NeuroImage: Clinical 30 (2021) 102651

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Discriminating cognitive motor dissociation from disorders of consciousness using structural MRI

Polona Pozeg [a,1], Jane Jöhr [b,1], Alessandro Pincherle [b,c], Guillaume Marie [a], Philippe Ryvlin [b], Reto Meuli [a], Patric Hagmann [a], Karin Diserens [b,*,2], Vincent Dunet [a,*,2]

[a] Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[b] Neurology and Acute Neurorehabilitation Unit, Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[c] Neurology Unit, Department of Medicine, Hopitaux Robert Schuman, Luxembourg, Luxembourg

ARTICLE INFO

Keywords:
Disorders of consciousness
Structural MRI
Support vector machine
Cognitive motor dissociation
Brain injury

ABSTRACT

An accurate evaluation and detection of awareness after a severe brain injury is crucial to a patient's diagnosis, therapy, and end-of-life decisions. Misdiagnosis is frequent as behavior-based assessments often overlook subtle signs of consciousness. This study aimed to identify brain MRI characteristics of patients with residual consciousness after a severe brain injury and to develop a simple MRI-based scoring system according to the findings.

We retrieved data from 128 patients and split them into a development or validation set. Structural brain MRIs were qualitatively assessed for lesions in 18 brain regions. We used logistic regression and support vector machine algorithms to first identify the most relevant brain regions predicting a patient's outcome in the development set. We next built a diagnostic MRI-based score and estimated its optimal diagnostic cut-off point. The classifiers were then tested on the validation set and their performance compared using the receiver operating characteristic curve.

Relevant brain regions predicting negative outcome highly overlapped between both classifiers and included the left mesencephalon, right basal ganglia, right thalamus, right parietal cortex, and left frontal cortex. The support vector machine classifier showed higher accuracy (0.93, 95% CI: 0.81–0.96) and specificity (0.97, 95% CI: 0.85–1) than logistic regression (accuracy: 0.87, 95% CI: 0.73 – 0.95; specificity: 0.90, 95% CI: 0.75–0.97), but equal sensitivity (0.67, 95% CI: 0.24–0.94 and 0.22–0.96, respectively) for distinguishing patients with and without residual consciousness.

The novel MRI-based score assessing brain lesions in patients with disorders of consciousness accurately detects patients with residual consciousness. It could complement valuably behavioral evaluation as it is time-efficient and requires only conventional MRI.

1. Introduction

Severe brain injuries often result in disorders of consciousness (DOC), in which the main components, wakefulness and awareness, are altered (Posner et al., 2007). The majority of DOC patients first fall into coma after brain lesion (Dikmen et al., 2003) and may then gradually recover consciousness by transitioning through ascending levels of arousal and awareness. The patient is clinically diagnosed with unresponsive wakefulness syndrome (UWS) when the sleep-wake cycles and autonomic and motor reflexes are present but no behavioral signs of awareness of self or environment are evidenced (Monti et al., 2010). The patient may transition into a minimally conscious state (MCS), whereby minimal and inconsistent, but definite behavioral signs of awareness are observed, e.g., following simple commands, or noxious stimuli

Abbreviations: DOC, Disorders of consciousness; UWS, Unresponsive wakefulness syndrome; MCS, Minimally conscious state; CRS-R, Coma Recovery Scale Revised; MRI, Magnetic resonance imaging; CMD, Cognitive motor dissociation; MBT-R, Motor Behavior Tool Revised; LR, Logistic regression; SVM, Support vector machine; ROC, Receiver operating characteristic.

* Corresponding authors at: Department of Radiology, Lausanne University Hospital, Rue du Bugnon 46, 1011 Lausanne, Switzerland. (V. Dunet). Department of Clinical Neurosciences, Lausanne University Hospital, Rue du Bugnon 46, 1011 Lausanne, Switzerland (K. Diserens).
E-mail addresses: karin.diserens@chuv.ch (K. Diserens), vincent.dunet@chuv.ch (V. Dunet).

[1] These authors contributed equally to the manuscript.
[2] These authors contributed equally and act both as co-senior authors and corresponding authors.

https://doi.org/10.1016/j.nicl.2021.102651
Received 22 December 2020; Received in revised form 12 March 2021; Accepted 26 March 2021
Available online 29 March 2021

localization (Giacino et al., 2002). The emergence from the MCS is diagnosed once a patient functionally communicates or uses objects (Giacino et al., 2002).

