

ORIGINAL WORK



Artificial Intelligence for Early Prediction of Intracranial Hypertension: From Single-Centre Development to Multicentre Validation

Peter Galos^{1*†} , Ludvig Hult^{2†}, Giuseppe Citerio³, Anders Lewén¹, Anders Hånell¹, Thomas B. Schön⁴, Per Enblad¹ In collaboration with The CENTER-TBI High-Resolution ICU (HR ICU) Sub-Study Participants and Investigators

© 2026 The Author(s)

Abstract

Background: Raised intracranial pressure is one of the most severe secondary insults after traumatic brain injury. Recognising physiological patterns before they progress to dangerous intracranial pressure elevations could enable preventive rather than reactive interventions. In our previous Uppsala single-centre study, an XGBoost machine learning model showed high discriminative performance for prediction of imminent intracranial pressure insults. The model identifies short-term physiological trajectories associated with imminent threshold crossing rather than predicting independent future clinical events. Whether these results are generalisable across centres is unknown. The aim was therefore to evaluate whether the XGBoost model trained on an Uppsala dataset maintains similar performance when applied to a multicentre TBI cohort - The CENTER-TBI High-Resolution ICU cohort.

Methods: The XGBoost model trained on the Uppsala high frequency traumatic brain injury cohort was externally validated using data from the multicentre CENTER-TBI cohort. Minute-level averages of intracranial pressure and invasive arterial blood pressure were used to predict the onset of an intracranial pressure insult within 30 min on the basis of a 4h historical window. A bidirectional validation design was applied and performance was assessed using multiple metrics.

Results: The cohorts comprised 602 patients from Uppsala and 255 patients from the CENTER-TBI database. When trained on the Uppsala cohort and tested on CENTER-TBI, the model achieved an area under the receiver operating characteristic curve (AUROC) of 97.4%, sensitivity of 93.8%, and specificity of 98.6%. Training on CENTER-TBI and testing on the Uppsala dataset yielded an AUROC of 96.1%, sensitivity of 81.2%, and specificity of 98.7%. Balanced accuracy exceeded 88% in both validation directions.

Conclusions: The XGBoost model retained high performance under external multicentre validation, indicating that pre-insult physiological patterns preceding increased intracranial pressure are shared across cohorts. Differences in sensitivity between validation directions were observed and may reflect cohort characteristics or dataset size. These findings support the need for local evaluation before clinical implementation.

Keywords: Traumatic brain injury, Intracranial pressure, Machine learning, Artificial intelligence, External validation, Neurointensive care

*Correspondence: peter.galos@uu.se

†Peter Galos and Ludvig Hult have contributed equally to this work.

¹ Department of Medical Sciences/Neurosurgery, Uppsala University, Uppsala, Sweden

Full list of author information is available at the end of the article

The CENTER-TBI High-Resolution ICU (HR ICU) Sub-Study Participants and Investigators are present in Acknowledgment section.

Introduction

Traumatic brain injury (TBI) still remains one of the most important causes of death and disability worldwide [1]. In severe cases, patients require neurointensive care (NIC), where the prevention of secondary insults and avoiding development of secondary brain injury is the primary therapeutic goal. Amongst all possible secondary clinical insults, raised intracranial pressure (ICP) is one of the most severe, as it may lead to cerebral ischaemia and in worst case total brain infarction if not treated promptly. Continuous monitoring of ICP enables early treatment when the values become increased, but recognising the patterns that precede dangerous ICP elevations remains challenging. A system recognising such patterns could warn before an ICP insult develops and enable increased surveillance, further evaluation, and, when appropriate, intervention, potentially preventing the insult from occurring.

Artificial intelligence (AI) is expected to play a significant role in medicine and is already being increasingly applied in several clinical fields. Using rich data sources, machine learning (ML) models have been used in radiology, pathology and gastroenterology, amongst others [2]. In neurosurgery and NIC, similar models have been applied to predicting the risk of vasospasm in subarachnoid hemorrhage [3], the need for surgical intervention after admission in severe TBI [4] and for early prediction of ICP insults in TBI [5, 6].

In our previous study, four different ML models were trained on the Uppsala High-Frequency TBI (UHF-TBI) dataset for short-term prediction of imminent ICP insults within 30 min [7]. Amongst them, the Extreme Gradient Boosting (XGBoost) model achieved high discriminative performance (area under the receiver operating characteristic curve; AUROC 98.7%, sensitivity 93.8% and specificity 94.6%) using ICP and arterial blood pressure (ABP) data only, with marginally improved performance when demographic data were added, whilst remaining computationally lightweight. Its combination of strong accuracy, modest computational demands and ready availability makes it an attractive ML model for clinical applications. The results achieved for ICP insult prediction were promising but were obtained from a single-centre dataset and may not be generalisable.

The Collaborative European NeuroTrauma Effectiveness Research in TBI (CENTER-TBI) project offers an ideal opportunity for external validation of the results on a multicentre dataset. The CENTER-TBI initiative was a European research project, which as one part collected a cohort of patients with TBI with high-resolution physiological recordings, defined here as

physiological monitoring data sampled at 100 Hz or higher, under standardised conditions.

External validation using this dataset would reveal whether a model trained on the Uppsala data performs similarly when applied to patients treated at other hospitals. This is a critical step towards understanding model generalisability and the feasibility of AI-based early warning systems in real clinical environments.

