Artificial Intelligence-Enhanced Neurocritical Care for Traumatic Brain Injury: Past, Present and Future

Kyung Ah Kim,1 Hakseung Kim,1 Eun Jin Ha,2 Byung C. Yoon,3 Dong-Joo Kim1,4
Department of Brain and Cognitive Engineering,1 Korea University, Seoul, Korea
Department of Critical Care Medicine,2 Seoul National University Hospital, Korea
Department of Radiology,3 Stanford University School of Medicine, VA Palo Alto Health Care System, Palo Alto, CA, USA
Department of Neurology,4 Korea University College of Medicine, Seoul, Korea

Running title: AI-Enhanced Neurocritical Care for TBI

• Received : September 6, 2023 • Revised: December 18, 2023 • Accepted : January 4, 2024

Address for correspondence : Dong-Joo Kim
Department of Brain and Cognitive Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea
Tel : +82-2-3290-5929, Fax : +82-2-3290-3970, E-mail : dongjookim@korea.ac.kr, ORCID : https://orcid.org/0000-0002-0988-2236

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Abstract

In neurointensive care units (NICU), particularly in cases involving traumatic brain injury (TBI), swift and accurate decision-making is critical because of rapidly changing patient conditions and the risk of secondary brain injury. The use of artificial intelligence (AI) in NICU can enhance clinical decision support and provide valuable assistance in these complex scenarios. This article aims to provide a comprehensive review of the current status and future prospects of AI utilization in the NICU, along with the challenges that must be overcome to realize this. Presently, the primary application of AI in NICU is outcome prediction through the analysis of pre-admission and high-resolution data during admission. Recent applications include augmented neuromonitoring via signal quality control and real-time event prediction. In addition, AI can integrate data gathered from various measures and support minimally invasive neuromonitoring to increase patient safety. However, despite the recent surge in AI adoption within the NICU, the majority of AI applications have been limited to simple classification tasks, thus leaving the true potential of AI largely untapped. Emerging AI technologies, such as generalist medical AI and digital twins, harbor immense potential for enhancing advanced neurocritical care through broader AI applications. If challenges such as acquiring high-quality data and ethical issues are overcome, these new AI technologies can be clinically utilized in the actual NICU environment. Emphasizing the need for continuous research and development to maximize the potential of AI in the NICU, we anticipate that this will further enhance the efficiency and accuracy of TBI treatment within the NICU.

Key Words: Artificial intelligence · Clinical decision support · Intensive care · Machine learning · Neurocritical care · Traumatic brain injury.

INTRODUCTION
Traumatic brain injury (TBI) is one of the leading causes of death and disability worldwide, with the estimated, global incidence of 27–69 million cases annually. TBI can result in a range of progressive long-term physical and psychosocial impairments, including difficulties with cognitive abilities such as attention and memory, neurological symptoms such as headaches and dizziness, neuropsychiatric disorders, and an increased likelihood of neurodegenerative diseases later in life.

TBI can be classified into two types: primary and secondary. The former refers to the immediate consequences of the initial trauma, and as such, these effects are irreversible. Secondary brain injury refers to a series of events triggered by physiological responses following initial damage that may result in additional brain injury and worsen patient prognosis. Therefore, the primary objective of TBI management in neurointensive care units (NICUs) is to prevent, evaluate, and treat these secondary insults. This involves implementing measures such as maintaining sufficient oxygenation and blood pressure, controlling intracranial pressure (ICP) and cerebral perfusion pressure (CPP), and providing appropriate nutritional support.

Despite major advances in our understanding of TBI, numerous challenges persist in managing this complex and heterogeneous condition. One of the primary challenges arises from the heterogeneity of injury mechanisms, patient characteristics, and injury severity. As a result, there is a pressing need for patient-specific treatments. To determine the most suitable treatment strategies, it is essential to take into account the degree of TBI severity, specific areas of the brain affected, and unique physiological responses. Artificial intelligence (AI) has the potential to play a critical role in meeting these needs by providing tools for personalized treatment plans and decision support. Particularly in the management of TBI, numerous studies are being conducted, exploring the diverse application possibilities of medical AI. It can perform tasks such as classification, prediction, and segmentation of TBI patients, thereby offering more precise diagnoses and efficient treatment directions. Table 1 lists some of these studies, demonstrating the specific roles AI can play in TBI management (Table 1). This table summarizes various research examples on how AI can contribute in phases like initial management, NICU management, and neuroprognostication for TBI patients, thus providing insights into the clinical
applicability of AI.

AI-enhanced neurocritical care has the potential to significantly enhance the treatment outcomes and quality of life for patients with TBI. Although AI techniques have rapidly gained acceptance within the neurosurgical community, the same level of acceptance does not extend to the utilization of AI for TBI management. This brief review provides the core concepts of machine learning (ML) and an overview of ML applications in NICUs with a focus on patients with TBI. It further explores the use of real-time AI in the NICU environment and concludes by investigating the prospective roles of generalist medical AI (GMAI) and digital twins (DT) in the NICU.