The detection of consciousness in a severely brain damaged patient presents a major clinical challenge. Although bedside neurobehavioral assessment remains the gold standard, misdiagnosis occurs in up to 40% (Schnakers et al., 2009; Andrews et al., 1996)of cases due to fluctuating arousal, evaluation bias, comorbidities or other accompanying factors that limit the patient's interaction (Seel et al., 2010; Pincherle et al., 2021). Currently, the most sensitive scale discriminating between DOC subtypes is the JFK Coma Recovery Scale Revised (CRS-R) (Seel et al., 2010; Giacino et al., 2004). However, as its criteria rely rigorously on motor behavior, it may fail to detect consciousness in patients with sensorimotor, communication, or drive deficits (Gill-Thwaites, 2006; Majerus et al., 2009). Recently, we developed an observational instrument, the Motor Behavior Tool-revised (MBT-r) (Pignat et al., 2016) to detect subtle signs of conscious motor behavior, which are overlooked by the CRS-R assessment. MBT-r was able to discriminate patients with a cognitive motor dissociation condition (CMD) from true DOC patients (Pincherle et al., 2021, 2019; Jöhr et al., 2020) and had a superior outcome predictability than the CSR-R alone (Pignat et al., 2016; Jöhr et al., 2020).

CMD was first described by Owen et al. in 2006 in their pioneering functional magnetic resonance imaging (MRI) approach, where they reported a patient diagnosed with UWS who nevertheless, showed command-following brain activity modulation through mental imagery in a neuroimaging task (Owen et al., 2006). Consequently, CMD was proposed as a separate category from DOC, defining patients who are behaviorally unresponsive or minimally responsive due to a blocked motor efferent system, but who show command following in neuroimaging paradigms (Chapter, 2015; Schiff, 2015; Edlow et al., 2017). The estimated prevalence of CMD among DOC patients is 15% (Kondziella et al., 2016).

Task-dependent neuroimaging and electrophysiological methods detecting covert consciousness are not always applicable as they require a patient's wakefulness, collaboration, and understanding of the instructions, and the patient's vigilance fluctuation and language function impairment in the acute phase may hinder their performance. Moreover, the methods are time-consuming and have certain contraindications (Chapter, 2015). Nevertheless, early detection of consciousness is crucial as the presence of residual awareness has better outcome prediction (Pincherle et al., 2019; Jöhr et al., 2020; Di et al., 2008; Stender et al., 2014)and important implications for accurate diagnosis, prognosis and life-death decision management (Pincherle et al., 2021; Königs et al., 2018).

Here, we present an MRI-based scoring system constructed on the MBT-r diagnosis for use as a radiology tool accompanying the neurobehavioral assessment of DOC patients. We developed two classifiers using logistic regression (LR) and support vector machine (SVM) to discriminate between patients with and without residual consciousness based on their brain lesion patterns. We report their classification performance, compare it to the CRS-R classification, and discuss the possible clinical application.

2. Methods

2.1. Participants

This retrospective study was conducted in compliance with the ethical standards of the declaration of Helsinki. The local ethics committee (CER-VD, reference number: 142/09) approved the study protocol. Informed consent to participate in the study was obtained from each patient's legal representative.

We screened the hospital database for adult patients (age $\geq$ 16 years), with severe brain injury who were admitted to the Acute Neurorehabilitation Unit in the period between 01. 11. 2011 and 05. 11. 2019. All admitted patients were evaluated with the CSR-R (Giacino et al., 2004) scale at entry and at weekly follow-up during their stay in the unit as part of the standard clinical protocol. This assessment was complemented with the MBT-r (Pincherle et al., 2019) tool, administered on admission to the unit. The MBT-r tool was designed to detect subtle behavioral signs of preserved consciousness in behaviorally unresponsive patients according to the CSR-R scale, and might indicate a possible CMD. The tool consists of seven items that assess the presence of various motor signs (ocular, facial, limb, or oral) as an intentional response to stimulation or environment, and spontaneous non-reflexive movements. The MBT-r evaluation score is based on a non-cumulative binary classification: the presence of residual cognition is identified as soon as one positive sign is observed. Experienced neurologists and neuropsychologists carried out all clinical evaluations.

The initial search in the unit's database yielded 171 patients. Twenty-four patients were diagnosed as non-DOC and 3 with locked-in syndrome; they were omitted from the analyses. We further excluded 16 patients due to unavailable MRI scans. The final sample consisted of 128 patients (46 MCS and 82 coma/UWS as classified by the CRS-R), who were split in two groups based on the MBT-r category: patients with signs of residual cognition were included in the CMD group (n = 111), while patients without detected positive signs were included in the DOC group (n = 17). The patient selection flowchart is shown in Fig. 1.

