Critical Care Medicine
XXX 2026 • Volume 54 • Number 00 • Pages XX–XX

CLINICAL INVESTIGATION

Explainable Machine Learning for Coma Outcome Prediction Based on Structural and Functional Brain MRI

Benjamine Sarton, MD, PhD[1,2]
Giulia Maria Mattia, PhD[2]
Eve Cervoni, MD[1,2]
Julie Decourt, MD[1,2]
Patrice Peran, PhD[2]
Beatrice Riu, MD[1]
Fanny Bounes, MD, PhD[3]
Edouard Naboulsi, MD[4]
Pablo Barttfeld, MD[5]
Jean Marc Olivot, MD, PhD[2]
Stein Silva, MD, PhD[1,2]
Sylvain Cussat-Blanc, PhD[6,7]

OBJECTIVES: Advanced MRI is recommended for the clinical evaluation of patients with coma. However, the implementation of these guidelines has been hindered by an inadequate identification of relevant markers among the vast amount of reported MRI-derived metrics. We developed and validated an innovative and explainable machine learning (ML) analytical pipeline to fill this critical knowledge gap.

DESIGN: Prospective cross-sectional study.

SETTING: Three Intensive Critical Care Units affiliated to the University in Toulouse (France).

PATIENTS: Patients with coma (Glasgow Coma Scale score at the hospital admission ≤ 9) from either traumatic or anoxic origin. Patient's neurologic outcome was assessed at 3 months by using the Coma Recovery Scale-Revised. Whole-brain advanced structural MRI data and functional connectivity analysis of resting-state networks known to contribute to conscious processing. A specifically designed ensemble of explainable ML methods was applied and cross-validated.

INTERVENTIONS: None.

MEASUREMENT AND MAIN RESULTS: Overall, 64 patients with coma due to either traumatic ($n = 26$) or anoxic ($n = 38$) brain injuries were studied and compared with 55 controls. The median delay between ICU admission and MRI scan was 9 days (interquartile range, 6–16 d). At 3 months, 50% patients (32/64) had an unfavorable outcome. All the models showed valuable generalization capacities: coma diagnosis (mean accuracy, 0.934%), primary brain injury discrimination (mean accuracy, 0.762 %), and neurologic outcome prediction (mean accuracy, 0.824 %).

CONCLUSIONS: A new ensemble of brain MRI-derived metrics was specifically related to coma state, its etiology, and the patient's potential for recovery at 3 months. The structural and functional integrity of mesocircuit and frontoparietal networks appeared to carry the most relevant information.

KEYWORDS: coma; explainable artificial intelligence; machine learning; mesocircuit; multimodal magnetic resonance imaging; prognosis

Prediction of consciousness recovery in the acute setting of severe brain injury remains one of the most challenging and impactful neurologic assessments (1). These determinations often guide whether life-sustaining treatment is continued or withdrawn; hence, they have major implications for patients with coma mortality, morbidity, and healthcare costs (2). It has been argued that available predictive markers of coma recovery, leveraging on clinical, electrophysiology, laboratory findings, and standard imaging

DOI: 10.1097/CCM.0000000000007068

KEY POINTS

Question: Despite recommendations for the use of advanced MRI in the evaluation of patients with coma, implementation has been constrained by the challenge of discerning clinically relevant markers within the extensive array of MRI-derived metrics.

Findings: The structural and functional integrity of mesocircuit and frontoparietal networks seemed most important for predicting a patient's recovery potential 3 months after coma onset.

Meaning: Leveraging on these results, future studies could focus on the implementation in real-world workflows of this information for clinical use and to improve our understanding of the fundamental foundations of consciousness.

(3–5), might lack accuracy because they do not reliably assess the critical integrity of brain networks supporting conscious processing (6–8). Due to these failings, prognostication is still an uncertain art for physicians treating disorders of consciousness (DoC) in the ICU and beyond.

Over the last decades, advanced neuroimaging has yielded new insights into the pathophysiological underpinnings of coma. Hence, both structural and functional neuroimaging have revealed the critical role of specific key brain regions, such as the brainstem tegmentum, thalamus, posterior cingulate cortex (PCC), and medial prefrontal cortex (mPFC), where lesions correlate with the current or future state of consciousness (9, 10). Actually, advanced structural MRI techniques such as diffusion tensor imaging (DTI) have demonstrated significant added value for either traumatic (11, 12) or anoxic (13, 14) coma neuroprognostication. On the other hand, functional MRI (fMRI) has been used to assess patients with coma resting-state brain networks (15–17) to improve the accuracy of neuroprognostication in this setting. These promising findings, coupled with the imperfect predictive value of conventional predictors and the high stakes of these evaluations, have prompted clinical guidelines to endorse advanced functional and structural MRI for comatose and DoC patient's neuroprognostication (18–21). However, there is an unresolved issue related to the overwhelming number and diversity of complex images and derived metrics.

As a paradigm shift from previous reports, which focused on the assessment of a limited number of isolated structural or functional neuroimaging data to predict coma patients' outcome, we have specifically designed and validated a new comprehensive machine learning (ML) analytical pipeline to generate "actionable intelligence" from such high-dimensional neuroimaging datasets. We hypothesized that this ML-based strategy would allow us to identify novel multimodal MRI-derived metrics that are specifically associated with the coma state, its etiology, and the coma patient's potential for recovery.

MATERIALS AND METHODS

Study Design and Participants

Cross-sectional study of traumatic and anoxic patients with coma and compared controls, from two previous prospectively collected datasets: Intrinsic Brain Activity in COMA (ACI-COMA) (NCT NCT03482115) (13, 15–17) and COMA-3D (NCT 01620957) (22, 23). Both studies were approved by the Ethics Committee of the University Teaching Hospital of Toulouse, France (ACI-COMA, January 08, 2015; Functional, Structural and Neuroimmune Connectomes in COMA [COMA-3D], February 22, 2018). These prospective studies aimed to study the predictive value of functional and structural MRI in coma and were undertaken in three ICUs affiliated with the University Hospital of Toulouse (Toulouse, France) between January 2015 and December 2017 (ACI-COMA) and February 2018 and February 2022 (COMA-3D). MRI and clinical assessment (Glasgow Coma Scale [GCS] and Full Outline of UnResponsiveness [FOUR]) (1) were performed at baseline. Patients were followed up at 3 months after primary brain injury, including scoring with the revised version of the Coma Recovery Scale (CRS-R) (2). In all instances, informed and written consent to participate in the study was obtained from the subjects themselves in the case of healthy subjects and from legal surrogates of the patients. Written consent to use previously collected data was obtained from patients who recovered consciousness across the study follow-up duration. All research procedures were followed in accordance with the ethical standards of the responsible committee on human experimentation of

the University Teaching Hospital of Toulouse and with the Helsinki Declaration of 1975.

