



OPEN

Continuous visualization and validation of pain in critically ill patients using artificial intelligence: a retrospective observational study

Naoya Kobayashi¹✉, Kazuki Watanabe², Hitoshi Murakami² & Masanori Yamauchi¹

Machine learning tools have demonstrated viability in visualizing pain accurately using vital sign data; however, it remains uncertain whether incorporating individual patient baselines could enhance accuracy. This study aimed to investigate improving the accuracy by incorporating deviations from baseline patient vital signs and the concurrence of the predicted artificial intelligence values with the probability of critical care pain observation tool (CPOT) ≥ 3 after fentanyl administration. The study included adult patients in intensive care who underwent multiple pain-related assessments. We employed a random forest model, utilizing arterial pressure, heart rate, respiratory rate, gender, age, and Richmond Agitation–Sedation Scale score as explanatory variables. Pain was measured as the probability of CPOT scores of ≥ 3 , and subsequently adjusted based on each patient's baseline. The study included 10,299 patients with 117,190 CPOT assessments. Of these, 3.3% had CPOT scores of ≥ 3 . The random forest model demonstrated strong accuracy with an area under the receiver operating characteristic curve of 0.903. Patients treated with fentanyl were grouped based on CPOT score improvement. Those with ≥ 1 -h of improvement after fentanyl administration had a significantly lower pain index ($P = 0.020$). Therefore, incorporating deviations from baseline patient vital signs improved the accuracy of pain visualization using machine learning techniques.

Many critically ill patients in the intensive care unit (ICU) experience intense pain regardless of whether they have undergone surgery^{1–5}. Pain has a wide variety of adverse effects on patients, including psychological stress, sleep disturbances, decreased respiratory function, increased heart rate and blood pressure, arrhythmia, and poor nutritional status that can lead to prolonged hospital stays, increased treatment costs, and poor life outcomes^{6–9}.

Although patient self-reporting is considered the most accurate and reliable way to assess pain, many critically ill patients are unable to self-evaluate due to intubation, tracheotomy, analgesic sedatives, or delirium; hence, alternative methods of assessing pain are necessary. Assessment scales, such as the Behavioral Pain Scale (BPS)¹⁰ and Critical Care Pain Observation Tool (CPOT, Supplementary Table S1)¹¹, which are used^{12–16} as alternatives, are easy for observers to use and allow for some standardization of assessments; however, they are human evaluations, which may lead to inter-rater differences, albeit acceptable. Nevertheless, such differences may be eliminated if the evaluation is performed by a machine. Furthermore, since only intermittent assessments are possible, pain treatment may be delayed.

Machine learning analysis of vital sign data can overcome the shortcomings of conventional pain assessment methods and predict pain at a given time with extremely high accuracy¹⁷. The comprehensive analysis of time-series data conducted by Kobayashi et al.¹⁷ allowed three types of machine learning methods to predict pain with high accuracy, with the best random forest method achieving an area under the receiver operating characteristics curve (AUROC) of 0.85 for predicting whether the patient had a CPOT score of ≥ 3 . However, individual differences in vital signs (e.g., variations in vascular, cardiac, and neurological functions) may affect the accuracy of pain prediction in conventional models, thereby reducing the prediction accuracy.

Therefore, the purpose of this study was to investigate (1) the possibility of improving the prediction accuracy by incorporating deviations from baseline patient vital signs into the prediction model and (2) the concurrence of the predicted artificial intelligence (AI) values with the scores obtained using the CPOT, a conventional objective pain assessment, during analgesic administration.

¹Department of Anesthesiology and Perioperative Medicine, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan. ²Hitachi Solutions East Japan, Ltd, Sendai, Miyagi, Japan. ✉email: naoya.kobayashi.a4@tohoku.ac.jp

