% CI 0.72 to 0.89). Calibration showed a non-significant Homer-Lemeshow goodness-of-fit test ( $p=0.97$ ). Although the EARS demonstrated low concern of applicability, the high risk of bias compromised the overall usability of this model.

**Conclusions** The current systematic review concluded that EARS has good discrimination performance but low usability to predict ED in a paediatric population. Further research is warranted to develop novel models for the prediction of ED in paediatric anaesthesia.

**PROSPERO registration number** CRD42019141950.

## INTRODUCTION

Emergence delirium (ED) is a common complication occurring in approximately 25% of paediatric general anaesthetics (GA)<sup>1</sup> and is associated with significant adverse effects including injury to the patient and

## Strengths and limitations of this study

- This is a first systematic review of risk prediction models for paediatric emergence delirium and adheres to recommendations made in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.
- An extensive systematic literature search was conducted through 11 databases including Medline (Ovid), Embase (Ovid) and Web of Science (Clarivate Analytics) from their inception to 17 November 2020.
- The current systematic review employed the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies for data extraction and the Prediction model study Risk of Bias Assessment Tool which determines the usability of the prediction models identified in the systematic review.
- Our systematic review revealed only one study reporting the performance characteristics of the Emergence Agitation Risk Scale, a predictive model with low usability for predicting the development of paediatric emergence delirium.
- Our results serve to highlight the need for the ongoing development and validation of robust predictive models for paediatric emergence delirium risk.

personnel, damage to incision sites, exacerbated parental anxiety and increased nursing requirements, further resulting in an increased burden on the healthcare system.<sup>2,3</sup> Identifying which patients are at highest risk for developing ED will allow practitioners to effectively apply multimodal prophylaxis in order to decrease the incidence of this significant complication.

Several risk factors have been identified for the development of ED, including age, preoperative anxiety, type of surgery and type of anaesthetic given.<sup>2,4</sup> Development of predictive scores which aim to integrate these risk factors to determine an individual's overall risk for developing ED has been attempted. However, no prior systematic review has been



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

### Correspondence to

Dr Kazuyoshi Aoyama;  
kazu.aoyama@utoronto.ca

conducted to determine the usability of these prediction models.

The aim of the current study is to conduct a systematic review to identify all existing prediction models for the development of ED in a paediatric population, assess model performance and determine the usability of these models for use in clinical practice.

## METHODS

The results were reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols (PRISMA) 2015 statement.<sup>5</sup>

### Search strategy

An information specialist with experience in systematic reviews searched eight databases (Medline (Ovid), PubMed, Embase (Ovid), Cochrane Database of Systematic Reviews (Ovid), Cochrane CENTRAL (Ovid), PsycINFO (Ovid), Scopus (Elsevier) and Web of Science (Clarivate Analytics) starting from their inception. Furthermore, three databases were searched for relevant recently completed or ongoing research (ClinicalTrials.gov, International Clinical Trials Registry Platform and ProQuest Digital Dissertations and Theses International). All searches were conducted on 17 May 2019 and updated on 17 November 2020. Search strategies were built to contain three sets of terms reflecting our search questions including the prediction models, the target condition (ED, emergence agitation (EA)) and the patient population (paediatric patients undergoing GA). Since ED and EA are sometimes used interchangeably in older literature, both these terms were included in our search strategy. In addition, reference lists of relevant trials and reviews were scanned. Refer to online supplemental appendix 1 for search strategies for all databases.

### Study selection and data extraction

Title, abstract and full-text screening were independently conducted by two reviewers (SK and BS) with conflicts resolved by a third reviewer (M-AP). Cohen's kappa was calculated to quantify the inter-rater reliability.<sup>6</sup>

Inclusion criteria were: randomised controlled trials (RCTs), cohort studies or case-control studies examining paediatric populations (<18 years old) undergoing GA investigating preoperative or intraoperative predictive models for the development of ED or EA in the postanaesthetic care unit (PACU). Inclusion and exclusion criteria are detailed in the online supplemental appendix 2.