Population

Patients with either traumatic or anoxic brain injury were included in the study after being diagnosed as being in a coma (GCS score at admission to the hospital ≤ 9 with motor responses < 6). Patients needed to maintain a GCS score less than or equal to 9 with motorsubscore less than 6 without sedation at the time of brain imaging acquisition. Patients not meeting this criterion at assessment were excluded (**Fig. 1**; and **eTable 1**, https://links.lww.com/CCM/H900). Additional exclusion criteria were significant neurologic or psychiatric illness before coma. All patients were managed according to standard of care recommendations by physicians blinded to neuroimaging data. All patient assessments were conducted at least 2 days (4 ± 2 d) after complete withdrawal of all sedative drug therapy and were performed under normothermic conditions. On the day of brain imaging, urine benzodiazepine and barbiturate screening tests were used in patients to rule out residual sedation in case of prolonged utilization of these drugs. Additional exclusion criteria were significant neurologic, psychiatric illness before coma, poor quality of structural MRI images, and motion in fMRI acquisition (> 3 mm in translation, > 3° in rotation).

Over the same recruitment period, 55 controls, matched by age, were recruited and included if they had normal neurologic examination results and no history of neurologic or psychiatric disorder. Control participants were scanned in the same MRI facility as patients and underwent the same MRI acquisition protocol.

Clinical Assessment

In patients, standardized clinical examination (GCS and FOUR) was performed by investigators (B.S., S.S.), blinded to neuroimaging data on the day of

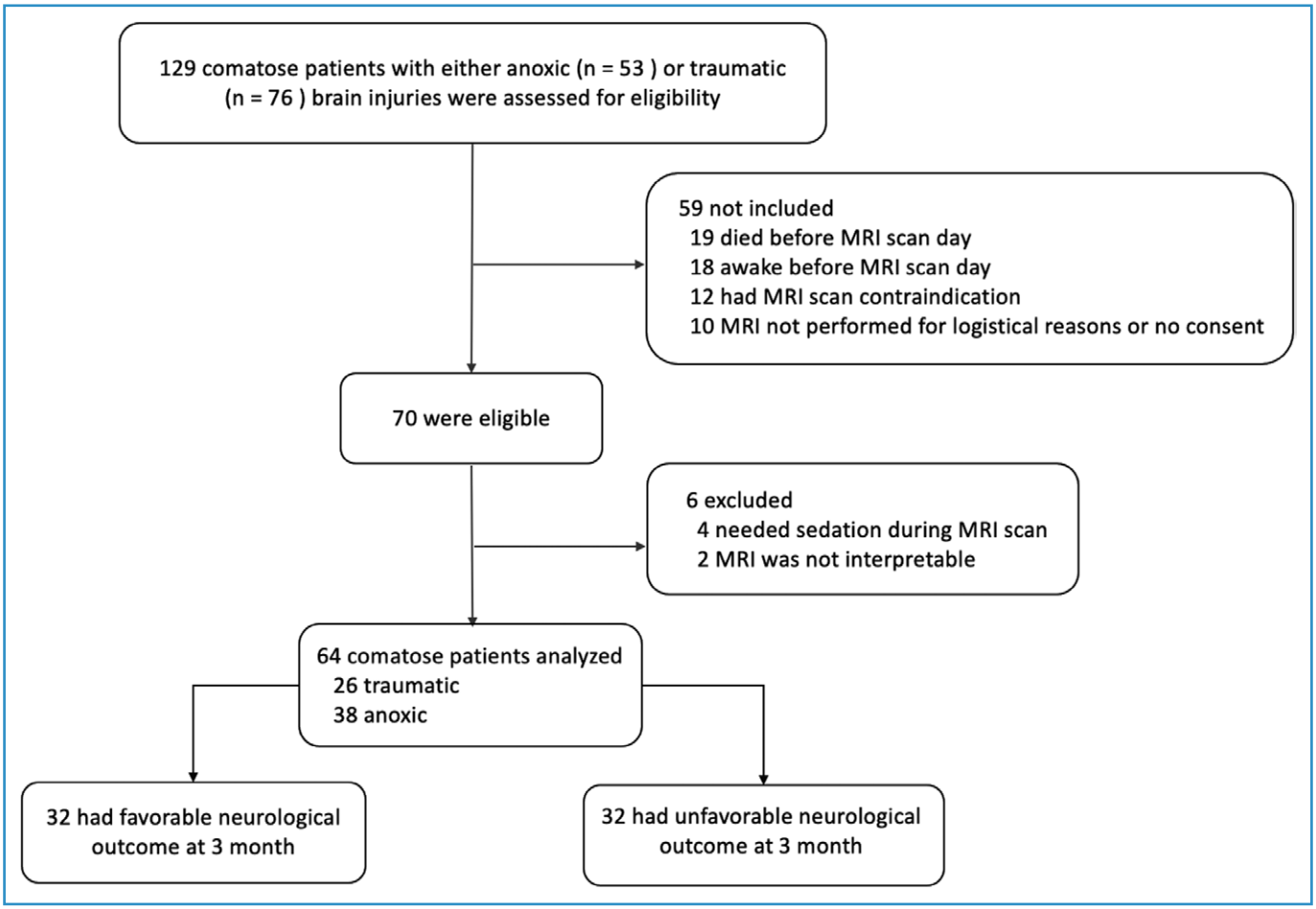

Figure 1. Study flow chart.

the patient's admission to the hospital and the day of MRI scanning. The patient's neurologic outcome was assessed 3 months after the primary brain injury by using the CRS-R. The CRS-R is a 23-point scale measuring arousal level, motor function, visual-perception, auditory, language, and communication abilities (2). This scale enables the distinction between patients in an unresponsive wakefulness syndrome (UWS, also known as vegetative state) and patients in a minimally conscious state (MCS) or patients who emerged from MCS (EMCS) and recovered consciousness as reflected by functional communication or functional use of objects. According to this scale, the patient's 90-day outcome was binarized as either favorable (MCS or EMCS) or unfavorable (UWS or deceased).