Results

Among the 10,299 eligible patients (Table 1), the CPOT assessment was conducted 117,190 times, and 3925 (3.3%) with a CPOT score of ≥ 3 were judged to have pain. The distribution of the Richmond Agitation-Sedation Scale (RASS) scores, evaluated at approximately the same time, was 0 at 42.7%, ≥ 1 at 8.3%, and ≤ -1 at 49.0%. The Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) scores indicated that 21.6% of the patients had delirium. The random forest method was used to predict whether a patient had a CPOT score of ≥ 3 points at the time of CPOT evaluation (probability of CPOT being ≥ 3 was defined as pain index), and the AUROC was 0.903, higher than that in the previous study (0.853) (Fig. 1). The importance of the features of the random forest model is shown in Supplementary Table S2. In addition, the generalization performance of the model was verified with an AUROC of 0.873 ± 0.020 over 10 trials with different cross-validation random numbers. The hyperparameters were as follows at the highest accuracy: max_depth: 14; estimates: 17. Accuracy comparisons with other machine learning (logistic regression analysis, LightGBM) performed as subanalyses are shown in Supplementary Fig. S1. The relationship between the pain index and the CPOT assessed by the health care provider is shown in Fig. 2. A trend toward a higher pain index with a higher CPOT was observed.

The probability of pain (probability of CPOT ≥ 3 : pain index, 0–100%) before and after bolus administration of fentanyl was visualized (Fig. 3, see Supplementary Fig. S2 for graphs subdivided by RASS). Patients included were those not used for modeling (Table 1, $N = 326$). The results showed that the pain index at 1 h after administration was significantly lower in the group whose CPOT improved by more than 1 than that of the group whose CPOT had not improved (at administration: 44.7 ± 27.4 ; 1 h after administration: 41.5 ± 26.7 , $P = 0.020$, Table 2).

Discussion

The purpose of this study was to investigate (1) the possibility of improving the prediction accuracy by incorporating deviations from baseline patient vital signs into the prediction model and (2) the concurrence of the predicted AI values with the scores obtained using the CPOT, a conventional objective pain assessment, during analgesic administration. The results showed that (1) pain visualization accuracy corresponded to AUROC 0.902; (2) pain index decreased in tandem only when the CPOT score decreased after a bolus dose of fentanyl.

Pain measurement tools can be divided into self-report tools and behavioral assessment tools. Patients who cannot self-report pain but have observable behaviors must be assessed using behavioral assessment tools¹⁸. The BPS and CPOT have shown the best validity and reliability to date. CPOT, in particular, has been optimized for use in patients with atypical behaviors, such as those with brain injury¹⁹. In contrast, neuromuscular blocking agents (NMBAs) are sometimes used in patients with severe breathing problems, such as patients with acute respiratory distress syndrome; however, they chemically paralyze the patient and hinder behavioral assessment²⁰. Therefore, periodic discontinuation of NMBA is recommended²¹, and alternative pain assessment methods must be used for the duration of its administration. Although vital signs can be used as a guide for the administration of analgesics and sedatives, they are not recommended as an appropriate indicator of pain due to the large variability that occurs when they are used for pain assessment^{22–25}. However, the AI in this study can predict pain with high accuracy despite the use of vital signs and may be applicable even to patients for whom behavioral indicators are not available. Since vital signs are data that are automatically obtained in many patients in ICUs, it may be possible to assess pain automatically and continuously if this AI can be put to practical use.

Recent topics in new objective pain assessment methods, such as the present study, include several electrophysiological tools. Pupil monitoring, a method of assessing pain by recording fluctuations in the pupils in relation to the sympathetic-parasympathetic responses, has been reported to give inconsistent results in critically ill patients²⁶. A method of assessing the pupillary response by applying gentle electrical stimulation to the skin has also been developed; however, it can only be used in patients who are under moderate-to-deep sedation for pain²⁷. The validity of adding new pain to patients for pain assessment is also questionable. The Analgesia Nociception Index (ANI), which indicates pain on a scale of 0 to 100, has been reported to be particularly useful during dressing changes, with a negative predictive value of 90% using a cut-off of 42.5 points. However, the sensitivity and specificity were 61.4% and 77.4%, respectively, and the number of participants was only a few dozen^{28,29}. The Nociception Level (NOL) is a new multiparameter pain assessment system that, similar to the ANI, can display pain on a scale from 0 to 100³⁰. In addition to HRV, the device combines photoplethysmography pulse wave amplitude, skin conductance, and body temperature. Pilot studies have shown that NOL is associated with NRS and CPOT during endotracheal suctioning and cuff inflation as well as chest tube removal^{31,32}. However, the number of applicable patients in both studies was limited and further validation must be performed in the future. The pain index validated in this study is a new tool to determine pain using vital signs without HRV. It was modeled on data from approximately 10,000 patients and can be varied in parallel with CPOT with an accuracy better than an AUROC of 0.9.