MACHINE LEARNING FOR NEUROCritical CARE DATA ANALYSIS

Core concepts of machine learning

ML is a subfield of AI that enables computers to identify patterns and rules inherent in data without being explicitly programmed. A key characteristic of ML models is their ability to continuously improve their performance through experience. Although more recent ML techniques, such as reinforcement learning and generative learning, have yielded promising results in certain areas (e.g., surgical planning and generating neuroimages, etc.), the conventional tasks of medical AI have primarily focused on two main approaches: unsupervised learning for identifying undefined patterns or reducing data dimensionality and supervised learning for classification and regression tasks. While both techniques can be used concurrently, the primary application of ML in medical fields has been centered around supervised classification and regression for image classification and segmentation, diagnosis, and prediction of outcome and complications.

While it is true that most of these ‘conventional’ tasks could theoretically be achieved through statistical methods, such statistical models rely on specific assumptions regarding data distribution, which may not be valid in numerous situations. Specifically, if the data have high dimensionality
(i.e., many variables) and/or if the relationship between data points is nonlinear, statistical assumptions are likely to fail. Nonlinearity and high dimensionality are common in medical data, making ML particularly advantageous. Logistic regression (LR), Naïve Bayes (NB), Decision Trees, K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM) are commonly used classification algorithms in the medical field\cite{6,11,33,123}. Recently, the utilization of ML has been increasing in the medical field, not only in clinical practice but also in personalized treatment\cite{70,110}. This trend has been further accelerated by the emergence of various decision tree variants and deep learning technologies\cite{9,44,63,68,113,142}.

**Role of machine learning in pre-admission data analysis**

It is crucial for patients with TBI to predict the risk factors and perform early diagnosis during the prehospitalization process\cite{34}. In addition, it is essential to optimize the triage process by appropriately classifying patients based on the severity of brain injury for treatment or transportation. In this regard, ML may enable a rapid and accurate diagnosis and treatment planning for patients with TBI.

The early recognition of TBI is particularly important because it effectively improves patient outcomes\cite{35,134,137}. Choi et al.\cite{19} used pre-hospital information from emergency medical service (EMS) providers, such as the EMS ambulance run sheet and EMS trauma in-depth registry, to identify the risk factors. Through the application of various ML techniques, including the elastic net algorithm, which exhibited the best performance, we identified that the three most significant predictors of TBI-related outcomes were loss of consciousness, Glasgow Coma Scale (GCS) score, and light reflex. In a study conducted by Hale et al.\cite{40}, an artificial neural network (ANN) using Computed Tomography (CT) data was employed to predict the likelihood of clinically relevant TBI in pediatric patients aged < 18 years who experienced non-penetrating head trauma.

Accurately classifying patients with TBI according to their trauma severity before or immediately after transfer to the emergency room is crucial for efficient early treatment and has a significant impact on
their prognosis\textsuperscript{1).} Abe et al.\textsuperscript{1)} aimed to detect traumatic intracranial hemorrhage (tICH) and differentiate patients with tICH from those without tICH using information collected by on-site EMS personnel, such as systolic blood pressure, heart rate, body temperature, and respiratory rate. In the testing set, extreme gradient boosting (XGBoost), a variant of decision tree, showed the best predictive performance among the various ML algorithms with an area under the ROC curve (AUC) of 0.80. Moyer et al.\textsuperscript{84)} developed a predictive model to identify patients requiring emergency neurosurgery among those diagnosed with moderate-to-severe TBI during the prehospital assessment process. Fifteen prehospital predictors, including the Glasgow Coma Scale (GCS), initial systolic blood pressure, initial diastolic blood pressure, and initial oxygen saturation, were used. The CatBoost model, another decision tree variant, outperformed the other models with an AUC of 0.81.

\textbf{Optimizing surgical timing with ML.}

Once TBI patients are admitted to the hospital, it is important to quickly identify which patients require surgical intervention and to intervene at the right time. ML can contribute to this treatment intervention process, helping medical professionals make better decisions. A number of studies have been conducted to classify the severity of patients and identify patients who need surgical intervention. For instance, Güler et al.\textsuperscript{38)} developed a classification model to rapidly evaluate the severity of TBI in emergency department admitted patients. An ANN model was used to classify the severity of TBI using trauma scores, GCS, and electroencephalography (EEG) data. The model showed high agreement with the neurologists' decisions, with an 87\% concordance rate. Habibzadeh et al.\textsuperscript{39)} focused on those with positive initial cranial CT scan results for moderate TBI, developing a model to differentiate between patients needing neurosurgical intervention and those who do not, based on age, gender, GCS, Marshall score, presence of hematoma, and midline shift score. The SVM model performed best with an F1 score of 0.83 and an AUC of 0.93. Currently, TBI patients take follow-up CT scans after initial scans to assess intracranial damage progression, leading to concerns over unnecessary radiation exposure. In this context, this research shows potential as an effective decision-making tool for predicting intracranial
damage and identifying patients in need of medical intervention without follow-up CT scans.

ML is not just limited to patient classification, but also plays an important role in determining the appropriate time for medical intervention according to the individual patient's situation\textsuperscript{46,55,76,125}. Some studies explore the general consequences of delayed neurosurgical intervention regardless of type or patient situation of the surgery, while others focus on detailing the appropriate time of intervention for specific surgical procedures and situations. Hanko et al.\textsuperscript{41} predicted mortality and functional outcomes for TBI patients who received primary Decompressive Craniectomy (DC) treatment, considering a variety of variables including medical records, clinical symptoms, and pre- and post-operative conditions. The RF-based predictive model used in this study achieved an AUC of 0.811 for 6-month mortality and 0.873 for outcomes, with surgery timing being an important variable. It concluded that early implementation of primary DC is essential for effectively relieving ICP and ongoing brain hemorrhage. Similarly, a study by Seelig et al.\textsuperscript{109} examined the outcomes of patients in a coma due to traumatic acute subdural hematoma, investigating various variables. It found that delay in surgery was a significant factor, particularly noting that patients operated within 4 hours of injury had a 30% mortality rate, while those operated after 4 hours had a mortality rate rising to 90%.