Study characteristics as well as primary and secondary outcomes data were extracted from relevant studies using a standardised data collection form in accordance to the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) framework<sup>7</sup> and the Transparent Reporting of a multi-variable prediction model for Individual Prognosis or Diagnosis (TRIPOD) checklist.<sup>8</sup> These included demographic information, anaesthetic characteristics (type

and dosages of anaesthetic used for induction and maintenance, use and dosage of premedication); type of scale used (components, model development methodology, model evaluation methodology, time span of prediction, and intended moment for the use of the model), discrimination characteristics (area under receiver operating curve (AUC) or corresponding c-statistic with 95% CI) and calibration characteristics (calibration plots, ratio between predicted vs observed incidence of ED, or Hosmer-Lemeshow goodness-of-fit statistic).

### Assessment of methodological quality

The methodological quality of the evidence was determined using the Prediction model study Risk Of Bias Assessment Tool (PROBAST) framework.<sup>9</sup> This consists of first determining the usability of the risk prediction model based on the risk of bias and concerns of applicability across four domains (participants, predictors, outcome and analysis), followed by a determination of the model's predictive performance (discrimination and calibration).<sup>9</sup> The model was considered to be 'usable' if it had a low risk of bias, low concern about applicability, and good predictive performance of discrimination and calibration. Good discrimination is defined as AUC  $\geq 0.8$ .<sup>10 11</sup>

### Software

References were collected and deduplicated using Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia).

### Patient and public involvement

It was not appropriate, possible or necessary to involve patients or the public in the design or conduct of this study, or to disseminate the results to them.