The limitations of this study include its retrospective, single-center design and the possible heterogeneity in general patient background data. In addition, CPOT score was also assessed in patients who were able to communicate to provide continuous pain visualization throughout their ICU stay, and this was used as training data to calculate the scores. The pain index indicates the probability of achieving a CPOT score of ≥ 3 . In this situation, there is an immediate risk of accidental tube removal and falls, and the ability to predict these and pain may be useful for patients in ICU. Another problem is that the pain index could not be displayed immediately after ICU admission because the system needs time to establish the baseline as the point at which the patient's condition stabilizes after ICU admission; the pain index is displayed after that point. Furthermore, the fentanyl dose per body weight was not constant in the validation step using the fentanyl bolus dosing data. Lastly, although the pain management protocol was followed, the final dosing decision was made by the charge nurse, and there may have been differences in decision criteria.

Characteristics	Model derivation and validation	Pain visualization test with fentanyl ^a
Number of patients	10,299	326
Age: year, median (IQR)	65 (73, 52)	67 (74, 52)
Age group: <i>N</i> (%)		
20–44	1847 (17.9%)	69 (21.1%)
45–64	3673 (35.7%)	85 (26.0%)
65≤	6007 (58.3%)	130 (39.8%)
Male: <i>N</i> (%)	6660 (58.0%)	202 (62.0%)
Days in ICU: days, median (IQR)	1 (3, 1)	4 (9, 1)
Days in ICU: hours, median (IQR)	21.6 (62.6, 17.6)	98.7 (222.2, 24.2)
APACHE II at ICU admission: median (IQR)	13 (17, 10)	16 (20, 13)
Mechanical ventilation: <i>N</i> (%)	2479 (24.1%)	88 (27.0%)
Mortality in ICU: <i>N</i> (%)	813 (7.9%)	31 (9.5%)
Reason for ICU admission: <i>N</i> (%)		
Medical	3985 (38.7%)	111 (34.0%)
Elective operation	5418 (52.6%)	176 (54.0%)
Emergency operation	896 (8.7%)	39 (12.0%)
ICU admission diagnosis: <i>N</i> (%)		
Heart	1321 (12.8%)	49 (15.0%)
Heart failure	399 (3.9%)	16 (4.9%)
Acute myocardial infarction	121 (1.2%)	6 (1.8%)
Aorta and vessels	1208 (11.7%)	47 (14.4%)
Kidney	487 (4.7%)	10 (3.0%)
Lung and mediastinum	1155 (11.2%)	44 (13.5%)
Acute respiratory failure	197 (1.9%)	10 (3.1%)
Digestive organs	3154 (30.6%)	85 (26.1%)
Metabolic and endocrine organs	198 (1.9%)	6 (1.8%)
Orthopedics	305 (3.0%)	9 (2.8%)
Genital organs	1053 (10.2%)	16 (4.9%)
Obstetrics	123 (1.2%)	1 (0.3%)
Neurosurgery	654 (6.4%)	5 (1.5%)
Others	1869 (18.1%)	54 (16.6%)
Pain, agitation and delirium assessments		
CPOT: <i>N</i> (%)	117,190	22,017
0–2	113,265 (96.7%)	21,437 (97.3%)
≥3	3925 (3.3%)	580 (2.6%)
RASS: <i>N</i> (%)	136,034	24,267
4	111 (0.1%)	11 (0.0%)
3	600 (0.4%)	117 (0.5%)
2	2899 (2.1%)	450 (1.9%)
1	7727 (5.7%)	1171 (4.8%)
0	58,174 (42.8%)	11,184 (46.1%)
–1	21,608 (15.9%)	3137 (12.9%)
–2	13,303 (9.8%)	1940 (8.0%)
–3	14,663 (10.8%)	2505 (10.3%)
–4	11,766 (8.6%)	2394 (9.9%)
–5	5183 (3.8%)	1358 (5.6%)
CAM-ICU: <i>N</i> (%)	71,497	11,737
Negative	55,999 (78.3%)	9105 (77.6%)
Positive	15,498 (21.7%)	2632 (22.4%)
Average dosage of sedatives and analgesics (<i>Number of patients receiving each drug, %</i>)		
Sedatives		
Dexmedetomidine (µg/kg/h)	0.38 (N: 2380, 21.1%)	0.32 (N: 234, 71.8%)
Propofol (mg/kg/h)	2.05 (N: 1802, 16.0%)	2.68 (N: 221, 67.8%)
Midazolam (mg/kg/h)	0.04 (N: 289, 2.6%)	0.04 (N: 60, 18.4%)
Analgesics		
Fentanyl (µg/kg/h)	1.88 (N: 2878, 25.5%)	1.35 (N: 326, 100.0%)
Continued		