**Outcome prediction via machine learning**

Seemingly moderate or minor cases of TBI can rapidly escalate to severe TBI, necessitating close surveillance\textsuperscript{119}. The generated variables and parameters offer valuable insights into prognostic assessments. Indeed, outcome prediction after TBI has long been a major topic in neurocritical care\textsuperscript{72}, and the increasing availability of ML techniques has stimulated their active incorporation into this topic. Tu et al.\textsuperscript{129} developed a predictive model for mortality risk using the Taiwan Triage and Acuity Scale (TTAS) and 11 characteristic variables, including age, sex, and obesity. The LR algorithm exhibited the best predictive performance with an AUC of 0.925. Another study by Thara and Thakul\textsuperscript{124} applied ML to predict functional outcomes in pediatric patients with TBI. GCS, hypotension, pupillary light reflex, and subarachnoid hemorrhage were established as important factors for predicting functional outcomes.
Among the classification algorithms, SVM showed the highest performance, with an accuracy of 0.94. Finally, Farzaneh et al.\textsuperscript{32}) proposed an interpretable ML-based framework for predicting long-term functional outcomes in patients with TBI. This study implemented the XGBoost classifier using 18 Electronic Health Record (EHR) variables. The final model achieved an accuracy of 0.7488 on the test set. The advantage of this study is that it provides SHAP (SHapley Additive exPlanations) contribution, allowing for easier interpretation of the model in medical decision-making, particularly for predicting long-term functional outcomes in patients with TBI.

**Machine learning for high-resolution data analysis during intensive care**

Management of severe TBI often necessitates continuous monitoring of various signals such as intracranial pressure, arterial pressure, and electrocardiography (ECG). The sheer scale, nonlinearity, and high dimensionality of the resulting data warrant analysis using ML. ML has undoubtedly demonstrated great potential for analyzing continuous neuromonitoring data from patients with TBI, allowing for the prediction of expected outcomes\textsuperscript{116,119,126}.

First, EEG is actively used to predict outcomes in patients with TBI\textsuperscript{7,81}). For example, Haveman et al.\textsuperscript{42}) used quantitative electroencephalography (qEEG) to predict the outcomes of patients with moderate-to-severe TBI 12 months after injury. This was accomplished by creating predictive models using an RF classifier based on qEEG features, age, and mean arterial blood pressure (MAP), which were collected over seven days post-admission. The model showed an AUC of 0.94 on the training set and 0.81 on the validation set. Similarly, Noor et al.\textsuperscript{88}) used the absolute power spectral density (PSD) in each EEG frequency band to predict the outcomes of patients with moderate TBI 12 months after injury. In this study, Random Under-Sampling Boosted Trees (RUSBoost) were used. It was found that the absolute PSD in the delta and gamma bands was the best outcome predictor. They demonstrated an AUC of 0.97 and 0.95, respectively.

In addition, ICP signal is a valuable tool that can be used to predict patient prognosis. Raj et al.\textsuperscript{96})
developed a dynamic mortality prediction model for patients with TBI in the NICU based on ICP and MAP. Dynamic features related to ICP, MAP, CPP, motor response, and eye response were generated. These were then used in a LR model to predict the mortality rate at 30 days post-injury. The algorithm achieved a maximum AUC of 0.84, demonstrating superior performance compared to IMPACT-TBI\textsuperscript{118}, which is one of the most widely accepted static prediction models with an AUC of 0.78. Pimentel et al.\textsuperscript{93} also attempted to predict the mortality rate of patients with TBI by extracting dynamic features from ICP and MAP signals. A Gaussian Process (GP) framework was used to extract dynamic features from physiological signals. This method has the advantage of performing well under intermittently disrupted monitoring conditions, which are common in the NICU. By combining GP-based and pressure reactivity index (PRx) features, the ability to predict patient outcomes significantly improved, achieving an AUC of 0.76.

REAL-TIME AI APPLICATIONS IN THE NICU ENVIRONMENT

In complex and unpredictable environments such as the NICU, the speed and accuracy of medical decisions are critical\textsuperscript{102,112}. The issue of rapidity and accuracy in the clinical decision-making process is further compounded by the sheer amount of data generated during neuromonitoring and altered levels of consciousness. The Clinical Decision Support System (CDSS)\textsuperscript{85} is intended to improve healthcare delivery by enhancing medical decisions with targeted clinical knowledge, patient information, and other health information\textsuperscript{90}. AI plays a variety of roles in CDSS, such as monitoring and analyzing a patient’s condition in real time and predicting future events. AI helps clinicians to quickly recognize situations and make appropriate medical decisions by analyzing large amounts of data in a complex environment.