The aim of this study was therefore to evaluate whether the XGBoost model trained on the UHF-TBI dataset maintains similar performance when applied to the CENTER-TBI dataset.

Methods

Patients and Management Protocol

The inclusion criteria were age 16 years or older and at least 10h of high-frequency ICP monitoring data collected. The final UHF-TBI dataset consisted of 602 patients with traumatic brain injury who were treated at the NIC unit at Uppsala University Hospital between 2008 and 2019 [7]. Patients were included regardless of whether the traumatic brain injury occurred as an isolated head injury or as part of polytrauma. In the CENTER-TBI database [8], 274 patients treated between 2014 and 2017 fulfilled the inclusion criteria. Systematic data integrity review revealed that 2 patients had missing demographic data and 17 had inconsistent high-frequency data, for example, detected unclear changes in sampling frequency, or were marked unreliable in the supplied CENTER-TBI data quality table. The final validation cohort, therefore, comprised 255 patients with complete and valid continuous recordings from 21 centres.

Both Uppsala and the CENTER-TBI centres follow international NIC guidelines [9], although minor local differences in treatment routines and monitoring practices may exist.

Data Collection and Data Management

In Uppsala, high-frequency physiological patient monitoring data, such as invasive arterial blood pressure (ABP) and ICP, are routinely collected from the networked bedside monitors by the ODIN system, a data monitoring and collection system developed in Edinburgh and Uppsala by Tim Howells et al. [10]. The data are stored on a central server at the university hospital.

In CENTER-TBI centres, patient monitoring data was collection by ICM+ and stored in the high-resolution repository (University of Cambridge, UK) and/or by the Moberg Neuromonitoring system (Moberg Research Inc., USA). Data from the CENTER-TBI centres were then collected through the Quesgen e-CRF (Quesgen

Systems Inc, USA), hosted on the INCF platform and LUMC Opal platform and extracted via the LUMC Opal platform (LUMC, NL). Version Opal 3.1 of the CENTER-TBI dataset was used in this study.

The monitoring data in the UHF-TBI dataset are sampled at 100–200 Hz and at 120–240 Hz in the CENTER-TBI dataset, depending on monitoring equipment and centre practices. Data points marked as low quality or invalid in the source databases were not treated as physiological measurements and were replaced with NaN, indicating that no reliable value had been produced by the monitoring system at that timepoint. Minute means of ICP and ABP were then calculated from valid data points, and missingness indicators were included where data were unavailable. Different medical monitoring and data collection equipment of centres had various clipping of data in their supplied data, thus ICP and ABP values outside the range –10 to 300 mm Hg were removed to reduce the risk of equipment- and centre-specific artefacts. Down-sampling to minute means was done as in previous work [7]; minute means of ICP and ABP were calculated on the basis of valid data points. Mean ICP for each minute was considered reliable if it had at least one data point per second. An indicator was added if data were missing. The data pipeline was reimplemented from previous work [7] and applied for both UHF-TBI and CENTER-TBI. This ensured that input features and pre-processing steps were identical across cohorts, allowing for a direct and fair comparison of model performance.

Mistral AI (Le Chat, Mistral AI, Paris, France) was used to create Fig. 1.

Study Design and Statistics

Descriptive and comparison statistics were computed over both databases. Comparisons of Glasgow Coma Scale (GCS) score components and total GCS score between the cohorts were performed using the two-sided Mann–Whitney *U* (Wilcoxon rank-sum). Differences in anticoagulant use and ICP insults between cohorts were assessed using the chi-squared test for independence, as anticoagulant treatment and ICP insults were binary variables compared between two independent cohorts.

Model training, testing and evaluation procedures followed the same methodological framework and the same performance metrics were used as in the original study [7]. The dataset included only features that were computable from mean ICP and ABP.

The cohorts could differ in patient mix and management practises, which can affect model generalizability negatively. Therefore, emphasis was put on descriptive statistics and two-way evaluation of the models. If one cohort is “easy” to predict and the other is “hard”,

this should be interpreted in the results. Therefore, the performance of models is evaluated when trained on one database and evaluated on the other. This two-way generalizability design helps to interpret the results in relation to the modelling strategy in general and the databases in specific.

Predictive Task and Definition of an ICP Insult

The task was to predict whether an ICP insult would begin within the next 30 min on the basis of a 4h historical window, together constituting a 4.5h time section. An insult start was defined as ICP exceeding 20 mm Hg for five consecutive min. The insult was considered to have ended when the ICP remained at or below 20 mm Hg for five consecutive min (Fig. 1A). This definition reflects clinically accepted thresholds in NIC, enabling comparison with previous work [7, 11]. Each time section was labelled according to whether an ICP insult occurred within the 30 min prediction window (Fig. 1B).

Model Training and Evaluation

The XGBoost model was implemented using the same configuration and class-balancing strategy as in the previous UHF-TBI study [7]. Model training was performed on the complete dataset from one cohort, using five-fold cross-validation for hyperparameter tuning, where hyperparameter tuning is the process of selecting parameters for the predictive model to yield the best average balanced accuracy. During hyperparameter tuning, folds were grouped by year in UHF-TBI and by centre in CENTER-TBI. The model was then refit on the entire training data, using the hyperparameters. The optimised model was then evaluated on the external cohort.