Characteristics	Model derivation and validation	Pain visualization test with fentanyl ^a
Hydromorphone	N: 31 (0.3%)	N: 10 (3.1%)
Ketamine	N: 15 (0.1%)	N: 10 (3.1%)
Epidural anesthesia (levobupivacaine)	N: 1235 (12.0%)	N: 0 (0.0%)

Table 1. Patient characteristics. *CPOT* critical care pain observation tool, *RASS* richmond agitation-sedation scale, *CAM-ICU* confusion assessment method for the intensive care unit, *ICU* intensive care unit, *IQR* interquartile range. ^aPatient groups not included in model derivation.

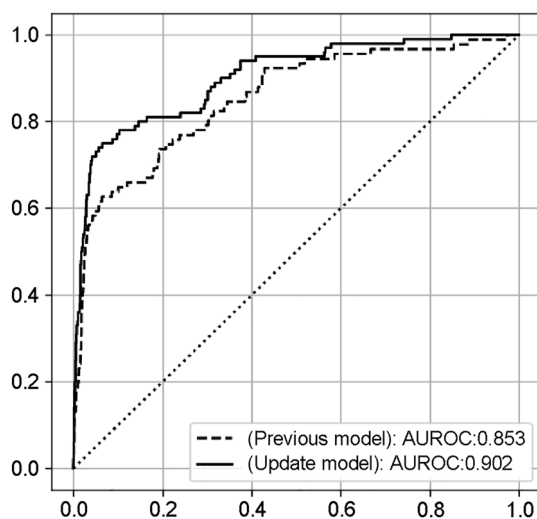


Figure 1. Accuracy of pain visualization. The “previous model” represents the previously reported model¹⁷, whereas the “updated model” represents the AUROC of the model proposed in this study. The x-axis and y-axis represent the negative sensitivity and specificity in the ROC curve, respectively. The accuracy of the test depends on the ability of the machine learning model to correctly determine whether the *CPOT* score is < 2 or > 3 . The accuracy is expressed by AUROC, where a range of 1 indicates a perfect test and a range of 0.5 indicates an inconclusive test. In the previous model, the sensitivity was 64.8%, the specificity was 88.2%, the positive predictive value was 11.0%, and the negative predictive value was 99.1%. In the updated model, the sensitivity was 73.0%, the specificity was 94.5%, the positive predictive value was 28.0%, and the negative predictive value was 99.2%.

In conclusion, this study confirmed that incorporating individual patient baseline data into a previously developed pain visualization model improved the accuracy and treatment follow-up. The next step in the practical application of this model is an open-label, prospective, randomized, controlled trial in a multicenter setting.

Methods

Design and study setting

This retrospective observational study was conducted in a single intensive care unit (ICU) at a single institution in Japan. Ethical approval was obtained from Ethics Committee Tohoku University Graduate School of Medicine (Study No. 2022-1-334). Owing to the retrospective design of the study, the requirement for obtaining written informed consent was waived from Ethics Committee Tohoku University Graduate School of Medicine. All methods used in this study were conducted in accordance with the tenets and regulations of the Declaration of Helsinki. The study design was entered into a database (ID: R000047019 UMIN000041179, URL: <https://www.umin.ac.jp/ctr/index.htm>).