MRI Acquisitions and Preprocessing

In all participants, high-resolution anatomical images, using 3D T1-weighted sequences, were acquired on the same 3T magnetic resonance scanner (Philips Achieva; Dstream, Best, The Netherlands). Monitoring of vital measures was performed by a senior intensivist throughout the experiment.

Structural MRI comprised 3D T1-weighted, fluid-attenuated inversion recovery (FLAIR), and DTI sequences were acquired. Brain extraction and normalization were performed on T1-weighted and FLAIR images with Statistical Parametric Mapping (Version SPM 12; Wellcome Trust Center for Neuroscience, London, United Kingdom; https://www.fil.ion.ucl.ac.uk/spm/). Before the computation of DTI models, images were skull-stripped and corrected for eddy current using the standard FMRIB Software Library pipeline. Fractional anisotropy (FA) and mean diffusivity (MD) maps were computed from DTI by fitting a standard tensor model, implemented in Python (Version 3.8.3 and 3.10.8; Python Software Foundation, Wilmington, DE) via the Dipy library (Version 1.3.0; DIPY development team, hosted via GitHub, San Francisco, CA). Analyses were conducted using Python (Python Software Foundation), DIPY (Version 1.3.0; DIPY development team, hosted via GitHub), scikit-learn (Version 1.3.2; scikit-learn developers, hosted via GitHub), and SHAP (Version 0.43.0; SHAP development team, hosted via GitHub). Code was version-controlled using GitHub (GitHub, Inc.).

Based on previous reports (2, 15–17, 22, 24), functional connectivity analysis was performed on resting-state fMRI data using a region of interest (ROI) vs. whole-brain approach, with the PCC as the considered ROI. fMRI data preprocessing was performed with Statistical Parametric Mapping (Version SPM 12; https://www.fil.ion.ucl.ac.uk/spm/), comprising realignment, slice-time correction, and co-registration to each subject's T1-weighted native space. All images were registered in the Montreal Neurological Institute template.

Machine Learning Processing

A full overview of ML methods is provided in **Figure 2**.

Dataset Constitution. Residual head motion was corrected in the preprocessing procedure performed by SPM 12 realign and unwarp procedure, co-registering, and resampling all scans to the first scan of the session using b-spline interpolation (integrated within the Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks (CONN) toolbox https://www.nitrc.org/projects/conn). Each image acquisition modality was first aggregated per area of the brain using: 1) Willard's functional atlas for PCC, mPFC, and precuneus and 2) Hammer's anatomical atlas (reference) for deep gray nuclei, that is, pallidum, putamen, and thalamus. Willard's atlas has been developed by the Functional Imaging in Neuropsychiatric Disorders Lab (Stanford University, Stanford, CA), describing 14 functional brain networks, including the DMN, with 499 functional ROIs (https://pyhrf.github.io/manual/parcellation_mask.html). Six cortical and subcortical structures, known to critically contribute to conscious processing (1, 8, 9), were considered: pallidum, putamen, thalamus, mPFC, PCC, and precuneus. For each patient in the study, the value of the voxels in a given structure was summed, providing one measure of the structure. This procedure resulted in the construction of a dataset that contains 119 subjects (patients with coma = 64; controls = 55) with 30 features per patient (five acquisition techniques times six areas). For each subject, we considered three possible classification targets: clinical status (coma vs. controls), patient's expected outcome (favorable vs. unfavorable), and the etiology of the primary brain injury (anoxic vs. traumatic). For the last two targets, only patients with coma were considered. A predictive analysis was then applied to each target separately, but the same procedure was used. The datasets were split into a train set and a test set with an 80/20 ratio. Python, Version 3.10.8, was used to preprocess the data.