Model performance was summarised using AUROC, sensitivity, specificity and balanced accuracy; given the use of a weighted loss function, which may distort predicted model risk scores, only moderate calibration was anticipated, consistent with previous work [7].

All analyses were made using Python 3.10 and 3.12 with the XGBoost package (version 1.7 [12]).

Ethics

For this type of study, formal consent is not required. The study was approved for UHF-TBI data by the local and the Swedish National Ethical Review Authority. Informed consent was obtained from patient or from their next of kin if the patient was unable to provide consent. It was also approved to include patients treated during the study period from whom consent not had been possible

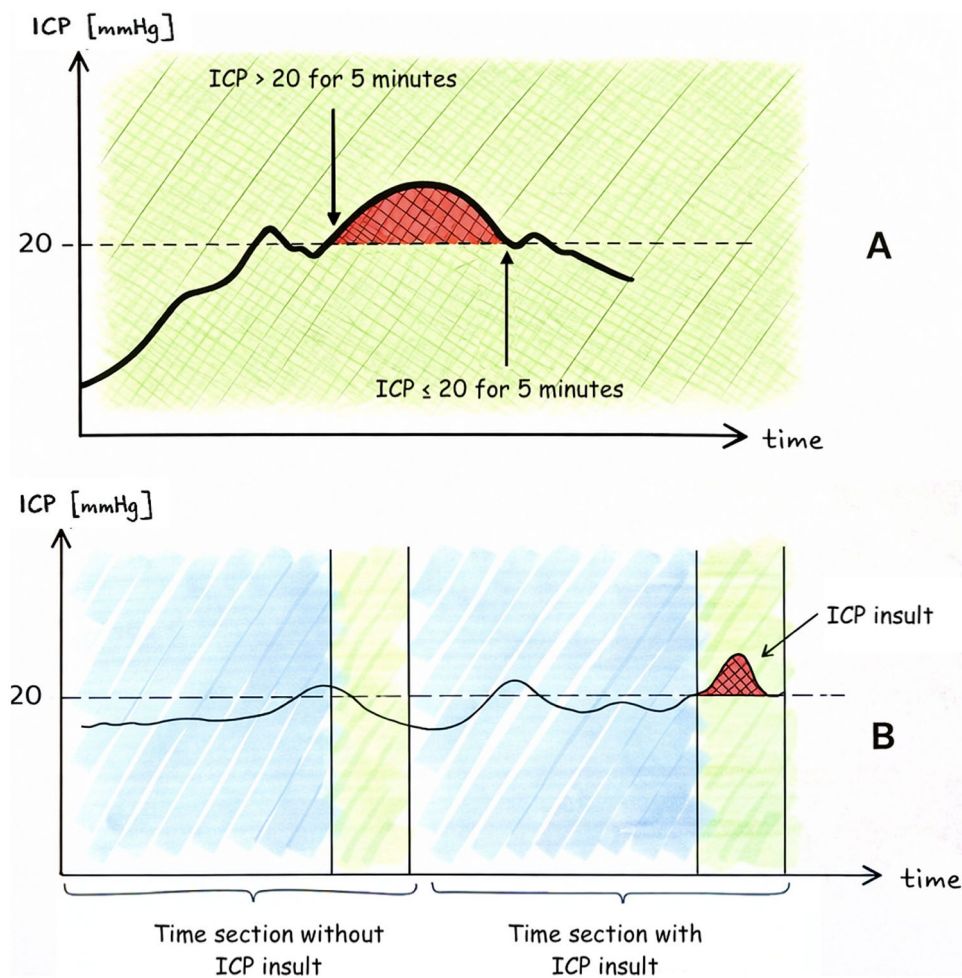


Fig. 1 Schematic illustration of the ICP insult definition and prediction task (the figure is intended to illustrate the study design and does not show actual physiological data). **A** ICP insult definition. The start of an insult is defined as ICP exceeding 20 mm Hg for five consecutive min. The insult is considered to have ended when the ICP remained at or below 20 mm Hg for five consecutive min. **B** Time section without and with an ICP insult. Each time section consists of a 4h historical window (blue) and a 30 min prediction window (green)

to obtain, for example, from deceased patients or when relatives had not been possible to reach.

The CENTER-TBI study (EC grant 602,150) has been conducted in accordance with all relevant laws of the EU if directly applicable or of direct effect and all relevant laws of the country where the Recruiting sites were located, including but not limited to, the relevant privacy and data protection laws and regulations (the “Privacy Law”), the relevant laws and regulations on the use of human materials and all relevant guidance relating to clinical studies from time to time in force including, but not limited to, the ICH Harmonised Tripartite Guideline for Good Clinical Practice (CPMP/ICH/135/95) (“ICH

GCP”) and the World Medical Association Declaration of Helsinki entitled “Ethical Principles for Medical Research Involving Human Subjects”. Informed Consent by the patients and/or the legal representative/next of kin was obtained, accordingly to the local legislations, for all patients recruited in the Core Dataset of CENTER-TBI and documented in the e-CRF. Ethical approval was obtained for each recruiting site. The list of sites, ethical committees, approval numbers and approval dates can be found on the website: <https://www.center-tbi.eu/project/ethical-approval>.