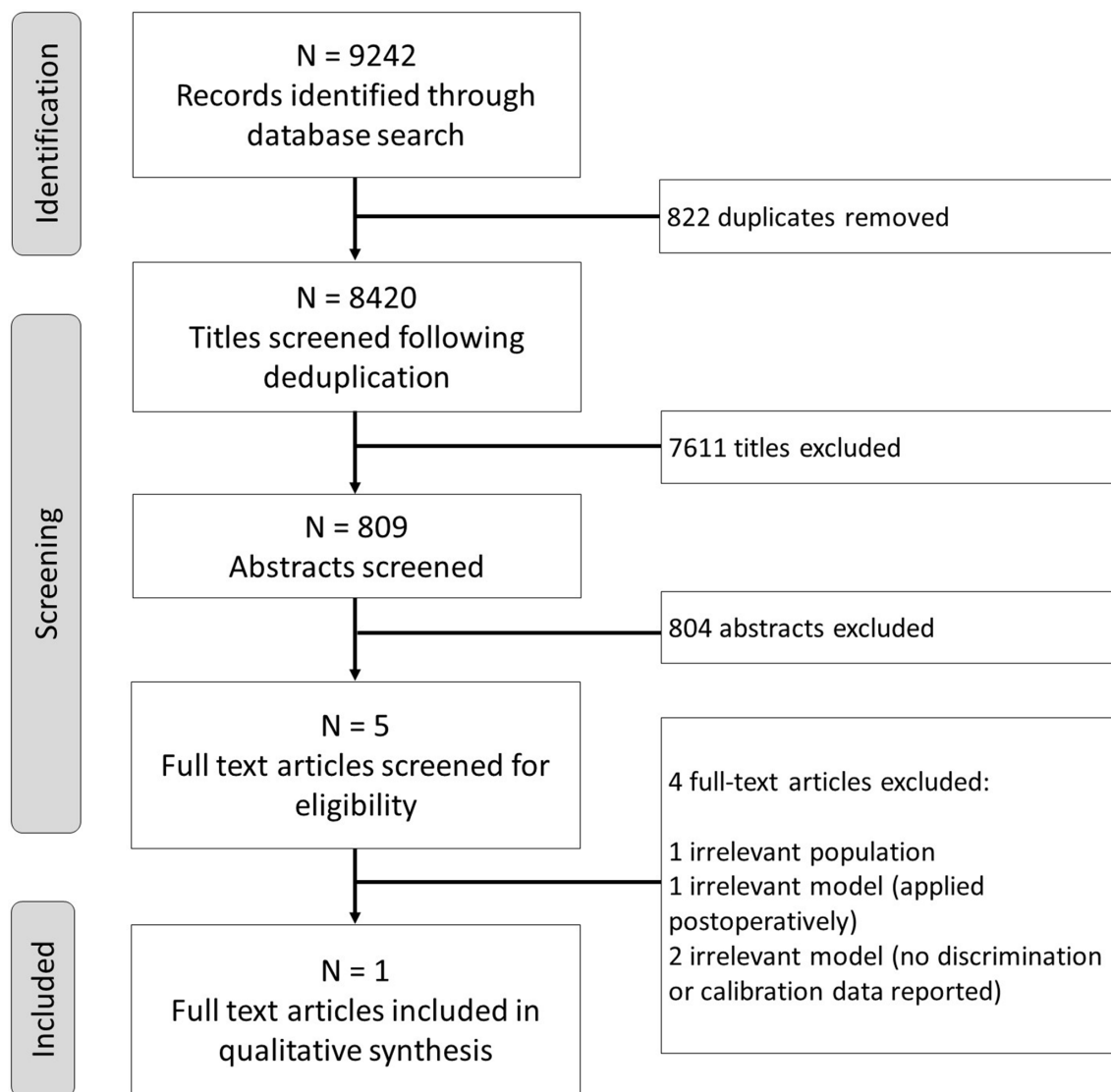
## RESULTS

### Search results

The literature search yielded 9242 citations, which, following title, abstract and full-text screening, yielded one study that met the inclusion criteria. The PRISMA flow diagram is shown in [figure 1](#). Four full-text articles were excluded due to irrelevant populations,<sup>12</sup> model being applied post-operatively<sup>13</sup> and lack of reported discrimination or calibration data<sup>4 14</sup> ([table 1](#)). The eligible paper investigated the development and validation of the Emergence Agitation Risk Scale (EARS).<sup>15</sup> Cohen's kappa showed good agreement between the two reviewers (kappa=0.99).

### Emergence Agitation Risk Scale

The included study, published in 2017, was a Japanese, single-centre, paediatric hospital-based study detailing the development and validation of the EARS. It comprised a total of 220 ASA class I or II patients with mean ages  $4.1 \pm 1.8$  years and undergoing minor surgery including tonsillectomy, adenoidectomy, myringotomy tube insertion, strabismus surgery, cryptorchidism repair



**Figure 1** PRISMA flow diagram for search and review strategy. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

or inguinal hernia repair under GA. Anaesthetic induction consisted of sevoflurane and nitrous oxide inhalation without premedication while maintenance was achieved with sevoflurane. Analgesia consisted of intravenous fentanyl, acetaminophen suppository and nerve blocks wherever indicated. The outcome of interest, EA, was measured in the PACU using the Paediatric Anaesthesia Emergence Delirium (PAED) scale with a cut-off of >12 and had an overall incidence of 36.4%.<sup>16</sup> Study characteristics are detailed in [table 2](#).

Development of EARS was conducted retrospectively in a cohort of 120 patients previously enrolled in an RCT investigating the use of acupuncture in the prevention of EA in children.<sup>17</sup> Logistic regression was used to test the ten candidate predictors including age, height, weight, sex, Paediatric Anaesthesia Behaviour (PAB) score, operative procedure, anaesthesia time, airway securing device (endotracheal tube or laryngeal mask airway), presence or absence of nerve block use, and total fentanyl dose. The Akaike information criterion stepwise selection was

used to identify four independently associated predictors for final inclusion in the EARS (age, PAB score, anaesthesia time and operative procedure).  $\beta$ -Coefficients were calculated and converted to integer scores for each predictor, yielding a score range of 1–23.

The validation phase of the score was conducted separately in a prospective observational cohort of 100 patients.

### Study findings

The development phase study population had mean age  $3.7 \pm 1.7$  years and incidence of EA of 34.2%. The c-statistic was 0.84 (95% CI 0.74 to 0.94) and Hosmer-Lemeshow statistic was non-significant ( $p=0.97$ ), indicating adequate discrimination and calibration, respectively.

The validation phase study population had mean age  $4.5 \pm 1.9$  years and EA incidence of 39%. The c-statistic for the validation phase was 0.81 (95% CI 0.72 to 0.89). The optimum cut-off point was found to be 11, yielding 87% sensitivity and 61% specificity for the development of ED.