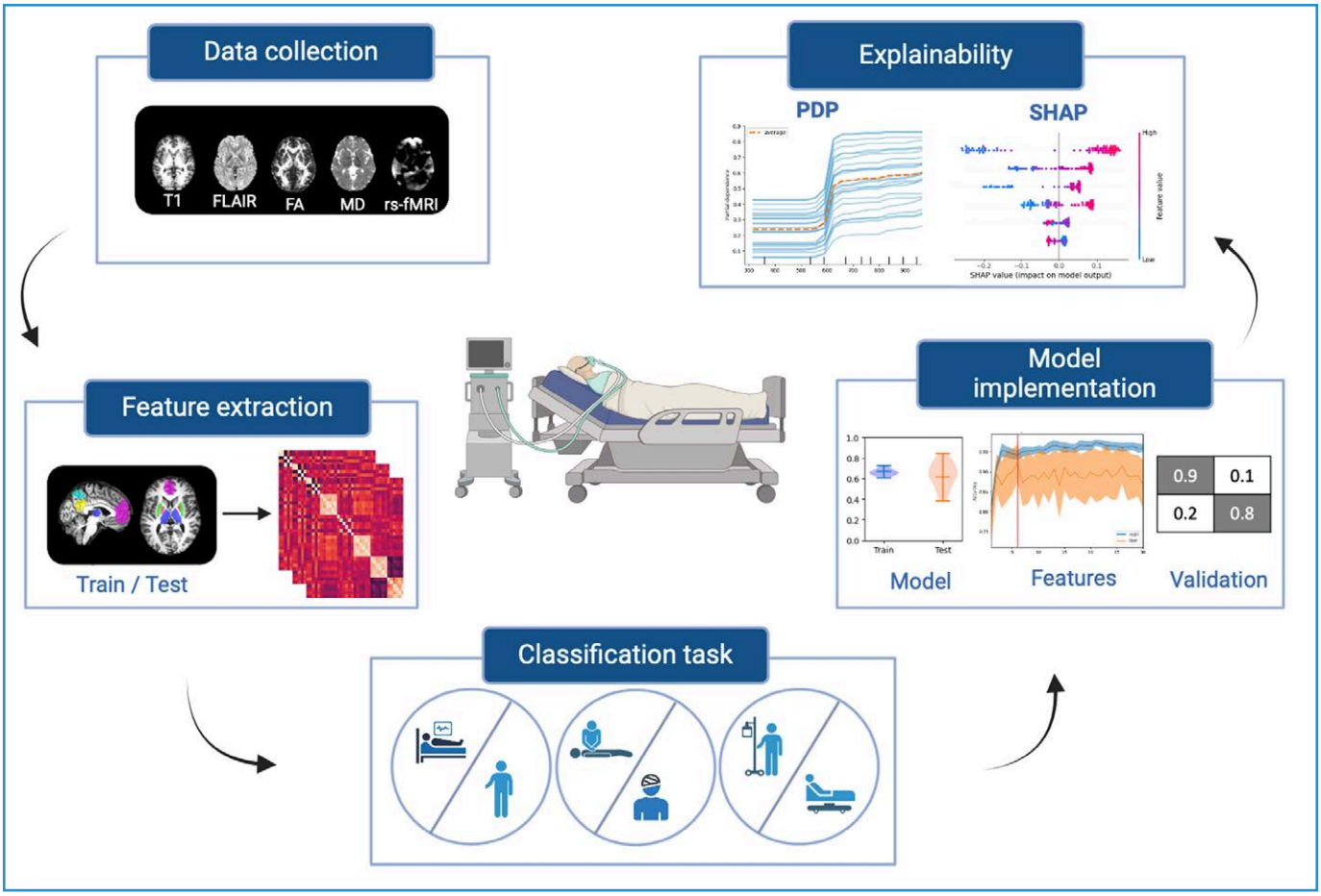

Figure 2. Methods overview. In all participants, structural and functional brain MRI was collected at the acute phase of coma. The whole dataset was split into a train and test sets (80/20 ratio) to build and validate independent machine learning models with the following classification targets: clinical status (coma vs. controls), primary brain injury mechanism (anoxic vs. traumatic), and patient's expected outcome at 3 mo from coma onset (favorable vs. unfavorable). Ultimately, to understand features importance in the model, model-agnostic visualization tools were applied. FA = fractional anisotropy, FLAIR = fluid-attenuated inversion recovery, MD = median diffusivity, PDP = partial dependence plot, rs-fMRI= resting-state functional MRI, SHAP = SHapley Additive exPlanations, T1 = T1-weighed sequence.

Regarding structural MRI data, T1 and FLAIR voxel intensities were directly used after brain extraction and spatial normalization to compute the sum of these intensity values for each ROI. To further assess the structural integrity of a region, DTI models were created to fit at each voxel, generating FA and MD maps, assessing fiber orientation and microstructural integrity, respectively. Each patient was described by a structural measure per region obtained by summing up voxel intensities according to the MRI modality. Concerning the fMRI analysis counterpart, rs-fMRI data were preprocessed using Statistical Parametric Mapping (Version SPM 12; https://www.fil.ion.ucl.ac.uk/spm/). The fMRI images were realigned, slice-time corrected, coregistered to each subject's T1-weighted image, and normalized to standard stereotaxic anatomical Montreal Neurological Institute space (reference). Seed-based connectivity maps, representing Fisher-transformed correlation coefficients for the regions of interest (ROI), were computed from the spatial maps of Pearson correlation coefficient, using the CONN toolbox (https://web.conn-toolbox.org/fmri-methods/connectivity-measures/seed-based). The ROI chosen as the seed was the PCC, whose activations were statistically compared with all other brain regions (whole-brain approach thereof).

Hyperparameter Optimization. The next step of the ML pipeline consisted in optimizing hyperparameters of five ML algorithms: K-Neighbors, Decision tree, Random Forest, Gaussian Naive Bayes, and Gradient Boosting. Grid search was employed with parameter ranges provided in the **Supplemental Digital Content** (**eFigs. 1–18**, https://links.lww.com/CCM/H900). A five-fold cross-validation was used during hyperparameter optimization to assess the generalization

capacity of the models. SKLearn, Version 1.3.2, was used to optimize hyperparameters and train models.

Model Selection. Once the hyperparameter was optimized for each ML algorithm, the best algorithm was chosen by evaluating its stability among reruns: each of the five ML algorithms with the corresponding optimized hyperparameter set was trained 30 times on various train/test splits (eFigs. 1–18, https://links.lww.com/CCM/H900). The best algorithms were chosen with higher accuracy, with the lower train/test difference, independently for each target.

Feature Selection. With the best ML algorithm and its optimal hyperparameter set, feature extraction was applied to identify key features from the dataset. Permutation importance from the SKLearn package was applied. The selected algorithm was then fitted 100 times with an increasing number of features, adding them one at a time in decreasing importance (eFigs. 1–18, https://links.lww.com/CCM/H900). For each target, the final number of features was manually chosen, choosing the highest accuracy with the lowest train/test difference. The corresponding most important features were kept for the next steps.

Model Explanation. With the selected model and feature, confusion matrices were generated and explanation using SHapley Additive exPlanations (SHAP) (25) and partial dependence plots (PDPs) (26). Summary plots were generated using SHAP library, Version 0.43.0 and Individual Conditional Expectation plots taken from the inspection package of SKLearn.

RESULTS

Neuroimaging Signatures of Coma

In the first classification task, we aimed to distinguish patients with coma from controls using the proposed automatic ML pipeline for structural and functional brain MRI data. In **Figure 3*A***, we present the classifier with the best performance, that is, random forest, after feature selection, showing a mean accuracy of 0.934% on the training set and 0.890% on the test set, showing the good generalization capacity of the generated