Both datasets were anonymised before analysis and handled in accordance with the Declaration of Helsinki.

Results

Baseline Characteristics of UHF-TBI and CENTER-TBI Cohorts

The cohorts were similar in age (Fig. 2) and sex distribution, with a clear male predominance, 79% CENTER-TBI and 78% UHF TBI. The distribution of the GCS components and the GCS sum score for the cohorts are presented in Table 1 (for details, see Supplementary Table 1). The median GCS sum score at admission was 5 (IQR 3–8) in the CENTER-TBI cohort (195 patients with incomplete/not applicable tests excluded) and 7

(IQR 6–10) in the UHF-TBI cohort (387 patients with incomplete/not applicable tests excluded) ($p < 0.001$); eye response 1 vs. 1, verbal response 1 vs. 1 and motor response 4 vs. 5 (Table 1). Neurological assessments were considered incomplete/not applicable when a valid component score was unavailable. This terminology was used because unavailable assessments in UHF-TBI are recorded as missing and CENTER-TBI explicitly records some assessments as untestable. Anticoagulants were used in 4% of the patients in the UHF-TBI cohort and in 16% of the patients in the

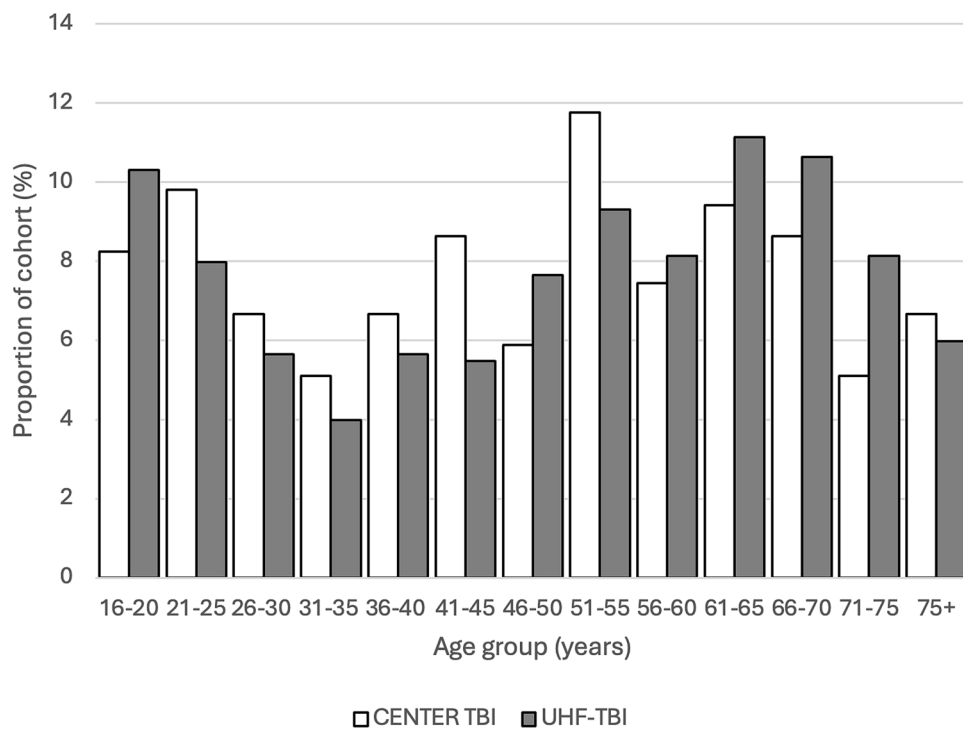


Fig. 2 Age distribution by cohort (%)

Table 1 Comparison of GCS components between cohorts; data are shown as mean (SD) and median (IQR) and differences between groups were assessed using the Mann–Whitney *U* test

GCS component	CENTER-TBI	UHF-TBI	<i>p</i> -value*
GCS Eye, mean (SD)/median (IQR)	1.5 (1.0)/1 (1–1)	1.7 (1.1)/1 (1–3)	0.001
GCS Eye recorded, n (%)	181 (71)	553 (92)	
GCS Verbal, mean (SD)/median (IQR)	1.4 (1.0)/1 (1–1)	1.9 (1.5)/1 (1–2)	0.07
GCS Verbal recorded, n (%)	62 (24)	226 (37)	
GCS Motor, mean (SD)/median (IQR)	3.2 (1.9)/4 (1–5)	4.7 (1.3)/5 (4–6)	< 0.001
GCS Motor recorded, n (%)	146 (57)	601 (100)	
GCS sum score, mean (SD)/median (IQR)	5.9 (3.5)/5 (3–8)	8.1 (3.5)/7 (6–10)	< 0.001
GCS Sum score recorded, n (%)	60 (24)	222 (37)	

* Mann–Whitney *U* test

CENTER-TBI cohort ($\chi^2 = 24.5$, $p < 0.001$) (see Supplementary Table 1).

Occurrence of ICP insults

ICP insults, according to the definition, were present in 16.8% of the 30 min prediction windows in the CENTER-TBI dataset and in 6.4% in the UHF-TBI dataset ($\chi^2 = 783.3$, $p < 0.001$).