Signal quality control
False alarms are prevalent in the NICU, adversely impacting staff alarm responsiveness, degrading the quality of care, and indirectly affecting patients by disrupting their sleep quality \(23,25,60,127\). Because of the complexity of neurophysiological signals, simple filter-based artifact rejection techniques are often insufficient. Thus, several studies have proposed methods to minimize false alarms by identifying and removing artifacts that are the main cause of false alarms or by improving alarm accuracy \(4,5,63,65,113,128,132\) (Table 2). The primary benefit of artifact elimination would be improved signal quality and a subsequent reduction of false alarms, and signal quality control may lead to significant changes in the prognostic values of essential neurophysiological signals. For instance, Kim et al.\(^{54}\) reported that the association between worse outcomes and the prevalence of hypotension significantly increased after the elimination of artifacts in arterial pressure.

**Event prediction**

Intracranial hypertension (IH) is common during the acute phase of TBI and can worsen secondary insult\(^{95}\). The prediction of IH could allow rapid intervention and, thus, result in a better outcome; hence, it has long been considered an important topic in the neurocritical care fields\(^{12,62}\). Schweingruber et al.\(^{108}\) reported the performance of long short-term memory (LSTM) models in predicting significant events at various time intervals ranging from 1 to 24 h. These models were trained using demographic variables and were continuously monitored for arterial blood pressure (ABP) and ICP. The model showed robust performance even at a distance of 24 h, with an AUC of 0.826. However, the most clinically relevant prediction was made 2 h before the significant event, achieving an AUC of 0.953. Additionally, this study defined two phases of IH: ‘long phase IH,’ lasting more than 2 h with IH above 22 mmHg, and ‘short phase IH,’ lasting less than 2 h. The LSTM model predicted the occurrence of both situations at a distance of 2 h, resulting in AUC of 0.953 and 0.728 for the long and short phases, respectively. Similarly, Güiza et al.\(^{37}\) developed a model that could predict increased ICP episodes lasting > 10 min at 30 mmHg, 30 min in advance, using only the dynamic characteristics of continuous ICP and MAP. The prediction model was built using multivariate LR and a GP, achieving a classification
accuracy of 77%, sensitivity of 82%, and specificity of 75% on the validation set. In addition to predicting simple IH events, Lee et al.\textsuperscript{62} focused on predicting ‘life-threatening’ intracranial hypertension (LTH) specifically. In this study, episodes of IH were classified as LTH if they were characterized by a PRx exceeding 0 and a pressure-time dose greater than 5. This study aimed to predict whether ongoing IH conditions would escalate to LTH within 5 min. CatBoost was used to predict using ABP and ICP-related parameters. This model achieved an AUC of 0.73 for predicting LTH on the test dataset. Furthermore, numerous studies have developed models using various variables and algorithms to predict ICP episodes early\textsuperscript{16,51,74,107,136} (Figure 1).

Minimally invasive neuromonitoring

Simple clinical assessments may not detect subtle changes in brain physiology or alterations in the neurological status that could appear at a later stage\textsuperscript{115}. This is particularly true for TBI, where the patient’s condition may rapidly deteriorate, necessitating the integration of clinical assessments with other neuromonitoring techniques\textsuperscript{112}. Even ICP (and CPP to some extent), which has conventionally served as a core measure in TBI, has yielded conflicting results concerning its prognostic value\textsuperscript{18,138}. Over the past decade, multimodal neuromonitoring has garnered increasing attention. For example, studies have shown that integrating brain tissue oxygen tension (PbtO2)-based treatments with conventional ICP/CPP-oriented protocols leads to a better neurological prognosis and lower death rates\textsuperscript{57,100}. Cerebral microdialysis can also provide a better understanding of hypoperfusion and metabolism when complemented by other modalities\textsuperscript{73}. Thus, multimodal neuromonitoring provides a more accurate understanding of each patient's individual circumstances and conditions, paving the way for the delivery of individualized therapy\textsuperscript{73}. However, most multimodal monitoring techniques are invasive, resulting in a range of problems. For example, invasive ICP monitoring using a ventricular catheter carries the risk of numerous complications including infection and intracranial hemorrhage\textsuperscript{31,112}. Cerebral microdialysis is not an exception to these concerns as it has additional problems such as tissue inflammatory response and biofouling\textsuperscript{115}. Therefore, many previous studies
have investigated neuromonitoring using minimally invasive or non-invasive methods\textsuperscript{3,106,131}. In addition, with the development of AI, it has become possible to non-invasively acquire and analyze accurate signals.

Various technologies spanning fluid dynamics, ophthalmic, otic, and electrophysiologic fields have been proposed for non-invasive ICP monitoring, including magnetic resonance imaging (MRI), transcranial Doppler ultrasound imaging, optic nerve sheath diameter assessment, and others. One of the most traditional methods for non-invasive ICP estimation is the use of cerebral blood flow velocity (CBFV) and ABP\textsuperscript{47,50,52}. According to Imaduddin et al.\textsuperscript{47}, a Pseudo-Bayesian Model was used to estimate ICP using non-invasive measurements such as ABP and CBFV. This model showed a mean absolute error (MAE) of 0.6 mmHg for mean ICP estimation, compared to measurements taken invasively. Similarly, Kashif et al.\textsuperscript{52} developed a patient-specific model-based algorithm to estimate ICP using routinely collected ABP and CBFV signals, eliminating the need for calibration. Using 35 hours of data from 37 TBI patients, the model achieved ICP estimates with a MAE of 1.6 mmHg and standard deviation of 7.6 mm Hg. The best performance of the model was a sensitivity of 90%, a specificity of 80%, and an AUC of 0.88, indicating improved accuracy and reduced variability compared to previous studies\textsuperscript{89}. Recently, attempts have been made to estimate the ICP using electrocardiogram (ECG). In 2021, Sadrivi et al.\textsuperscript{105} conducted a study using ECG and a U-Net-based deep convolutional autoencoder (MA-UDCAE) system to noninvasively estimate the mean ICP. The mean ICP evaluation had an MAE of 2.404 ± 0.043 mmHg, and Pearson’s linear correlation coefficient was 0.89. Although the results were inferior to those of previous studies that estimated ICP using CBFV and ABP, they are still valuable because they use completely non-invasive ECG signals.