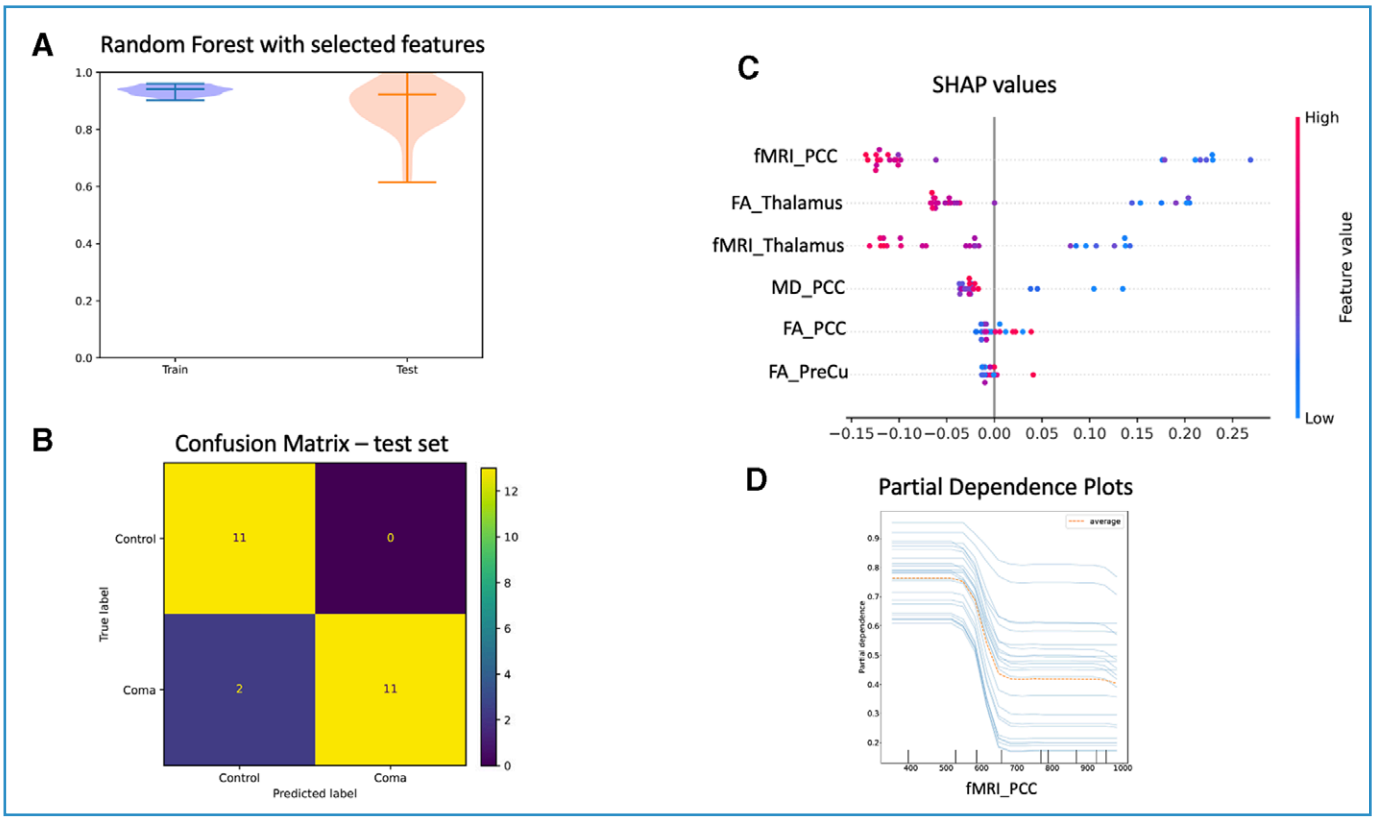

Figure 3. Neuroimaging signatures of coma. Model classification of clinical status (coma vs. controls) at the MRI scan time. **A**, Best machine learning model (Random Forest) with selected features. **B**, Confusion matrix from the test dataset. Model explainability investigation using SHapley Additive exPlanations (SHAP, **C**) and partial dependence plots (**D**). FA = fractional anisotropy, fMRI = functional MRI, MD = median diffusivity, PCC = posterior cingulate cortex, PreCu = precuneus.

models. The confusion matrix in **Figure 3*B*** of the test set (eFig. 4, https://links.lww.com/CCM/H900) reports that almost all samples were correctly classified, except for one control mistaken for patients with coma. With this model validated in terms of diagnosis capacity, we analyzed what were the most impactful features. First, of the 30 original features, only six were sufficient to reach the maximal accuracy score: fMRI-PCC, fMRI-thalamus, FA-thalamus, MD-PCC, FA-PCC, and FA-precuneus. Regarding the model's interpretability, the SHAP analysis allowed for the comparison between the selected features (**Fig. 3*C***). We found that the resting-state fMRI index from two ROIs, that is, PCC and thalamus, respectively, ranked as first and second in order of importance. The third place was occupied by the FA index of the thalamus, followed by MD in PCC in fourth position, and the FA-PCC and precuneus in the fifth and sixth positions. Interestingly, the lower the value of the fMRI-PCC, the more the prediction shifted toward the coma state. A similar trend was observed for the fMRI-thalamus, FA-thalamus, and MD-PCC, but was opposite for structural features, such as FA-PCC and FA-precuneus. The value of fMRI-PCC and the FA-thalamus to disentangle coma from conscious states was confirmed and expanded by PDP analysis (**Fig. 3*D***).

Multimodal Features Related to Primary Brain Injury

For the second classification task, we analyzed coma etiology by discriminating the causes of coma in the considered set of patients, namely, either anoxic or traumatic brain injuries. In this case, the best selected model was Gaussian Naive Bayes, achieving 0.762 as average accuracy on the training set and 0.733 on the test set, once again showing good generalization capacity of the generated models (**Fig. 4*A***). Confusion matrices also showed the classification capacity of the selected model; however, it mainly misclassifies traumatic brain injuries (**Fig. 4*B***). SHAP analysis was used to identify the brain MRI features that relate most with the primary brain injury mechanism responsible for coma. fMRI-PCC and FA-pallidum were isolated as the most relevant markers allowing distinction between traumatic and anoxic brain injuries (**Fig. 4*C***). It is worth noting that PDP analysis substantiates this result and permits to isolate a linear relationship between