Participants

Patients admitted to the ICU between October 2016 and March 2019 (1) aged ≥ 20 years with (2) at least five *CPOT*, *RASS*, and *CAM-ICU* assessments and (3) electrocardiography (ECG) and arterial pressure monitoring for at least 30 min were included in the study. Data from the following patients whose vital signs differed significantly from those of the general adult population were excluded: (1) patients who had undergone cardiopulmonary bypass; (2) pregnant patients; (3) patients who had undergone organ transplantation, artificial heart transplantation, extracorporeal membrane oxygenation, and intra-aortic balloon pump surgery; and (4) patients with a do-not-resuscitate order. Data were obtained from the electronic medical records system of the institution (PrimeGaia, Nihon Kohden Corporation, Tokyo, Japan). Patient identification information was not collected. In addition, patients admitted to the ICU between April 2019 and April 2022 who had received a bolus

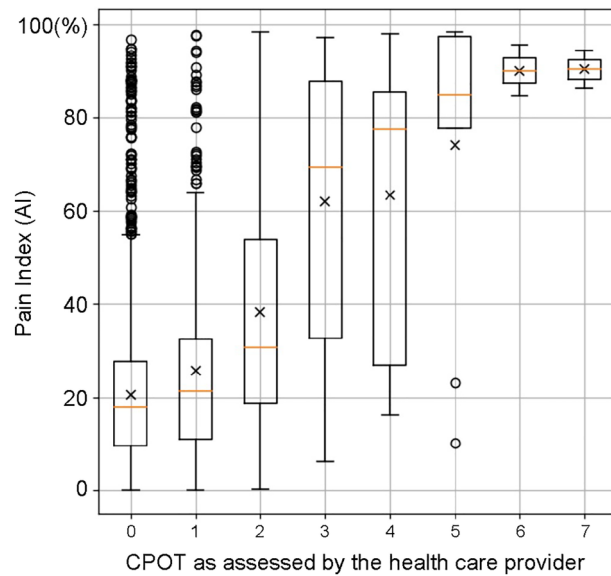


Figure 2. The relationship between the pain index and the CPOT assessed by the healthcare provider. The top and bottom edges of the boxes indicate the quartile range, the horizontal line indicates the median and the crosses indicate the mean. In the CPOT improvement group, the CPOT score decreased by at least 1. The dashed line indicates the median. The upper and lower colored ranges indicate the 75th and 25th percentiles, respectively.

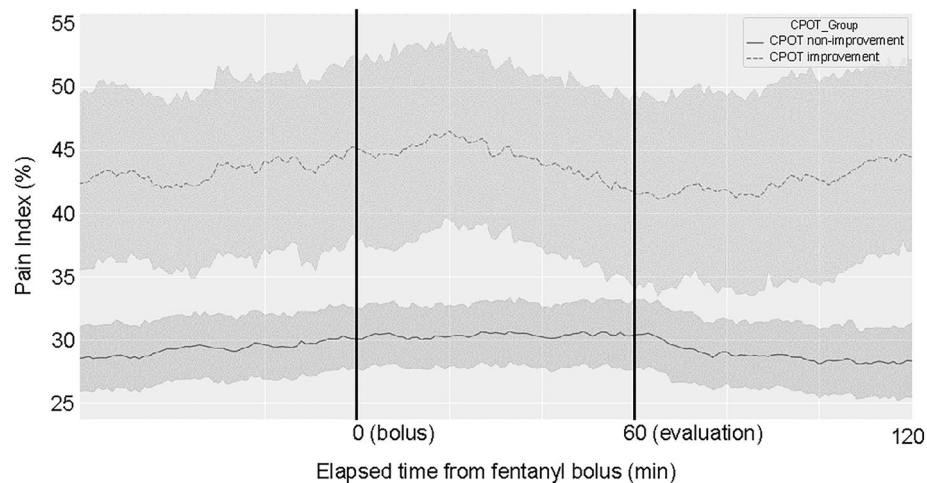


Figure 3. Change over time in pain index with fentanyl bolus administration. The probability of having a CPOT score of ≥ 3 calculated by the artificial intelligence was defined as the pain index, and changes in the pain index were plotted in chronological order before and after fentanyl administration. The time of fentanyl administration was set at 0 min, and the pain index displayed at 60 min was compared between the two groups based on whether the CPOT improved by 1 or more points. The dashed and solid lines in the figure show the median values for the CPOT improvement and no improvement groups, respectively. Statistical evaluation was performed 60 min after the administration of the fentanyl bolus.

dose of fentanyl (25–100 μg) with vital sign data available for each minute were selected to test the treatment response of a dataset not involved in the creation and validation of these machine learning models.