Table 2 Results from the different models using alternative evaluation metrics

Training data	UHF-TBI	CENTER TBI	Galos et al. 2025 [7]*
Test data	CENTER TBI	UHF-TBI	
AUROC%	97.4	96.1	98.7 (0.53)
Sensitivity%	93.8	81.2	93.8 (3.68)
Specificity%	98.6	98.7	94.6 (2.27)
Brier score	0.080	0.037	0.04 (0.01)
Accuracy%	88.9	95.1	94.6 (1.98)
Balanced Accuracy%	90.9	88.6	94.2 (1.71)
f1%	74.0	68.1	66.8 (5.55)
PPV%	61.1	58.7	52.0 (5.98)

* Values in parentheses denote Standard Deviations across year-wise cross-validation folds

Model Performance Metrics

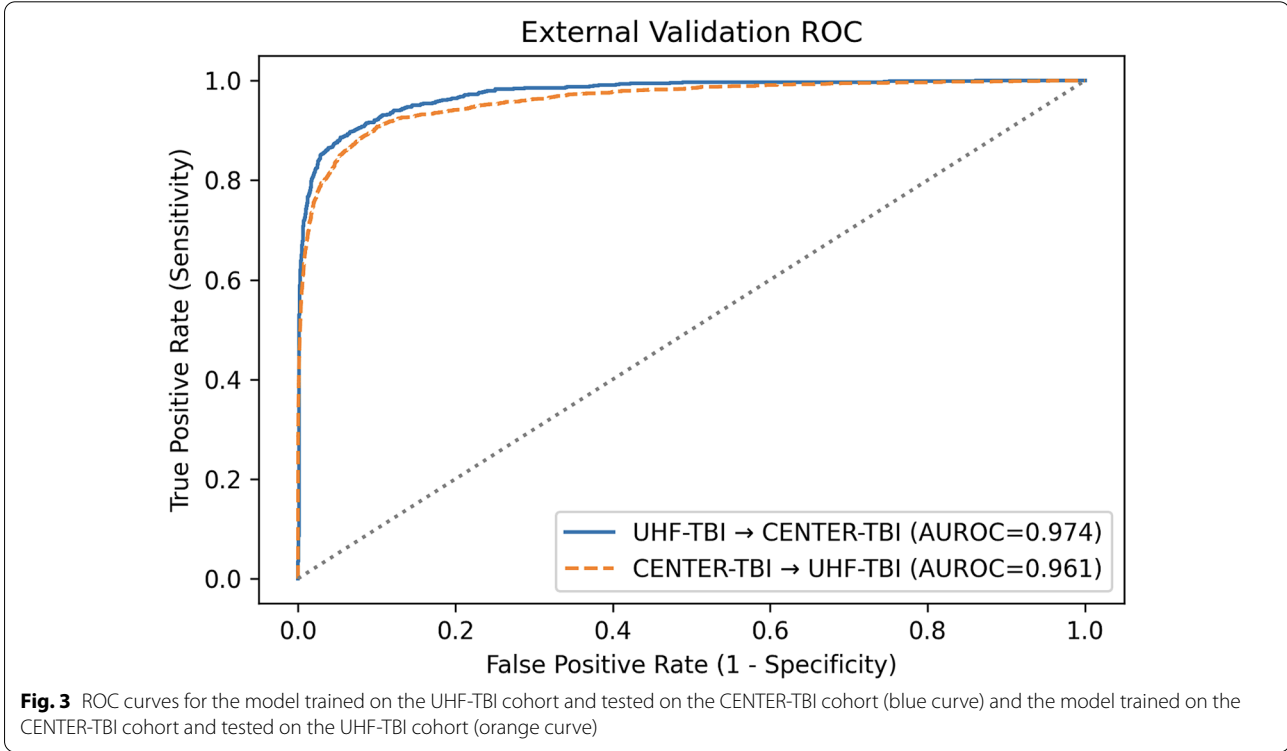
Table 2 and Fig. 3 summarise model performance across all training–test configurations. When trained on UHF-TBI and tested on CENTER-TBI, the model achieved an AUROC of 97.4%, with a sensitivity of 93.8% and specificity of 98.6%. Training on CENTER-TBI and testing on UHF-TBI yielded an AUROC of 96.1% with a sensitivity of 81.2% and a specificity of 98.7%.

Balanced accuracy was 90.9% for UHF-TBI to CENTER-TBI and 88.6% for CENTER-TBI to UHF-TBI.

Across both validation directions, accuracy exceeded 88% whilst F1-score was below 0.7 and positive predictive value ranged from 58 to 70%.

Discussion

This study externally validated an XGBoost model that had been trained on a single-centre cohort for short-term prediction of imminent ICP insults and was found to have high performance in that setting [7]. The model identifies short-term physiological trajectories associated with imminent threshold crossing, rather than predicting independent future clinical events. Overall, the findings support the external validity of the model, but also show that performance may differ between cohorts.



Definition and Occurrence of ICP Insults

ICP insults were defined using a sustained threshold-based criterion consistent with commonly used clinical practice [9, 13, 14]. The proportion of the ICP insults (according to the definition) in the 30-min prediction windows were higher in the CENTER-TBI dataset (16.8% vs. 6.4% in the UHF-TBI dataset).

Whilst local thresholds and clinical responses to ICP elevations may vary between centres to some extent, a uniform definition was applied for retrospective model evaluation. Such differences are therefore unlikely to affect retrospective performance estimates but may influence model behaviour when used in clinical practice, showing a need for evaluation and local retraining, when needed.