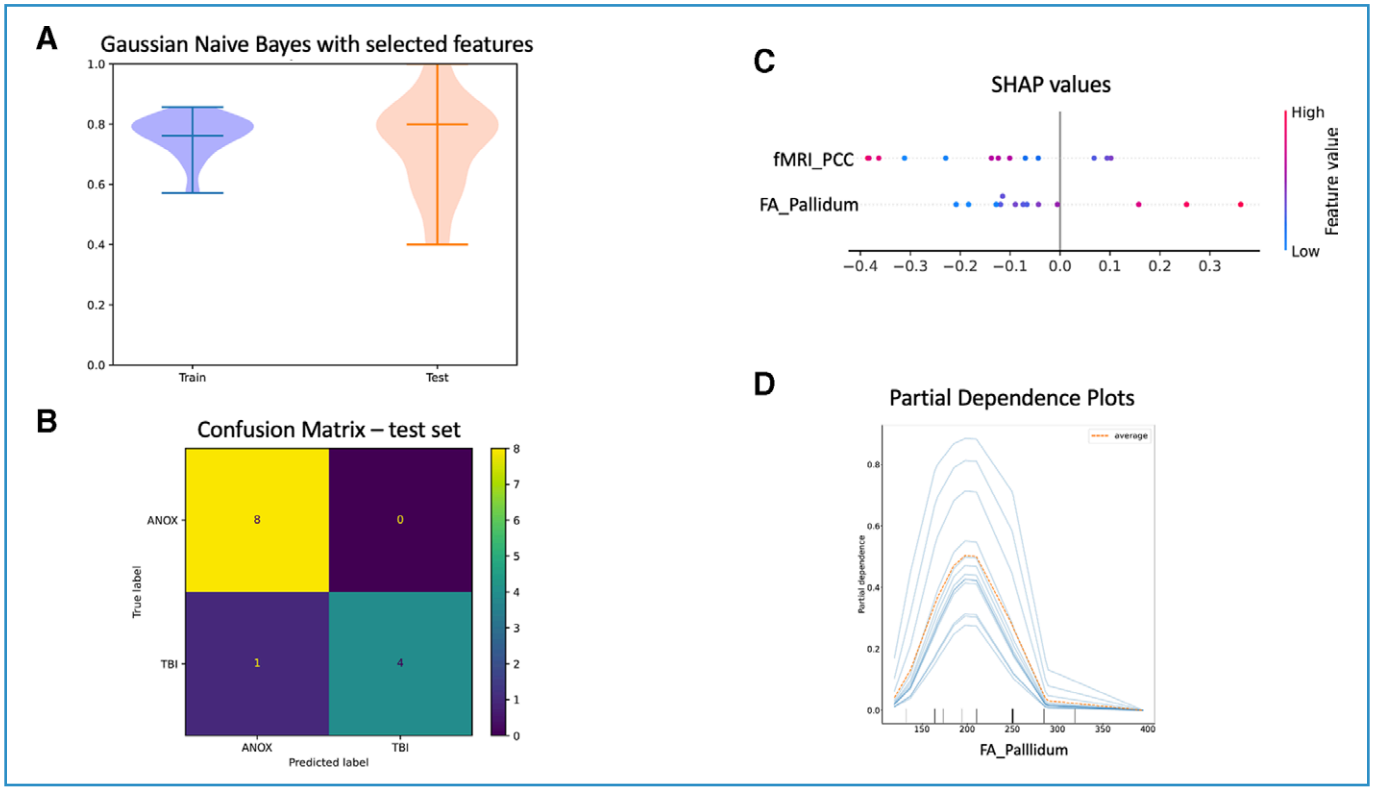

Figure 4. Multimodal features related to primary brain injury. Model classification of coma etiology (traumatic vs. anoxic [ANOX]). **A**, Best machine learning model (Gaussian Naive Bayes) with selected features. **B**, Confusion matrix from the test dataset. Model explainability investigation using SHapley Additive exPlanations (SHAP, **C**) and partial dependence plots (**D**). FA = fractional anisotropy, fMRI = functional MRI, PCC = parietal cingulate cortex, TBI = traumatic brain injury.

fMRI-PCC and primary brain injury classification, with lower fMRI-PCC values related to brain trauma conditions (**Fig. 4*D***). Additional analyses concerning features related to etiological heterogeneity and data preprocessing are available in the Supplemental Digital Content (**eFigs. 28–30**, https://links.lww.com/CCM/H900).

MRI-Derived Markers of Neurologic Outcome

The last classification task comprised the distinction between patients having either favorable or unfavorable outcomes at 3 months from coma onset. Among the five features leading to the best results, structural indexes (T1 and FLAIR) seemed to carry the most relevant predictive information. Gaussian Naive Bayes was the best classifier, although test performance presented a great variability and was significantly lower than the test set (accuracy of 0.694) in comparison to the training set (accuracy of 0.824) (**Fig. 5, *A*** and ***B***). Taking that into account in the following analysis, SHAP analysis (**Fig. 5*C***), corroborated with PDPs (**Fig. 5*D***), shows that higher values of T1 in PCC and precuneus are of better prognostic in terms of outcome. Further analysis examining neurological outcome classification are presented in Supplementary Material (**eFigs. 19–24**, https://links.lww.com/CCM/H900).

Data Availability

Anonymized data may be shared on request to the corresponding author from a qualified investigator for noncommercial use, subject to restrictions according to participant consent and data protection legislation.

Code Availability

Source code is available for noncommercial use on GitHub.

DISCUSSION

Converging evidence suggests that patients with DoC and preserved structural integrity within the anterior