Assessment and treatment of analgesia, sedation, and delirium

The CPOT score was used as the training target of the model to assess the pain level of all eligible patients. The CPOT assessments were performed by ICU nurses every 8 h and when obvious pain was observed. The RASS was used to assess the sedation level. Delirium was assessed using the CAM-ICU scale. The RASS, CAM-ICU, and CPOT assessments were used simultaneously by several nurses to ensure agreement among the data recorded by them. In cases of disagreement, the final decision was made by the intensivist. The patients were treated according to our pain management protocol (Supplementary Fig. S3).

CPOT	N	Pain Index (%)		P value
		Before fentanyl bolus	Post fentanyl bolus	
All	326	32.3 ± 21.4	31.2 ± 21.6	0.174
Improvement group	58	44.7 ± 27.4	41.5 ± 26.7	0.020*
Non-improvement group	268	29.6 ± 18.9	28.9 ± 19.6	0.632

Table 2. Changes in pain index with fentanyl bolus administration. The CPOT score decreased by at least 1 in the improvement group. The Pain Index indicates the probability of a human-rated CPOT score of ≥ 3 . Fisher's exact probability test was used for statistical analysis. CPOT critical care pain observation tool. * $P < 0.05$.

Vital sign data

The heart rate, arterial oxygen saturation, and arterial pressure were recorded every minute. The arterial pressure was continuously monitored via the left or right radial artery. The respiratory rate was measured by electrocardiographic impedance. The noise was removed using the methods described in Supplementary Table S3.

If any of the abovementioned items were applicable to any record, corresponding to a 1-min interval, all vital signs from that record were excluded from the analysis.

Primary outcome

A CPOT score of 3 was considered positive and a score of 2 was considered negative for pain, and the probability of being positive was calculated using machine learning (pain index). In this study, the predicted probability was defined as the pain index. A grid search was performed on the training data, and the hyperparameters that yielded the highest accuracy for each model were selected. In addition, the harmonic means of the sensitivity, specificity, and AUROC were calculated.

Machine learning

We used a machine learning algorithm to predict pain when the CPOT score was assessed by a healthcare professional. Random forest, which was reported to be the most accurate algorithm¹⁷, was used in this study. Two other machine learning methods (logistic regression and LightGBM) were also used for supplementary analysis. The characteristics of each machine learning method are shown in Supplementary Table S4. The following explanatory variables were included: arterial pressure (systolic, mean, and diastolic), heart rate, respiratory rate, sex, stratified age (20–44, 45–64, and 65+ years), and the RASS score. The CAM-ICU score was used for exploratory analysis but not for building the final model. Information processing was conducted to reflect the individual differences in the vital signs of each patient in the model.

Noise canceling and oversampling

The noise in the vital sign data was removed according to the requirements listed in Supplementary Table S3. Owing to the relatively small number of pain-positive patients, the positive group was oversampled using adaptive synthetic sampling. The building process illustrated in Supplementary Fig. S4 was followed to build the subsequent model to avoid the effects of overfitting. Tuning of hyperparameters was targeted for $n_estimator$ and max_depth . See Supplementary Fig. S5 for details.

Baseline evaluation

Baselines were obtained for each patient and adjusted as shown below (Supplementary Fig. S6):

1. Hourly records of patients admitted to the ICU were extracted.
2. Approximately 90% of the hourly data were evaluated for availability and two-tailed linear completion was performed.
3. The hourly data were divided into three chronological orders. Friedman tests were performed to evaluate statistically significant differences among the three age groups. The condition was considered unstable if a significant statistical difference was observed, and the condition was returned to condition (ii) after 20 min.
4. Variability (standard deviation/mean) was calculated to ensure that the condition did not deviate from the threshold values established for each type of vital sign in previous studies. The condition was considered unstable if the threshold was exceeded, and the patient was moved to (ii) after 20 min.
5. The 5-min moving average during this period was used as the baseline.

Calculation of the estimated probability of pain (pain index)

The estimated probability of pain during the CPOT assessment was also calculated. This probability was defined as the pain index. Vital signs from 1 h before the evaluation were extracted. After linear interpolation in both directions for the interval where more than 90% of the data existed in 1 h, the area over or under the baseline curve was obtained, and the difference from the baseline was calculated. The calculation methods and formulas for these characteristics are shown in Supplementary Fig. S7. Lastly, the age group, sex, and RASS scores were linked to the training data.