**Table 1** Detailed explanations for studies excluded following full-text assessment

Study reference	Predictive model investigated	Population	Reason for exclusion
Beringer <i>et al</i> , 2014 <sup>14</sup>	Paediatric Anaesthesia Behaviour scale	Children aged 2–12 years undergoing general anaesthesia for dental extractions	No discrimination or calibration parameters were reported
Kain <i>et al</i> , 2004 <sup>4</sup>	Yale Preoperative Anxiety Scale	Children undergoing surgery with general anaesthesia	No discrimination or calibration parameters were reported
Oh <i>et al</i> , 2013 <sup>12</sup>	Automatic PREdiction of DELirium in Intensive Care Units (ICU)	Delirium in adult ICU setting	Irrelevant population
Sadhasivam <i>et al</i> , 2009 <sup>13</sup>	Perioperative Adult Child Behavioural Interaction Scale (PACBIS)	Children aged 3–12 years undergoing general anaesthesia for tonsillectomy and/or adenoidectomy	PACBIS measured in the recovery room concurrently with PAED score No discrimination or calibration parameters were reported

PAED, Paediatric Anaesthesia Emergence Delirium.

The grey zone, delimited by the points of the scale where the sensitivity and specificity become 90%, was 10 to 13 and comprised 38% of patients. Calibration data was not reported for the validation phase.

### Score usability

Usability analysis based on PROBAST criteria, revealed low concern of applicability across all three domains probed (participants, predictors, outcomes). However, there was a high risk of bias, primarily due to deficits in the analysis domain. In particular, the number of events (n=41) per candidate predictor (n=10) during the development phase was 4.1 (at least 10 and preferably 20 would be considered adequate). Furthermore, during the validation phase the population size was inadequate, yielding only 39 patients experiencing the primary outcome, which fell short of the recommended 100 patients experiencing the outcome of interest. The authors did not present a calibration plot to illustrate the goodness-of-fit associated with the statistically significant Hosmer-Lemeshow statistic. And lastly, the authors did not account for overfitting and optimism in the context of the small events per candidate predictor (online supplemental table S1).

Taken together, given the good discrimination performance but insufficient calibration evaluation, as well as the high risk of bias despite a low concern of applicability of the model to paediatric populations for predicting EA, the EARS was given a designation of low ‘usability’ (table 3).

## DISCUSSION

The current systematic review revealed only one scale targeted at predicting the risk of developing ED in children, a complication associated with significant morbidity occurring in approximately 25% of paediatric GA. The PROBAST assessment of usability indicated that the discriminative

performance and applicability of the scale were both good, but insufficient calibration evaluation and deficiencies in the analysis placed the scale at high risk of bias, reducing its overall usability in clinical practice.

To date, this has been the only systematic review looking at prediction scales of ED in a paediatric population. The review was extensive, including eight major databases of peer-reviewed literature as well as 3 databases of protocols, with the search including both ED as well as EA, which have been interchangeably used in previous literature. The study protocol was published in the PROSPERO database before the start of the review. The review process used rigorous methodology, employing two independent reviewers and careful adherence to the PRISMA,<sup>5</sup> CHARMS,<sup>7</sup> TRIPOD<sup>8</sup> and PROBAST<sup>9</sup> guidelines for systematic reviews and prediction scales. While the scale was found to have low usability, the current study highlights potential required improvements in the methodology of future studies validating the EARS, with problems arising from high risk of bias due to the analysis phase. Particular attention should be given to the number of events per variable to avoid the risk of overfitting and underfitting of the model.

A considerable limitation of the current review was that only one study investigating a single predictive model was found, which met all the inclusion criteria. While this hampered our ability to run a meta-analysis on the predictive properties of the scale, a need for further validation in other settings is underscored, as is the need for the development of other scales in this domain. Literature in this field may have been previously lacking due to considerable variations in the definition of ED, with different diagnostic scales being used, although the PAED scale is the only validated scale to diagnose ED in children at the moment.<sup>16</sup> Furthermore, several confounding factors are present in the postoperative period, which makes timely and accurate diagnosis of paediatric ED difficult in clinical and research settings.<sup>1–3</sup> As such, strategic