External Validation Between Cohorts

Model performance differed depending on the direction of external validation. Training on UHF-TBI and testing on CENTER-TBI gave high AUROC, sensitivity and specificity. Training on CENTER-TBI and testing on UHF-TBI also gave high AUROC and specificity, but lower sensitivity. The AUROC remained above 96% in both validation directions supporting overall model generalisability and suggests that pre-insult physiological patterns are, at least partly, shared across cohorts. However, the lower sensitivity in one validation direction suggests some domain shift, that is, that the relation between physiological patterns and upcoming ICP insults may differ between the cohorts. Thus, external validation does not guarantee identical clinical performance across settings.

As presented in Table 1 and in Supplementary Table 1, admission GCS, type of injury and use of anticoagulants differed significantly between the cohorts. The comparison should be interpreted with some caution because many neurological assessments were recorded as not applicable. Furthermore, injury-type comparisons should be interpreted cautiously because the databases used different registration structures. In UHF-TBI, only the predominant TBI lesion type was registered for each patient, whereas CENTER-TBI allowed multiple lesion types to be registered for the same patient. The overall impression is, however, that the CENTER-TBI cohort includes more severe cases.

Differences in admission neurological status, ICP insult occurrence, distribution of injury type and so on, may have contributed to the observed difference in sensitivity, despite similar preprocessing and feature definitions.

Class Imbalance and Prevalence-Dependent Metrics

ICP insults were uncommon, resulting in class imbalance, that is, the model was trained and tested on fewer

time sections with ICP insults than without ICP insults, especially in the UHF-TBI cohort. This should be considered when interpreting model performance, as several performance metrics, such as Positive Predictive Value (PPV), accuracy and F1-score, depend strongly on the class balance: a model that is equally good per prediction instance over two different cohorts may still look worse if the class balance changes.

The class imbalance is significantly different across CENTER-TBI and UHF-TBI and therefore these metrics must be interpreted with care. The higher insult rate in CENTER-TBI skews the PPV to a greater number, and the accuracy may decrease, compared with a model on UHF-TBI. The F1-score is a type of average across sensitivity and PPV, and would therefore also be expected to increase. Conversely, PPV and F1-score would be expected to decrease when the model is tested in a cohort with lower insult prevalence, such as UHF-TBI.

Local Threshold Setting and Clinical Implementation

Local data characteristics may influence clinically relevant trade-offs between sensitivity, specificity, PPV and alarm burden. Before clinical use, the decision threshold should be evaluated locally. A threshold with high sensitivity may be useful for early warning, but could also increase the number of false alarms, whereas a higher threshold may reduce alarm burden at the cost of missing more true ICP insults. Local recalibration may therefore be needed before implementation.

In a future clinical setting, an alarm would trigger a search for suboptimal physiological conditions which may lead to an ICP insult, for example, insufficient sedation, hyponatraemia, impaired venous outflow, altered PaCO₂, fever or seizures. However, preventive action may change the physiological trajectory on which the model was trained. Real-time deployment studies are therefore needed to assess clinical utility, alarm burden and user response.

Methodological Considerations and Future Work

The results highlight a balance between generalisability and local optimisation. Whilst cross-centre robustness is desirable, locally trained models may achieve superior sensitivity in practice. Future work should evaluate adaptive strategies, including, for example, site-specific retraining or use of transfer learning [15].

Future studies including treatment variables may improve prediction performance and contribute to a better understanding of the underlying pathophysiology. In addition, separate analyses of time periods during ongoing ICP insults and periods without active insults may help clarify the stability of model performance across different physiological states.

Limitations

The feature set was restricted to physiological data and did not include baseline variables (e.g. sex, age and GCS) and treatment variables. Regarding baseline variables, our previous study of the Uppsala cohort including baseline data showed that the performance was marginally improved when included [7]. Differences in monitoring practices and data completeness may have influenced results. Exclusion of incomplete recordings may have introduced selection bias. Furthermore, the low prevalence of ICP insults at the time-window level limited statistical power for subgroup analyses.

Conclusions

The XGBoost model, trained on physiological data, retained strong discrimination for the early prediction of ICP insults under external validation on a multicentre dataset.

Model performance remained high across cohorts, supporting the presence of shared pre-insult physiological patterns in NIC, although sensitivity varied depending on the direction of validation.

The findings underscore the importance of local evaluation before clinical implementation, including assessment of model performance in the intended patient population, data acquisition pipeline, insult prevalence and clinical workflow, with local retraining performed when needed.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s12028-026-02633-w>.

Author details

¹ Department of Medical Sciences/Neurosurgery, Uppsala University, Uppsala, Sweden. ² Department of Medical Sciences/Clinical Epidemiology, Uppsala University, Uppsala, Sweden. ³ School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy. ⁴ Department of Information Technology/Division of Systems and Control, Uppsala University, Uppsala, Sweden.

Acknowledgements

The authors acknowledge the CENTER-TBI HR-ICU Investigators and participants for providing data from their cohort.