Evaluation of treatment response (visualization)

The pain index was visualized using a new dataset that was not used to build and validate the model. The patients who received a bolus dose of fentanyl and for whom vital data could be obtained each minute were divided into two groups according to CPOT score improvement by more than 1. The transition was graphed for each group. Other analgesics (ketamine and morphine) had only temporary and limited use and were analyzed for periods when these drugs were not administered.

Statistical analysis

Data analysis was performed using JMP v15 (SAS Institute Inc., Cary, NC, USA). Normally distributed data were reported as mean \pm standard deviation, and non-normally distributed data were reported as median and interquartile range. The AUROC was used to compare the accuracy and was classified as low (0.5–0.7), moderate (0.7–0.8), and high (≥ 0.8).

Data availability

The datasets generated and/or analyzed during the current study are available in the UMIN-ICDR repository at <http://www.umin.ac.jp/icdr/index-j.html>, with permission from the authors.

Received: 15 August 2023; Accepted: 13 October 2023

Published online: 14 October 2023

References

1. Tesarz, J. *et al.* Widespread pain is a risk factor for cardiovascular mortality: Results from the Framingham Heart Study. *Eur. Heart J.* **40**, 1609–1617 (2019).
2. Macfarlane, G. J., Barnish, M. S. & Jones, G. T. Persons with chronic widespread pain experience excess mortality: Longitudinal results from UK Biobank and meta-analysis. *Ann. Rheum. Dis.* **76**, 1815–1822 (2017).
3. Gelinas, C. Management of pain in cardiac surgery ICU patients: Have we improved over time?. *Intensive Crit. Care Nurs.* **23**, 298–303 (2007).
4. Chanques, G. *et al.* A prospective study of pain at rest: Incidence and characteristics of an unrecognized symptom in surgical and trauma versus medical intensive care unit patients. *Anesthesiology* **107**, 858–860 (2007).
5. Payen, J. F. *et al.* Current practices in sedation and analgesia for mechanically ventilated critically ill patients: A prospective multicenter patient-based study. *Anesthesiology* **106**, 687–695 (2007).
6. Morris, J. L. *et al.* Determinants of pain assessment documentation in intensive care units. *Can. J. Anaesth.* **68**, 1176–1184 (2021).
7. Smith, D., Wilkie, R., Croft, P., Parmar, S. & Mcbeth, J. Pain and mortality: Mechanisms for a relationship. *Pain* **159**, 1112–1118 (2018).
8. Vaegter, H. B. *et al.* Cause-specific mortality of patients with severe chronic pain referred to a multidisciplinary pain clinic: A cohort register-linkage study. *Scand. J. Pain* **19**, 93–99 (2018).
9. Puntillo, K. A. *et al.* Determinants of procedural pain intensity in the intensive care unit. The Europain[®] study. *Am. J. Respir. Crit. Care Med.* **189**, 39–47 (2014).
10. Payen, J. F. *et al.* Assessing pain in critically ill sedated patients by using a behavioral pain scale. *Crit. Care Med.* **29**, 2258–2263 (2001).
11. Gelinas, C., Fillion, L., Puntillo, K. A., Viens, C. & Fortier, M. Validation of the critical-care pain observation tool in adult patients. *Am. J. Crit. Care* **15**, 420–427 (2006).
12. Marques, R. *et al.* Validation testing of the European Portuguese critical-care pain observation tool. *Healthcare* **10**, 1075 (2022).
13. Kontou, P. *et al.* Evaluation of pain scales and outcome in critically ill patients of a Greek ICU. *J. Pain Palliat. Care Pharmacother.* **37**, 34–43 (2022).
14. Waladani, B., Yuniar, I., Setianingsih, E., Rahmah, N. N. & Yunita, I. P. Application of pain scale assessment in patients attached mechanical ventilator in intensive care unit. In: *4th International Conference on Sustainable Innovation 2020–Health Science and Nursing (ICoSIHSN 2020)* 332–336 (Atlantis Press, 2021).
15. Gélinas, C. *et al.* A psychometric analysis update of behavioral pain assessment tools for noncommunicative, critically ill adults. *AACN Adv. Crit. Care* **30**, 365–387 (2019).
16. Barr, J. *et al.* Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit. Care Med.* **41**, 263–306 (2013).
17. Kobayashi, N. *et al.* Semi-automated tracking of pain in critical care patients using artificial intelligence: A retrospective observational study. *Sci. Rep.* **11**, 5229 (2021).
18. Devlin, J. W. *et al.* Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit. Care Med.* **46**, e825–e873 (2018).
19. Gélinas, C. *et al.* Validation of the critical-care pain observation tool-neuro in brain-injured adults in the intensive care unit: A prospective cohort study. *Crit. Care* **25**, 142 (2021).
20. Chanques, G. & Gélinas, C. Monitoring pain in the intensive care unit (ICU). *Intensive Care Med.* **48**, 1508–1511 (2022).
21. Chanques, G. *et al.* Analgesia and sedation in patients with ARDS. *Intensive Care Med.* **46**, 2342–2356 (2020).
22. Boitor, M., Martorella, G., Arbour, C., Michaud, C. & Gélinas, C. Evaluation of the preliminary effectiveness of hand massage therapy on postoperative pain of adults in the intensive care unit after cardiac surgery: A pilot randomized controlled trial. *Pain Manag. Nurs.* **16**, 354–366 (2015).
23. Chen, H. J. & Chen, Y. M. Pain assessment: Validation of the physiologic indicators in the ventilated adult patient. *Pain Manag. Nurs.* **16**, 105–111 (2015).
24. Kapoustina, O., Echegaray-Benites, C. & Gélinas, C. Fluctuations in vital signs and behavioural responses of brain surgery patients in the intensive care unit: Are they valid indicators of pain?. *J. Adv. Nurs.* **70**, 2562–2576 (2014).
25. Gelinas, C. & Arbour, C. Behavioral and physiologic indicators during a nociceptive procedure in conscious and unconscious mechanically ventilated adults: Similar or different?. *J. Crit. Care* **24**(628), e7–628.e17 (2009).
26. Bernard, C. *et al.* Assessing pain in critically ill brain-injured patients: A psychometric comparison of 3 pain scales and videopupillometry. *Pain* **160**, 2535–2543 (2019).
27. Vinclair, M. *et al.* Using pupillary pain index to assess nociception in sedated critically ill patients. *Anesth. Analg.* **129**, 1540–1546 (2019).
28. Chanques, G. *et al.* Analgesia nociception index for the assessment of pain in critically ill patients: A diagnostic accuracy study. *Br. J. Anaesth.* **119**, 812–820 (2017).