Unfortunately, there are a number of challenges to applying those noninvasive methods in clinical practice. Almost all review articles on this topic have concluded that it is impossible to estimate ICP noninvasively in all situations, for all patients, and over a long period of time without calibration.\textsuperscript{31,58,94,99,103,141}. For instance, CBFV not only exhibits high intra- and inter-user variability but is also significantly affected by other physiological changes such as medications, and autoregulation. These
factors compromise the reliability of CBFV as a source signal, furthermore, making it challenging for long-term continuous use. Additionally, other non-invasive ICP monitoring measures also have distinct limitations such as insufficient validation, inability for continuous monitoring, poor signal quality, variations due to different operator expertise, and inter-observer variability. For example, neurological pupillary index measurement using pupillometry has currently attracted much attention in research and clinical practice\(^{17,48,91}\), but there is still controversy about its suitability for continuous monitoring, accuracy in reflecting changes in ICP, and precision in estimating ICP\(^{69,101,117}\).

Despite these limitations, noninvasive methods of measuring ICP can be a viable and attractive alternative for patients who are not amenable to invasive monitoring. Continuous improvements and validations are being made in non-invasive methods, and particularly, the development of multimodal non-invasive neuromonitoring technologies that integrate various methods is expected to significantly reduce existing limitations. The advancement of such technologies will mark an important turning point in non-invasive ICP measurement.

**AI-ENHANCED NEUROCRITICAL CARE IN THE NEAR FUTURE**

In neurocritical care, clinical decisions require careful consideration of the vast amount of information generated by neuromonitoring and patient management. In this endeavor, several decision support AI applications have been proposed; at present, AI applications have been designed and utilized for relatively limited tasks such as patient monitoring, event prediction, and prognosis forecasting, that is, simple classification tasks\(^{130}\). Furthermore, most, if not all, existing AI applications are designed to analyze signals, images, or tabular data separately without integration. However, with rapid advancements in AI technology, newer and more advanced approaches have emerged. Generalist medical AI (GMAI) and DT technology are two such examples that hold promise for revolutionizing the landscape of neurocritical care. Although somewhat conceptual, these technologies have the
potential to significantly enhance the AI capabilities in this field.

Concepts of digital twins and generalist medical AI (GMAI)

A DT is a virtual replica of a physical system or object and is a technological concept for efficiently testing and simulating products or processes in a digital environment\textsuperscript{61,122}. In the medical field, DT can be used as real-time digital replicas of patients to develop effective treatment strategies. Additionally, it could be useful for optimizing medical processes and hospital management strategies\textsuperscript{21}.

The first application of DT in the medical field was the Archimedes project in 2003, where a DT approach was used and validated to model the complex management of diabetes\textsuperscript{27}. Later, Lal et al.\textsuperscript{59} developed a DT model of intensive care unit (ICU) patients to predict the therapeutic response during the first 24 h of sepsis and conducted pilot testing in an actual ICU. In this study, various technologies such as agent-based modeling, discrete event simulation, and Bayesian networks were used to visualize major organs, including the nervous and cardiovascular systems, and simulate treatment effects. The pilot testing of the DT showed a moderate level of agreement (fair level, Cohen's kappa = 0.41) for the first response and a significantly high level of agreement (good level, Cohen's kappa = 0.65) for the second response. In recent years, DT models have been developed in other medical fields, such as cardiology and endocrinology, to provide tailored treatment for patients\textsuperscript{20,24,30}.

The concept of the GMAI stems from the shared limitations of existing medical AI models, which are designed to perform specific tasks and do not generalize to new tasks or datasets\textsuperscript{36,43,92}. GMAI aims to address the limitations of traditional medical AI models and relies on a foundation model pretrained on large amounts of data to flexibly perform various tasks\textsuperscript{83}. Conceptually, the GMAI can conduct diverse medical tasks without altering the trained network, thereby ensuring robustness and versatility\textsuperscript{66}. According to Moor et al.\textsuperscript{83}, the tremendous potential of the GMAI can support clinical decision-making by drafting radiology reports considering patient history, summarizing real-time patient data, and complex electronic health records to predict future patient status or compare treatment options. In
addition, interactive notetaking would be possible by utilizing the development of voice-to-text models to draft reports based on conversations with clinicians; this advantage of GMAI could also be exploited as chatbots for individual patients.