**Table 2** Summary characteristics of studies included in systematic review

Study	Score name and composition	Population	Inclusion and exclusion criteria	Anaesthetic type	Outcome definition	Model development and validation	Model performance
15	Emergence Agitation Risk Scale (EARS): 1. Age 2. Operative procedure (strabismus surgery, tonsillectomy) 3. Paediatric Anaesthesia Behaviour scale 4. Anaesthesia time	n=220 (120 during development and 100 during validation phases)	<u>Inclusion criteria:</u> 1.5–8 years old; ASA 1–2; undergoing general anaesthesia for inguinal hernia repair, adenoidectomy, tonsillectomy, strabismus repair, tympanostomy tube, cryptorchidism, and other minor surgery <u>Exclusion criteria:</u> mental retardation; taking psychotropic drugs	Induction with N <sub>2</sub> O and sevoflurane, maintenance with sevoflurane Analgesia with intravenous fentanyl and suppository acetaminophen, and nerve block where appropriate	PAED score >12 as assessed by trained nurse or anaesthesiologist in PACU Incidence: 34.2% of patients in the development phase 39% of patients in the validation phase	<u>Development phase:</u> 1. Data from prior randomised controlled trial 2. Multicollinearity between 10 candidate predictors: age, height, weight, sex, operative procedures (tonsillectomy, strabismus surgery, other), airway management (ETT, LMA), nerve block, fentanyl dose, anaesthesia time, PAB score 3. Multivariate logistic regression with Akaike information criterion stepwise selection conducted to determine the optimal combination of 7 remaining predictors (age, sex, operative procedure, airway management, anaesthesia time, PAB score) 4. Beta-coefficient was calculated for each predictor <u>Validation phase:</u> 1. Prospective cohort data. 2. Best cut-off point determined by Youden index 3. Grey zone identified as range between the two points on receiver operating curve where sensitivity and specificity are 90%	<u>Development phase:</u> Discrimination: C-index 0.84 (95% CI, 0.74 to 0.94) Calibration: Hosmer-Lemeshow goodness-of-fit test non-significant (p=0.97) <u>Validation phase:</u> Discrimination: C-index 0.81 (95%CI 0.72-0.89)

ASA, American Society of Anesthesiologists; EARS, Emergence Agitation Risk Scale; ETT, endotracheal tube; LMA, laryngeal mask airway; PAB, Paediatric Anaesthesia Behaviour Scale; PACU, postanesthesia care unit; PAED, Paediatric Anaesthesia Emergence Delirium scale.

**Table 3** Prediction model usability assessment

Study	Predictive performance		Risk of bias	Applicability	Usability
	Discrimination	Calibration			
Hino <i>et al</i> , 2017 <sup>15</sup> EARS model	<u>Development:</u> C-index 0.84 (95% CI 0.74 to 0.94) <u>Validation:</u> C-index 0.81 (95% CI 0.72 to 0.89)	<u>Development:</u> Hosmer-Lemeshow p=0.97 <u>Validation:</u> Not reported	Overall high concern	Overall low concern	Low

EARS, Emergence Agitation Risk Scale.

prevention of paediatric ED based on a precise and validated prediction scale is preferable to treatment when the event occurs.

The current systematic review highlights several potential clinical and research implications. The first is the need to identify validated risk factors for ED based on prior research, which can then be used to generate new prediction models using prospective cohort methodology to accurately and precisely identify those patients at the highest risk for developing this side effect.

Important lessons can be drawn from the three studies excluded in our systematic review due to a lack of reported discrimination and calibration parameters. Discrimination and calibration characteristics are essential features in the assessment of the usability of a predictive scale. Discrimination values such as the AUC or the C-statistic indicate how well a model differentiates those patients who are likely to develop a condition in comparison to those who are not likely to develop a condition. Calibration, on the other hand, as measured by a visual representation of the relationship between observed and predicted values and by the Hosmer-Lemeshow statistic, is a measure of the scale's ability to predict the absolute risk of developing the endpoint in question. The results of this study indicate that future scale development and validation should include a systematic assessment of the predictive properties of such scales. Indeed, following the analytical guidance provided by the PROBAST guidelines would ensure the analytic rigour required to incorporate these scores into clinical practice.

Lastly, future research, quality improvement projects and cost-benefit analyses should focus on determining whether implementing such scales into clinical practice to help target prophylaxis and treatment of ED is adequate and acceptable to healthcare providers and whether it ultimately results in improved patient outcomes.

In conclusion, the current systematic review has revealed a single study of predictive scores for the development of paediatric ED. While this score showed low usability, our results do highlight the need for the development of more such scales as well as the requirement for methodologically rigorous validation of such scales.