CENTER TBI High Resolution (HR ICU) Sub Study Participants and Investigators Audny Anke (Department of Physical Medicine and Rehabilitation, University Hospital Northern Norway), Ronny Beer (Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria), Bo-Michael Bellander (Department of Neurosurgery & Anaesthesia & intensive care medicine, Karolinska University Hospital, Stockholm, Sweden), Erta Beqiri (NeuroIntensive Care, Niguarda Hospital, Milan, Italy), Andras Buki (Department of Neurosurgery, Medical School, University of Pécs, Pécs, Hungary; Neurotrauma Research Group, János Szentágotai Research Centre, University of Pécs, Pécs, Hungary), Manuel Cabeleira (Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK), Marco Carbonara (Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy), Arturo Chiericato (NeuroIntensive Care, Niguarda Hospital, Milan, Italy), Giuseppe Citerio (NeuroIntensive Care Unit, Department of Anaesthesia & Intensive Care, ASST di Monza, Monza, Italy; School of Medicine and Surgery,

Università Milano Bicocca, Milano, Italy), Hans Clusmann (Department of Neurosurgery, Medical Faculty RWTH Aachen University, Aachen, Germany), Endre Czeiter [Department of Neurosurgery, University of Pécs, Pécs, Hungary; MTA-PTE Clinical Neuroscience MR Research Group and Janos Szentágotai Research Centre, University of Pécs, Hungarian Brain Research Program (Grant No. KTIA 13 NAP-A-II/8), Pécs, Hungary], Marek Czosnyka (Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK), Bart Depreitere (Department of Neurosurgery, University Hospitals Leuven, Leuven, Belgium), Ari Ercole (Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK), Shirin Frisvold (Department of Anaesthesiology and Intensive care, University Hospital Northern Norway, Tromsø, Norway), Raimund Helbok (Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria), Stefan Jankowski (Neuro-intensive Care, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK), Daniel Kondziella (Departments of Neurology, Clinical Neurophysiology and Neuroanaesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark), Lars-Owe Koskinen (Department of Clinical Neuroscience, Neurosurgery, Umeå University, Umeå, Sweden), Ana Kowark (Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany), David K. Menon (Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK), Geert Meyfroidt (Intensive Care Medicine, University Hospitals Leuven, Leuven, Belgium), Kirsten Moeller (Department Neuroanesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark), David Nelson (Department of Neurosurgery & Anaesthesia & intensive care medicine, Karolinska University Hospital, Stockholm, Sweden), Anna Piippo-Karjalainen (Helsinki University Central Hospital, Helsinki, Finland), Andreea Radoi (Department of Neurosurgery, Vall d'Hebron University Hospital, Barcelona, Spain), Arminas Ragauskas (Department of Neurosurgery, Kaunas University of Technology and Vilnius University, Vilnius, Lithuania), Rahul Raj (Helsinki University Central Hospital, Helsinki, Finland), Jonathan Rhodes (Department of Anaesthesia, Critical Care & Pain Medicine NHS Lothian & University of Edinburgh, Edinburgh, UK), Saulius Rocka (Department of Neurosurgery, Kaunas University of Technology and Vilnius University, Vilnius, Lithuania), Rolf Rossaint (Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany), Juan Sahuquillo (Department of Neurosurgery, Vall d'Hebron University Hospital, Barcelona, Spain), Oliver Sakowitz (Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany; Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany), Peter Smielewski (Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK), Nino Stocchetti (Department of Pathophysiology and Transplantation, Milan University, and Neuroscience ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milano, Italy), Nina Sundström (Department of Radiation Sciences, Biomedical Engineering, Umea University, Umea, Sweden), Riikka Takala (Perioperative Services, Intensive Care Medicine, and Pain Management, Turku University Central Hospital and University of Turku, Turku, Finland), Tomas Tamosiutis (Neuro-intensive Care Unit, Kaunas University of Health Sciences, Kaunas, Lithuania), Olli Tenovuo (Rehabilitation and Brain Trauma, Turku University Central Hospital and University of Turku, Turku, Finland), Andreas Unterberg (Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany), Peter Vajkoczy (Neurologie, Neurochirurgie und Psychiatrie, Charité - Universitätsmedizin Berlin, Berlin, Germany), Alessia Vargiolu (NeuroIntensive Care Unit, Department of Anaesthesia & Intensive Care, ASST di Monza, Monza, Italy), Rimantas Vilcinis (Department of Neurosurgery, Kaunas University of Health Sciences, Kaunas, Lithuania), Stefan Wolf (Department of Neurosurgery, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany), Alexander Younsi (Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany) and Frederick A. Zeiler (Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK; Section of Neurosurgery, Department of Surgery, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada).

Author Contributions

Conceptualisation: PG, LH and PE; methodology: PG and LH; data curation: PG, LH, AH, AL, GC and PE; formal analysis: PG, LH and AH; writing—original draft: PG, LH and PE; and writing—review and editing: PG, LH, PE, AL and GC. All authors contributed to the interpretation of the data.

Source of Support

Open access funding provided by Uppsala University. This study was supported by the Swedish Research Council (agreements no. 2022–06725 and no. 2018–05973), the Kjell och Mårta Beijer Foundation, the Swedish state under the agreement between the Swedish government and Uppsala county council, the ALF-agreement and the European Union 7th Framework programme (EC grant 602,150).

Conflicts of Interest

The authors declare no conflicts of interest, except G.C., who reviews grants and personal fees as Speakers' Bureau and Advisory Board member from Integra and NeuroOptics, all outside the submitted work.