Promises of digital twins and GMAI in TBI

With DT and GMAI, clinicians may virtually navigate treatment options, crafting personalized medical plans and rehabilitation programs. Furthermore, patient-friendly chatbots and personalized reports can help patients intuitively understand their own condition, enabling them to participate more actively in their healthcare, treatment and recovery processes. The vast spectrum of symptoms and outcomes in TBI patients, driven by the injury's cause, location, and severity, demands personalized treatment approaches – a long-standing challenge in this field. This heterogeneity in TBI has long been recognized as a significant challenge. By leveraging data on a patient's demographic characteristics, risk factors, and injury patterns, a DT model can be built to aid in the establishment of personalized treatment plans. McIver proposed a model that creates patient-specific finite element meshes based on patients' MRI data. The model, by using DT technology, reflects the structure and function of an individual's brain, and can simulate the effects of various types of head trauma. In addition, it enables proactive identification and prediction of patients who may suffer severe long-term damage from repeated head impacts, contributing to prevention and effective management strategies for TBI. Advances in GMAI may potentially allow for the expansion of AI's application beyond tasks such as prediction and classification. For instance, by integrating medical imaging with clinical information, GMAI can create integrated radiological reports that include not only predictive analysis but also recommendations for personalized treatment and surgical plans for clinicians. Based on this information, chatbot services can be provided to patients, offering easy-to-understand consultations and personalized health education materials.

DT and GMAI can also be used to monitor rehabilitation status and design personalized rehabilitation programs by comprehensively considering a patient's medical history, cognitive status, and motor skills.
For example, Huanxia⁴⁵) presented a method to monitor the gait of patients with gait disorders in real time using a powered exoskeleton and DT technology. The powered exoskeleton supports the patient's movement, while DT analyzes gait patterns by digitally replicating the patient's actual physical state. This system is capable of providing fall alerts and evaluating the effectiveness of rehabilitation, thereby aiding clinicians in implementing appropriate interventions based on the patient's condition.

**Pitfalls and remaining challenges**

The combination of DT and GMAI could offer significant benefits to both clinicians and patients by supporting medical decision-making at every step and directing personalized care effectively. However, despite their tremendous potential, several challenges must be addressed to implement and utilize these technologies in practice.

**Data scale**

To develop models that can be utilized for various purposes in the medical field, a wide range of diverse and huge datasets are needed, as well as standardized specifications for signals, devices, and other components to ensure data uniformity. Thus, collaboration between different centers, guidelines, and unified storage methods for data standardization is necessary. Additionally, even when data are available, there are concerns about the substantial costs of model training as the size of the foundation model for DT and GMAI increases, as well as the potential environmental concerns associated with it.

**Privacy and security**

Medical AI handles sensitive personal information related to patient health. Specifically, the DT and GMAI models require a substantial amount of personal information to achieve high performance. DT represents a virtual replica of an individual, and security becomes an even more critical issue. Hence,
relevant data must be collected, stored, and processed in a secure and privacy-compliant manner.

Validation

Because of the complexity of the DT and GMAI models, validation poses significant challenges. Unlike the current task-specific AI models, these models perform diverse tasks, making it highly challenging to anticipate all possible scenarios. In addition, given that the DT and GMAI models directly affect patient health when used in the medical field, the validation process is especially important. Thus, diverse and precise validation procedures are required.

Application

Integration of AI into real medical practice, especially in the NICU for TBI patients, presents significant challenges. First, one of the major challenges for AI applications in clinical settings is the issue of compatibility between software and hardware. An algorithm optimized for a specific system may not function properly in another system. For example, an algorithm trained on a particular MRI scanner device may not work well on a different brand or model of scanner. Second, medical AI is still too immature to function effectively especially, in complex medical situations such as TBI. The inherent learning curve of AI demands a significant amount of data and information for actual clinical application; thus, it would take some time for AI to gain a sufficient level of clinical proficiency. While it is crucial to evaluate and validate the performance of algorithms, such as accuracy, it is more important to validate the compatibility, effectiveness, stability, and other related aspects of algorithms when applied in actual clinical settings. Such validation necessitates close collaboration between doctors, computer scientists, and data managers. This will give us a deep understanding of the clinical value of medical AI technology. Finally, until all of these medical AI validation processes are completed, clinicians should avoid over-reliance on AI decisions and always keep in mind the possibility of errors or misdiagnosis.
CONCLUSION

The application of AI, particularly ML, and concepts such as DT and GMAI hold substantial promise for improving neurocritical care for patients with TBI. It has potential across a wide range of applications from prognostication, decision support, patient specific treatment planning, and even rehabilitation. However, although these opportunities are significant, challenges remain, including the need for high-quality data, privacy concerns, ethical considerations, and further validation of these technologies. Despite these hurdles, the future of AI in neurocritical care seems bright, and continuous research and development in this field is necessary to unlock its full potential.

AUTHORS' DECLARATION

Conflicts of interest

No potential conflict of interest relevant to this article was reported.

Informed consent

This type of study does not require informed consent.

Data sharing

None

Preprint
None

**ORCID**

Kyung Ah Kim  https://orcid.org/0009-0003-3472-9991
Hakseung Kim  https://orcid.org/0000-0001-5407-293X
Eun Jin Ha  https://orcid.org/0000-0003-3278-0550
Byung C. Yoon  https://orcid.org/0000-0003-2929-5677
Dong-Joo Kim  https://orcid.org/0000-0002-0988-2236

**Acknowledgements**

This work was supported by a National Research Foundation of Korea (NRF) Grant funded by the Korean government (Ministry of Science and ICT, MSIT) (No. 2022R1A2C1013205).