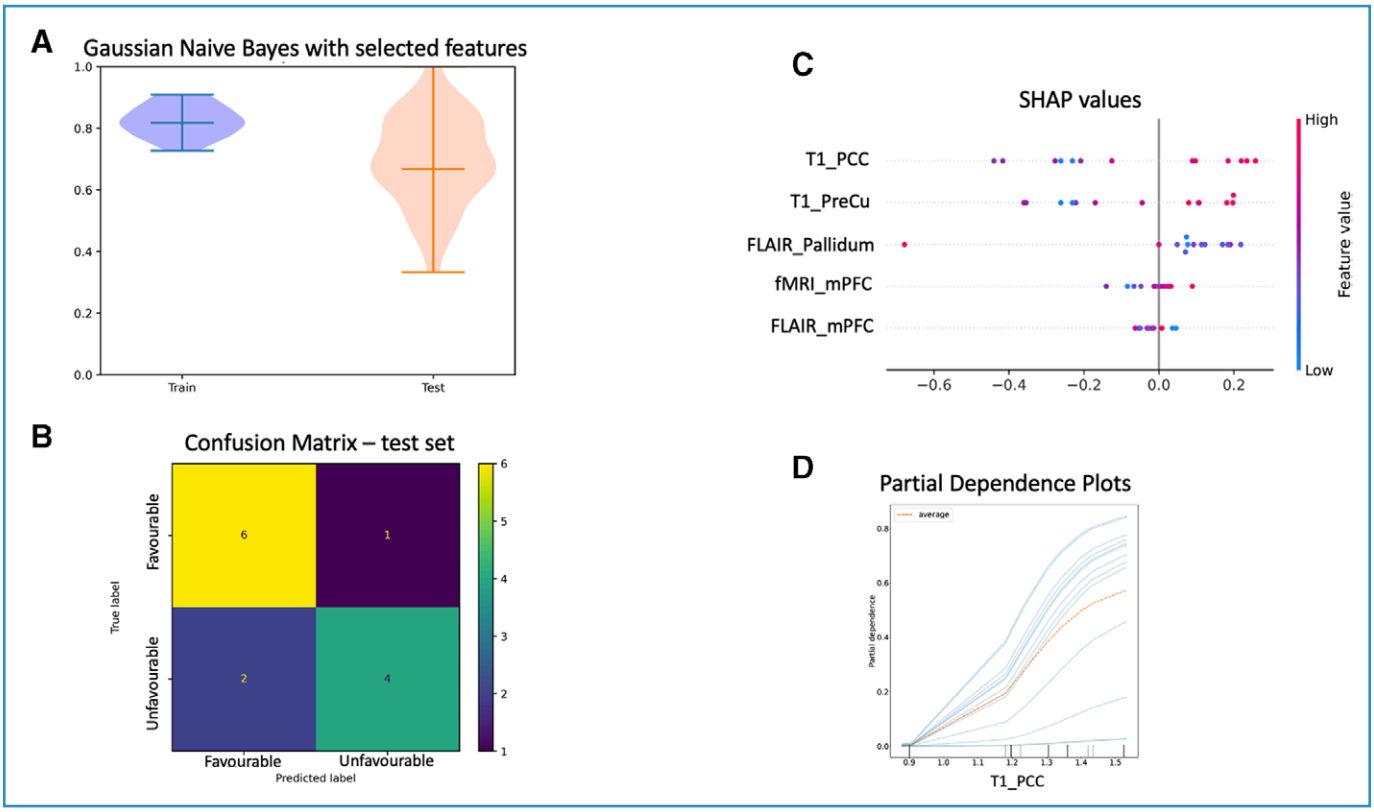

Figure 5. MRI-derived markers of neurologic outcome. Model classification of neurologic outcome at 3 mo (favorable vs. unfavorable). **A**, Best machine learning model (Gaussian Naive Bayes) with selected features. **B**, Confusion matrix from the test dataset. Model explainability investigation using SHapley Additive exPlanations (SHAP, **C**) and partial dependence plots (**D**). FLAIR = fluid-attenuated inversion recovery, fMRI = functional MRI, mPFC = median prefrontal cortex, PCC = parietal cingulate cortex, PreCu = precuneus, T1 = T1-weighed sequence.

forebrain mesocircuit display more favorable neurologic evolution (1). Regarding the functional counterpart of reported MRI-derived predictors of DoC recovery, default mode network connectivity, encompassing both within and internetwork correlations between key functional hubs (thalamus, the PCC, and the mPFC), has been intensely investigated in this context, showing that functional disruption increases with consciousness impairment, ranging from a MCS, UWS, to coma (1). Here, we have developed and validated an innovative, hypothesis-driven, and broadly explainable analytical pipeline, which was capable of accurately identifying MRI-derived signatures of coma, its etiology, and potential for neurologic recovery. It is worth noting that selected clinical features (age, sex, time from coma onset and MRI scanning, GCS at hospital admission) weakly contribute to classification models accuracy (**eFigs. 25–27**, https://links.lww.com/CCM/H900) and suggest that including these MRI biomarkers improves classifications beyond what clinical variables alone can achieve.

The generated ML model showed good accuracy and generalization capacity to discriminate coma from conscious state based on early brain MRI data. Interestingly, by following a without a priori hypothesis analytical approach, we observed that functional baseline dysconnection between the whole-brain and either the thalamus, PCC, or mPFC was associated with coma. These results are in line with current theoretical and neurobiological frameworks of consciousness, which posit that these cortical and subcortical brain structures are key information hubs for conscious processing (1, 2, 9). Hence, grounded on the current theoretical framework for conscious processing, we deliberately focus on the study of MRI-derived metrics of neuronal integrity and baseline function of mesocircuit and frontoparietal networks' key components. Further studies should explore the added value of lesions identified outside these chosen networks, such as the reticular ascending activation system, the dorsal attention, and salience networks (15–17). Regarding the structural imaging counterpart, FA anomalies—a measure of microstructural integrity in diffusion MRI—within the thalamus and the nigrostriatal system were also associated with coma state.

Finally, the structural integrity of PCC, the thalamus, and the nigrostriatal structures, assessed by widely available and cost-effective MRI metrics (FA, FLAIR), conveyed valuable information in terms of neurologic outcome prediction. Our findings support current theoretical hypotheses about the implication of such neocortical and subcortical brain structures in consciousness abolition and neurologic recovery from severe brain injury.

The proposed structural MRI and fMRI data analysis pipeline was developed and evaluated within a specific clinical workflow, involving patients with acute severe brain injuries (5–15 d post-injury) who were treated in a tertiary ICU. However, our results must be interpreted with caution, and a number of limitations should be borne in mind. First, we cannot exclude the possibility of unmeasured confounding factors in this observational study. For example, even though we controlled for ruling out residual sedation the day of the MRI scan, other essential medications were administered at the time of neuroimaging data collection, and several of them could affect the resting-state brain activity. Second, despite the promise held by advanced neuroimaging in this setting, performing multisequence MRI in this setting is logistically challenging and can be hindered by cost and accessibility issues that need specific workflow improvements. Third, despite being one of the largest reported to date prospective multimodal MRI cohorts in coma patients, the limited sample size of our study hinders the generalizability of our findings. Fourth, imaging on the same scanner raises the possibility of site-specific bias and hinders the generalizability of the reported models in external populations scanned using different MRI platforms. Fifth, DTI quantifies the 3D movement of water molecules by assuming Gaussian diffusion within a voxel. Main metrics derived from DTI, such as FA and MD, provide valuable insights into brain microstructure, but can be biased in the case of a voxel containing crossing fibers of a mixture of white matter and freely extracellular water molecules. Alternative DTI methods, such as free-water elimination modeling and multishell angular resolution diffusion imaging, have been proposed to accurately address these issues.

Our study addresses the crucial need for understanding how artificial intelligence-/ML-based approaches are taking their decisions, which is of central interest, in particular, in critical applications such as healthcare. While our approach is not yet fully

interpretable, we argue that proposing SHAP and PDPs on human-extracted features from images is a step toward this direction. This allowed us to shed light on the most relevant features for each classification task, thus providing additional elements for increasing trust in these methods. Nevertheless, these explanations are descriptive rather than causal and should not be interpreted as evidence of underlying mechanistic relationships.

In the current study, we developed a hypothesis-driven ML-analytical pipeline that was capable of accurately and comprehensibly identifying multimodal MRI-derived metrics that were specifically associated with coma per se, its etiology, and the patient's potential for recovery. The structural and functional integrity of mesocircuit and frontoparietal brain hubs appeared to carry the most relevant information in terms of the model's classification. However, these results should be considered preliminary prognostic signatures that still require external validation.