Ethical Approval/Informed Consent

The CENTER-TBI study (EC grant 602150) has been conducted in accordance with all relevant laws of the EU if directly applicable or of direct effect and all relevant laws of the country where the recruiting sites were located, including but not limited to, the relevant privacy and data protection laws and regulations (the "Privacy Law"), the relevant laws and regulations on the use of human materials and all relevant guidance relating to clinical studies from time to time in force including, but not limited to, the ICH Harmonised Tripartite Guideline for Good Clinical Practice (CPMP/ICH/135/95) ("ICH GCP") and the World Medical Association Declaration of Helsinki entitled "Ethical Principles for Medical Research Involving Human Subjects". Informed consent by the patients and/or the legal representative/next of kin was obtained, accordingly to the local legislations, for all patients recruited in the Core Dataset of CENTER-TBI and documented in the e-CRF. Ethical approval was obtained for each recruiting site. The list of sites, Ethical Committees, approval numbers and approval dates can be found on the website: <https://www.center-tbi.eu/project/ethical-approval>. Informed consent was obtained from all participants/waived due to the retrospective design. This is a retrospective study and for this type of study, formal consent is not required. The study was approved for UHF-TBI data by the local and the Swedish National Ethical Review Authority. Informed consent was obtained from patient or from their next of kin if the patient was unable to provide consent. It was also approved to include patients treated during the study period from whom consent not had been possible to obtain, for example, from deceased patients or when relatives had not been possible to reach.

Artificial Intelligence

Artificial intelligence (AI) was used solely to create one figure.

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/>.

Publisher's Note

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

Received: 21 April 2026 Accepted: 23 July 2026

Published online: 14 October 2026

References

1. Maas AIR, Menon DK, Manley GT, et al. Traumatic brain injury: progress and challenges in prevention, clinical care, and research. *Lancet Neurol*. 2022;21(11):1004–60. [https://doi.org/10.1016/S1474-4422\(22\)00309-X](https://doi.org/10.1016/S1474-4422(22)00309-X).
2. Rajpurkar P, Chen E, Banerjee O, et al. AI in health and medicine. *Nat Med*. 2022;28(1):31–8. <https://doi.org/10.1038/s41591-021-01614-0>.
3. Palermo M, D'Arrigo S, Olivi A, et al. Machine learning models for predicting vasospasm following ruptured intracranial aneurysms: a systematic review and meta-analysis. *Acta Neurochir (Wien)*. 2025;167(1):314. <https://doi.org/10.1007/s00701-025-06725-y>.
4. Hinds S, Robertson C, Chen J, et al. High-frequency physiological measures predict post-admission surgical intervention after severe traumatic brain injury. *J Neurotrauma*. 2025. <https://doi.org/10.1177/08977151251395740>.
5. Van Hal ST, Van Der Jagt M, Van Genderen ME, et al. Using artificial intelligence to predict intracranial hypertension in patients after traumatic brain injury: a systematic review. *Neurocrit Care* [Online ahead of print]; <https://doi.org/10.1007/s12028-023-01910-2>.
6. McNamara R, Meka S, Anstey J, et al. Development of traumatic brain injury associated intracranial hypertension prediction algorithms: a narrative review. *J Neurotrauma*. 2022. <https://doi.org/10.1089/neu.2022.0201>.
7. Galos P, Hult L, Zachariah D, et al. Machine learning based prediction of imminent ICP insults during neurocritical care of traumatic brain injury. *Neurocrit Care*. 2025;42(2):387–97. <https://doi.org/10.1007/s12028-024-02119-7>.
8. Steyerberg EW, Wiegers E, Sewalt C, et al. Case-mix, care pathways, and outcomes in patients with traumatic brain injury in CENTER-TBI: a European prospective, multicentre, longitudinal, cohort study. *Lancet Neurol*. 2019;18(10):923–34. [https://doi.org/10.1016/S1474-4422\(19\)30232-7](https://doi.org/10.1016/S1474-4422(19)30232-7).
9. Carney N, Totten AM, O'Reilly C, et al. Guidelines for the management of severe traumatic brain injury. Fourth Edition *Neurosurg*. 2017;80(1):6–15. <https://doi.org/10.1227/NEU.0000000000001432>.
10. Howells TP, Jones PA, Piper IR, Souter MJ, Miller JD. Design of a research database for the study of secondary insults following head injury. *J Neurotrauma*. 1995;12:471.
11. Galos P, Nyholm L, Howells T, et al. Capturing of intracranial pressure treatment during neurointensive care in patients with acute brain injury using a novel tablet-based method. *J Clin Monit Comput*. 2022. <https://doi.org/10.1007/s10877-022-00820-6>.
12. Van Rossum G, Drake FL. Python 3 reference manual. Scotts Valley, CA: CreateSpace; 2009.
13. Riparbelli AC, Capion T, Möller K, et al. Critical ICP thresholds in relation to outcome: is 22 mmHg really the answer? *Acta Neurochir (Wien)*. 2024;166(1):63. <https://doi.org/10.1007/s00701-024-05929-y>.
14. American College of Surgeons Committee on Trauma. Best practices guidelines: the management of traumatic brain injury. Chicago, IL: American College of Surgeons; 2015.
15. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng*. 2010;22(10):1345–59. <https://doi.org/10.1109/TKDE.2009.191>.