**References**


31. Evensen KB, Eide PK: Measuring intracranial pressure by invasive, less invasive or non-
invasive means: limitations and avenues for improvement. **Fluids and Barriers of the CNS** 17 : 1-33, 2020


Maas MB, Naidech AM, Batra A, Chou SH-Y, Bleck TP: Comment on “Can Quantitative Pupillometry be used to Screen for Elevated Intracranial Pressure? A Retrospective Cohort Study”. *Neurocritical Care* 37: 597-598, 2022


78. McIver KG: The Application of High-Performance Computing to Create and Analyze Simulations of Human Injury: Purdue University Graduate School, 2022


92. Pesapane F, Codari M, Sardanelli F: Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. European radiology experimental 2: 1-10, 2018


94. Popovic D, Khoo M, Lee S: Noninvasive monitoring of intracranial pressure. Recent Patents on Biomedical Engineering (Discontinued) 2: 165-179, 2009


119. Stocker RA : Intensive care in traumatic brain injury including multi-modal monitoring and neuroprotection. Medical Sciences 7 : 37, 2019


noninvasive means of measuring intracranial pressure: a review. *Physiological measurement* 38: R143, 2017


143. 공은정, 방창석: 인공지능을 이용한 의료영상 판독의 현재와 미래. *대한소화기학회지* 82: 43-46, 2023
Figure Legends

**Fig. 1.** A visual representation of the entire process of the ICP related event prediction model. (A) Signal data such as ABP, ICP, and ECG, and EHR data such as personal information, medical history, and radiology reports are collected from patients during their hospitalization. (B) Before using the collected data as model input, preprocessing is required. Signal data is filtered according to predefined ranges and segmented into windows of specific sizes. Additionally, considering the quality of signal, noise is removed. Non-signal data from patients undergoes processes such as encoding, normalization, and imputation of missing data, depending on the type of data. (C) An appropriate model is selected and trained using the processed ICP-related signal parameters and patient health-related parameters. (D) This model, by utilizing signal data from specific window and preprocessed EHR data, can predict life-threatening clinical conditions, such as IH, within a defined time interval. (E) In a clinical setting, such predictions can assist clinicians in making decisions for early interventions like drug treatment, surgery,
and supportive care. ABP, arterial blood pressure; ICP, intracranial pressure; ECG, electrocardiogram; EHRs, electronic health records; BMI, body mass index; CT, computed tomography; MRI, magnetic resonance imaging; IH, intracranial hypertension; LightGBM, light gradient boosting machine; LSTM, long short term memory; EVD, external ventricular drainage; DC, decompressive craniectomy.
<table>
<thead>
<tr>
<th>Phases of TBI management</th>
<th>Reference</th>
<th>Task</th>
<th>Indicators / Strategies</th>
<th>Description</th>
<th>Method</th>
</tr>
</thead>
</table>
| Initial Management       | Hanko et al.\(^{41}\)  
                          Yu et al.\(^{139}\)  
                          Cui et al.\(^{22}\) | Triage | Prehospital variables | Utilize ML algorithms such as XGBoost and LR to help triage patients in the prehospital setting by classifying their severity or predicting their risk of premature death. | ML |
|                          | Ellethy et al.\(^{28}\)  
                          Brossard et al.\(^{14}\)  
                          Jain et al.\(^{39}\)  
                          Keshavamurthy et al.\(^{53}\) | Diagnosis | CT | RF, SVM, etc. and computer vision are used to rapidly auto-detect lesions in CT scans, and effectively classify patients when integrated with clinical data. | ML, DL |
| Intensive Care Unit Management | Ye et al.\(^{136}\) | Multimodality monitoring | ICP | Continuous and early prediction of ICP in using only ICP data with LSTM model. | DL |
|                          | Schweingruber et al.\(^{108}\)  
                          Jung et al.\(^{51}\)  
                          Lee et al.\(^{52}\) | Multimodality monitoring | ICP, ABP | For predicting the critical phases of ICP at various levels of accuracy, depending on the time between the prediction moment and the actual risk phase, models such as LSTM and boosting algorithms were utilized. Using multivariate analyses and LR, the study found a statistically significant correlation between parameters such as ONSD and ventricular width extracted from CT images, which could be used to predict children with elevated ICP. | ML, DL |
<p>|                          | Miyagawa et al.(^{82}) | Multimodality monitoring | ONSD |  | SA, ML |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Medical Management</th>
<th>Treatment</th>
<th>Data Sources</th>
<th>Method/Algorithm</th>
<th>Results/Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melinosky et al.</td>
<td>Multimodality</td>
<td>ECG, PPG</td>
<td></td>
<td>ML</td>
<td>To predict early neurological damage after TBI, this study analyzed ECG and PPG data in the first hour post-resuscitation using a multivariable logistic model, yielding higher predictive accuracy than clinical data alone. In patients undergoing BCT, improved survival and neurological outcomes were noted, emphasizing uncontrolled RICH as a key mortality predictor. Studies highlight the importance of BCT, especially when alternatives like DC are not feasible. Studies identified statistically benefits of therapeutic hypothermia in reducing mortality, enhancing brain function recovery, and decreasing stress responses in TBI patients. It also emphasized that these effects vary depending on factors like the depth and duration of hypothermia. The study statistically compared the effects of HTS and Mannitol on all-cause mortality, ICP changes, and serious adverse events. It was found that both treatments were similar in terms of average ICP reduction. Notably, HTS demonstrated a lower risk of ICP treatment failure. This study explored using brain EIT to evaluate mannitol treatments for IH, applying clustering methods and SVMs for analysis. The findings showed that EIT could effectively categorize patient responses to mannitol, thus promising a</td>
</tr>
<tr>
<td>Marshall et al.</td>
<td>Medical</td>
<td>BCT</td>
<td></td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td>Ryu et al.</td>
<td>Medical</td>
<td></td>
<td></td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td>McIntyre et al.</td>
<td>Medical</td>
<td>Hypothermia</td>
<td></td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td>Wang et al.</td>
<td>Medical</td>
<td></td>
<td></td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td>Burgess et al.</td>
<td>Medical</td>
<td>Hyperosmolar therapy</td>
<td></td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td>Li et al.</td>
<td>Medical</td>
<td>Hyperosmolar therapy</td>
<td></td>
<td>ML</td>
<td></td>
</tr>
</tbody>
</table>
more individualized and effective treatment approach.