#### Author affiliations

<sup>1</sup>Department of Pediatric Anesthesia, Montreal Children's Hospital, McGill University, Montreal, Quebec, Canada

<sup>2</sup>John A Burns School of Medicine at University of Hawai'i at Mānoa, Honolulu, Hawaii, USA

<sup>3</sup>Department of Anesthesia and Critical Care Medicine, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

<sup>4</sup>Library and Information Services, University Health Network, Toronto, Ontario, Canada

<sup>5</sup>Department of Anesthesia and Pain Medicine, Hospital of Sick Children, University of Toronto, Toronto, Ontario, Canada

<sup>6</sup>Child Health Evaluative Sciences, SickKids Research Institute, Toronto, Ontario, Canada

**Correction notice** This article has been corrected since it first published. The provenance and peer review statement has been included.

**Twitter** Nan Gai @nanesthetist

**Contributors** M-AP, BS, SK, ME and KA: study concept, protocol writing and registration, analysis and interpretation of the data, drafting of manuscript, manuscript revision, approval of final version. NG and AP: study concept, protocol writing, manuscript revision, approval of final version.

**Funding** This work was supported by Outcomes Research Award, Department of Anesthesia and Pain Medicine, Hospital for Sick Children 2018–2019, and 2019 Perioperative Services Summer Studentship Program and Facilitator Grants Program, Hospital for Sick Children.

**Disclaimer** The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

**Competing interests** All authors have completed the ICMJE uniform disclosure form at [http://www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years, no other relationships or activities that could appear to have influenced the submitted work.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data of the current study is present in the main manuscript, figures, tables and online supplemental material.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Kazuyoshi Aoyama <http://orcid.org/0000-0002-7502-0896>

## REFERENCES

- 1 Ortiz AC, Atallah AN, Matos D. Intravenous versus inhalational anaesthesia for paediatric outpatient surgery. *Cochrane database Syst Rev* 2014;2:CD009015.
- 2 Voepel-Lewis T, Malviya S, Tait AR. A prospective cohort study of emergence agitation in the pediatric postanesthesia care unit. *Anesthesia & Analgesia* 2003;96:1625–30.
- 3 Vljakovic GP, Sindjelic RP. Emergence delirium in children: many questions, few answers. *Anesthesia & Analgesia* 2007;104:84–91.
- 4 Kain ZN, Caldwell-Andrews AA, Maranets I, et al. Preoperative anxiety and emergence delirium and postoperative maladaptive behaviors. *Anesthesia & Analgesia* 2004;99:1648–54.
- 5 Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
- 6 O'Connor D, Green S, Higgins JP. Defining the review question and developing criteria for including studies. In: *Cochrane Handbook for systematic reviews of interventions Cochrane book series*, 2008: 81–94.
- 7 Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.
- 8 Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Eur Urol* 2015;67:1142–51.
- 9 Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.
- 10 Aoyama K, D'Souza R, Pinto R, et al. Risk prediction models for maternal mortality: a systematic review and meta-analysis. *PLoS One* 2018;13:e0208563.
- 11 Hosmer DW, Lemeshow S, Sturdivant RX. Assessing the Fit of the Model. In: *Applied Logistic Regression. 3rd Edition*. Hoboken, NJ, USA: John Wiley & Sons, 2013: 153–225.
- 12 Oh S-H, Park E-J, Jin Y, et al. Automatic delirium prediction system in a Korean surgical intensive care unit. *Nurs Crit Care* 2014;19:281–91.
- 13 Sadhasivam S, Cohen LL, Szabova A, et al. Real-Time assessment of perioperative behaviors and prediction of perioperative outcomes. *Anesth Analg* 2009;108:822–6.
- 14 Beringer RM, Greenwood R, Kilpatrick N. Development and validation of the Pediatric Anesthesia Behavior score--an objective measure of behavior during induction of anesthesia. *Paediatr Anaesth* 2014;24:196–200.
- 15 Hino M, Mihara T, Miyazaki S, et al. Development and validation of a risk scale for emergence agitation after general anesthesia in children. *Anesthesia & Analgesia* 2017;125:550–5.
- 16 Sikich N, Lerman J. Development and psychometric evaluation of the pediatric anesthesia emergence delirium scale. *Anesthesiology* 2004;100:1138–45.
- 17 Hijikata T, Mihara T, Nakamura N, et al. Electrical stimulation of the heart 7 acupuncture site for preventing emergence agitation in children. *Eur J Anaesthesiol* 2016;33:535–42.