29. Chanques, G. *et al.* Is there a single non-painful procedure in the intensive care unit? It depends!. *Intensive Care Med.* **44**, 528–530 (2018).
30. Ben-Israel, N., Kliger, M., Zuckerman, G., Katz, Y. & Edry, R. Monitoring the nociception level: A multi-parameter approach. *J. Clin. Monit. Comput.* **27**, 659–668 (2013).
31. Gélinas, C. *et al.* Exploration of a multi-parameter technology for pain assessment in postoperative patients after cardiac surgery in the intensive care unit: The Nociception Level Index (NOL) (TM). *J. Pain Res.* **14**, 3723–3731 (2021).
32. Shahiri, T. S., Richard-Lalonde, M., Richebé, P. & Gélinas, C. Exploration of the Nociception Level (NOL™) Index for pain assessment during endotracheal suctioning in mechanically ventilated patients in the intensive care unit: An observational and feasibility study. *Pain Manag. Nurs.* **21**, 428–434 (2020).

Acknowledgements

We wish to thank Hitachi Solutions East Japan, Ltd. (Sendai, Miyagi, Japan) for their timely help and Nihon Kohden Corporation (Tokyo, Japan) for analyzing the large quantity of data. We would like to thank Editage (www.editage.com) for English language editing.

Author contributions

N.K. prepared the main manuscript text. K.W. and H.M. analyzed the data and prepared the figures and tables. M.Y. supervised the entire study. All the authors read and approved the final manuscript.

Funding

This study was supported by the JSPS KAKENHI (Grant Number JP19K18343) and the research grant from Hitachi Solutions East Japan, Ltd. JSPS played no role in this study. Hitachi Solutions East Japan had a collaborative research relationship with this study and cooperated mainly in the aspect of machine learning.

Competing interests

NK received a research grant from Hitachi Solutions East Japan, Ltd. The other authors declare that they have no conflicts of interest to disclose.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44970-2>.

Correspondence and requests for materials should be addressed to N.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023