<table>
<thead>
<tr>
<th>Neuro prognostication</th>
<th>Medical Management</th>
<th>DC</th>
<th>ML models were utilized to predict postoperative mortality and unfavorable outcomes in TBI patients who underwent DC. The best performance was achieved by the RF model for predicting 6-month outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thara and Thakur (1^{24})</td>
<td>Outcome Prognostication</td>
<td>Clinical scales, Vital signs, Neuroimaging</td>
<td>Utilizing XGBoost, RF, LR, ANN, and DT, research predicts outcomes like mortality and brain function decline in TBI patients. Predictions are based on in-hospital indicators and target various time frames, including 30 days and six months post-injury.</td>
</tr>
<tr>
<td>Farzaneh et al. (3^{2})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haveman et al. (4^{2})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raj et al. (6^{6})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pimentel et al. (9^{3})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abujaber et al. (2^{1})</td>
<td>Weaning Process</td>
<td>Ventilation</td>
<td>LR, ANN, SVM, RF, DT, and boosting algorithms are employed to predict successful weaning in patients requiring prolonged mechanical ventilation.</td>
</tr>
<tr>
<td>Lin et al. (6^{7})</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TBI, traumatic brain injury; ML, machine learning; DL, deep learning; SA, statistical analyses; CT, computed tomography; ICP, intracranial pressure; ABP, arterial blood pressure; ONSD, optic nerve sheath diameter; ECG, electrocardiogram; PPG, photoplethysmography; RICH, refractory intracranial hypertension; BCT, barbiturate coma therapy; DC, decompressive craniectomy; HTS, hypertonic saline; EIT, electrical impedance tomography; XGBoost, extreme gradient boosting; LR, logistic regression; RF, random forest; SVM, support vector machine; LSTM, long short term memory; ANN, artificial neural network; DT, decision tree
<table>
<thead>
<tr>
<th>Scope</th>
<th>Reference</th>
<th>Monitoring Parameter</th>
<th>Method</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhancing alarm accuracy</td>
<td>Au-Yeung et al. 5)</td>
<td>ECG, PPG, BP</td>
<td>RF</td>
<td>Score*: 83.03</td>
</tr>
<tr>
<td></td>
<td>Wang et al. 100)</td>
<td>ABP, ECG</td>
<td>SVM</td>
<td>Accuracy: 90.39% FA suppression rate: 83.99% TA suppression rate: 5.25%</td>
</tr>
<tr>
<td></td>
<td>Li and Clifford 50)</td>
<td>ABP, PPG, ECG</td>
<td>RVM</td>
<td>FA Suppression rate Asystole: 86.4% EB: 100% ET: 27.8% VT: 30.5%</td>
</tr>
<tr>
<td>Artifact removal</td>
<td>Athaya and Choi 3)</td>
<td>PPG</td>
<td>RF</td>
<td>Accuracy: 84.00 ± 2.89%</td>
</tr>
<tr>
<td></td>
<td>Tsien et al. 96)</td>
<td>ECG, BP, CO2, O2</td>
<td>DT</td>
<td>Accuracy of eliminating artifacts ECG: 98.5% BP: 98.9% CO2: 98.3% O2: 99.8%</td>
</tr>
<tr>
<td>Authors</td>
<td>Data</td>
<td>Model</td>
<td>Accuracy of eliminating artifacts</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>-----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Son et al.</td>
<td>ABP</td>
<td>DBN</td>
<td>ABP: 95.9%</td>
<td></td>
</tr>
<tr>
<td>Lee et al.</td>
<td>ABP, ICP</td>
<td>STAE-CNN</td>
<td>ABP: 97%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ICP: 94%</td>
<td></td>
</tr>
</tbody>
</table>

ECG, electrocardiogram; PPG, photoplethysmography; BP, blood pressure; ABP, arterial blood pressure; CO2, carbon dioxide concentration; O2, oxygen concentration; ICP, intracranial pressure; RF, random forest; SVM, support vector machine; RVM, relevance vector machine; DT, decision tree; DBN, deep belief network; STAE-CNN, stacked autoencoder-convolutional neural network; FA, false alarm; TA, true alarm; EB, extreme bradycardia; ET, extreme tachycardia; VT, ventricular tachycardia.

*Score*= \( \frac{(TP+TN)}{(TP+TN+FP+5\times FN)} \times 100. \)