1 *Critical Care Unit, University Teaching Hospital of Purpan, Toulouse, France.*

2 *Toulouse NeuroImaging Center, Toulouse University, Toulouse, France.*

3 *Critical Care Unit, University Teaching Hospital of Rangueil, Toulouse, France.*

4 *Neurocritical Care Unit, University Teaching Hospital of Purpan, Toulouse, France.*

5 *Cognitive Science Group, Instituto de Investigaciones Psicológicas, Facultad de Psicología Universidad Nacional de Córdoba–CONICET, Córdoba, Argentina.*

6 *Artificial and Natural Intelligence Toulouse Institute ANITI, Toulouse, France.*

7 *Institute of Research in Informatics (IRIT) of Toulouse, Toulouse, France.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (https://journals.lww.com/ccmjournal).

This work was supported by the "Association des Traumatisés du Crâne et de la Face," "Fondation de l'Avenir," and grant funding from the University Hospital of Toulouse.

Dr. Peran disclosed government work. Dr. Olivot received funding from Roche, Boehringer Ingelheim, and British Medical Society. Dr. Cussat-Blanc's institution received funding from French National Research Agency (ANR) Cube; he received support for article research from ANR. The remaining authors have disclosed that they do not have any potential conflicts of interest.

Drs. Silva and Cussat-Blanc contributed equally to this article.

For information regarding this article, E-mail: silvastein@me.com; silva.s@chu-toulouse.fr

REFERENCES

1. Edlow BL, Claassen J, Schiff ND, et al: Recovery from disorders of consciousness: Mechanisms, prognosis and emerging therapies. *Nat Rev Neurol* 2020; 17:135–156
2. Giacino JT, Fins JJ, Laureys S, et al: Disorders of consciousness after acquired brain injury: The state of the science. *Nat Rev Neurol* 2014; 10:99–114
3. Maas AIR, Menon DK, Adelson PD, et al; InTBIR Participants and Investigators: Traumatic brain injury: Integrated approaches to improve prevention, clinical care, and research. *Lancet Neurol* 2017; 16:987–1048
4. Nolan JP, Sandroni C, Bottiger BW, et al: European Resuscitation Council and European Society of Intensive Care Medicine Guidelines 2021: Post-resuscitation care. *Resuscitation* 2021; 161:220–269
5. Sandroni C, Grippo A, Nolan JP: ERC-ESICM guidelines for prognostication after cardiac arrest: Time for an update. *Intensive Care Med* 2020; 46:1901–1903
6. Menon DK, Maas AI: Traumatic brain injury in 2014. Progress, failures and new approaches for TBI research. *Nat Rev Neurol* 2015; 11:71–72
7. Sharp DJ, Scott G, Leech R: Network dysfunction after traumatic brain injury. *Nat Rev Neurol* 2014; 10:156–166
8. Thibaut A, Schiff N, Giacino J, et al: Therapeutic interventions in patients with prolonged disorders of consciousness. *Lancet Neurol* 2019; 18:600–614
9. Dehaene S, Changeux JP: Experimental and theoretical approaches to conscious processing. *Neuron* 2011; 70:200–227
10. Schiff ND: Recovery of consciousness after brain injury: A mesocircuit hypothesis. *Trends Neurosci* 2010; 33:1–9
11. Galanaud D, Perlbarg V, Gupta R, et al; Neuro Imaging for Coma Emergence and Recovery Consortium: Assessment of white matter injury and outcome in severe brain trauma: A prospective multicenter cohort. *Anesthesiology* 2012; 117:1300–1310
12. Sidaros A, Engberg AW, Sidaros K, et al: Diffusion tensor imaging during recovery from severe traumatic brain injury and relation to clinical outcome: A longitudinal study. *Brain* 2008; 131:559–572
13. Silva S, Peran P, Kerhuel L, et al: Brain gray matter MRI morphometry for neuroprognostication after cardiac arrest. *Crit Care Med* 2017; 45:e763–e771
14. Velly L, Perlbarg V, Boulier T, et al; MRI-COMA Investigators: Use of brain diffusion tensor imaging for the prediction of long-term neurological outcomes in patients after cardiac arrest: A multicentre, international, prospective, observational, cohort study. *Lancet Neurol* 2018; 17:317–326
15. Malagurski B, Peran P, Sarton B, et al: Neural signature of coma revealed by posteromedial cortex connection density analysis. *Neuroimage Clin* 2017; 15:315–324
16. Malagurski B, Peran P, Sarton B, et al: Topological disintegration of resting state functional connectomes in coma. *Neuroimage* 2019; 195:354–361
17. Silva S, de Pasquale F, Vuillaume C, et al: Disruption of posteromedial large-scale neural communication predicts recovery from coma. *Neurology* 2015; 85:2036–2044

18. Farisco M, Salles A: American and European Guidelines on Disorders of Consciousness: Ethical challenges of implementation. *J Head Trauma Rehabil* 2022; 37:258–262
19. Giacino JT, Katz DI, Schiff ND, et al: Practice guideline update recommendations summary: Disorders of consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. *Neurology* 2018; 91:450–460
20. Rajajee V, Muehlschlegel S, Wartenberg KE, et al: Guidelines for neuroprognostication in comatose adult survivors of cardiac arrest. *Neurocrit Care* 2023; 38:533–563
21. Scolding N, Owen AM, Keown J: Prolonged disorders of consciousness: A critical evaluation of the new UK guidelines. *Brain* 2021; 144:1655–1660
22. Peran P, Malagurski B, Nemmi F, et al: Functional and structural integrity of frontoparietal connectivity in traumatic and anoxic coma. *Crit Care Med* 2020; 48:e639–e647
23. Sarton B, Tauber C, Fridman E, et al: Neuroimmune activation is associated with neurological outcome in anoxic and traumatic coma. *Brain* 2024; 147:1321–1330
24. Mattia GM, Sarton B, Villain E, et al: Multimodal MRI-based whole-brain assessment in patients in anoxoischemic coma by using 3D convolutional neural networks. *Neurocrit Care* 2022; 37:303–312
25. Wang Y, Zhang L, Jiang Y, et al: Multiparametric magnetic resonance imaging (MRI)-based radiomics model explained by the Shapley Additive exPlanations (SHAP) method for predicting complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: A multicenter retrospective study. *Quant Imag Med Surg* 2024; 14:4617–4634
26. Cox LA Jr: Objective causal predictions from observational data. *Crit Rev Toxicol* 2024; 54:895–924