2.2. MRI acquisition

The MRI data were acquired during the patient's stay in the Acute

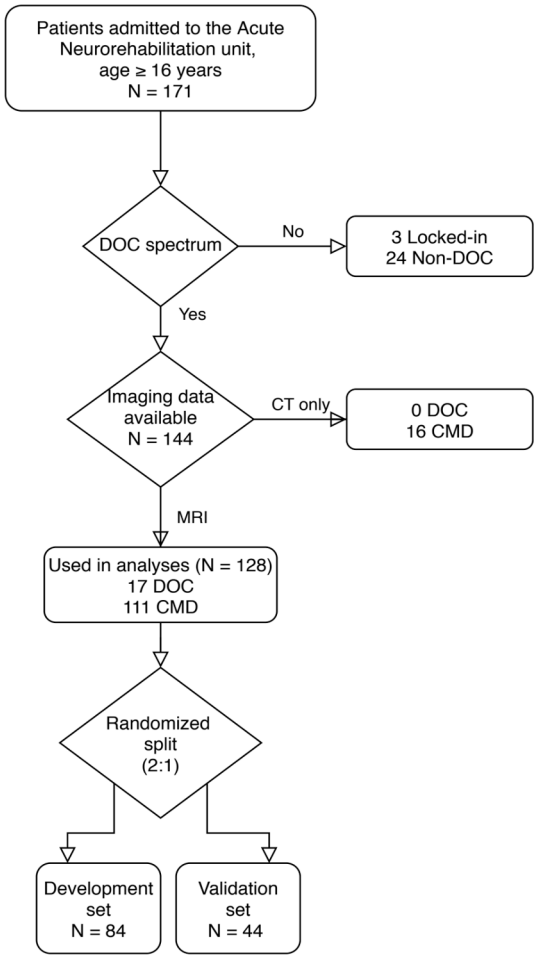

Fig. 1. Patient selection flowchart. Flowchart showing the selection procedure and number of included patients in each step of the procedure based on the clinical criteria.

Neuro-rehabilitation Unit, at the latest at patient discharge. The mean time between the initial injury and MRI scan was 32.1 ± 56.9 days (range: 0 – 418). The MRI scans were acquired on a 3 T Siemens scanner (TrioTim, Skyra or Prisma, Siemens Healthcare, Erlangen, Germany) situated in our institution (Lausanne University Hospital, Lausanne, Switzerland). The standard imaging protocol included a 3D T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) or a sagittal T1-weighted gradient echo sequence, a T2 axial spin-echo or a 3D fluid attenuated inversion recovery sequence, a T2 gradient echo or a susceptibility weighted sequence, a diffusion weighted sequence of the entire brain, and a 3D T1 MPRAGE when clinically relevant. The MR images were reviewed and evaluated on the institutional PACS viewing software (VuePacs, Carestream, Rochester, NY, USA).

2.3. MRI lesion evaluation

An experienced/senior neuroradiologist, blinded to the patients' clinical assessments and diagnoses, evaluated the patients' MRI scans by visual inspection. Four peripheral (frontal, temporal, parietal, and occipital lobe) and five deep (basal ganglia, thalamus, mesencephalon, pons and cerebellum) regions were evaluated bilaterally, resulting in 18 regions. Each region was given a binary score, as follows: 0, if the region was not affected or with only a small focal lesion, and 1, if diffuse (hemorrhagic traumatic axonal injury seen on susceptibility-weighted images or anoxic injury seen on diffusion-weighted images) or a large lesion (lesion volume ≥ 30% of the region) was observed. To note, the term "diffuse lesion" should not be confused with the term "widespread injury", also used in the text and referring to the number of affected brain regions.

The volume cut-off was defined based on the previous studies, which found that lesions affecting 30% or more of the lobar (Broderick et al., 1993) or basal ganglia volume (Rohaut et al., 2019) are significant predictors of a negative outcome in patients suffering from intracerebral hemorrhage. One half of the MRI data (n = 64 patients) was re-analyzed by a second neuroradiologist to estimate the inter-rater reliability. The same MRI data subset was re-evaluated by both neuroradiologist after 6 months to estimate the intra-rater reliability. Examples of diffuse and large brain lesions are shown in Fig. 2.

2.4. Statistical analyses

All statistical analyses were conducted with the R statistical software (version 4.0.0, R Foundation for Statistical Computing, Vienna, Austria). Continuous variables are presented as mean ± standard deviation (SD) and categorical variables as number or percentage. The MBT-r was considered as the standard of reference to classify patient outcome as CMD or DOC. The sample (n = 128), was randomly divided into two datasets with a 2:1 ratio: a dataset to develop the prediction model (development group, 84 patients) and a dataset to validate the model (validation group, 44 patients). The ratio of CMD and DOC patients was comparable between both groups. 2-tailed Fisher's exact test and Wilcoxon's unpaired test were conducted to test for the statistical differences between the development and validation groups for the most relevant clinical and demographic characteristics. Intra- and inter-rater reproducibility of the scoring was estimated with the Gwet's agreement coefficient1 index (Gwet, 2008) to overcome Cohen's kappa test paradoxes due to marginal distribution considerations (Wongpakaran et al., 2013).

The aim of the analyses was to build classification models and identify patients with possible residual cognition based on their brain lesion pattern. We used LR and SVM classifiers to select the most relevant brain regions predicting the patient's outcome. An MRI scoring system with an optimal cut-off score was created based on the selected predictors of each statistical method, and their classification performance was compared. We used the Caret (Kuhn, 2008) and e1071 (Dimitriadou et al., 2009) R packages for the development of predictive models, and the pROC package (Robin et al., 2011) for the receiver operating characteristic (ROC) analyses.

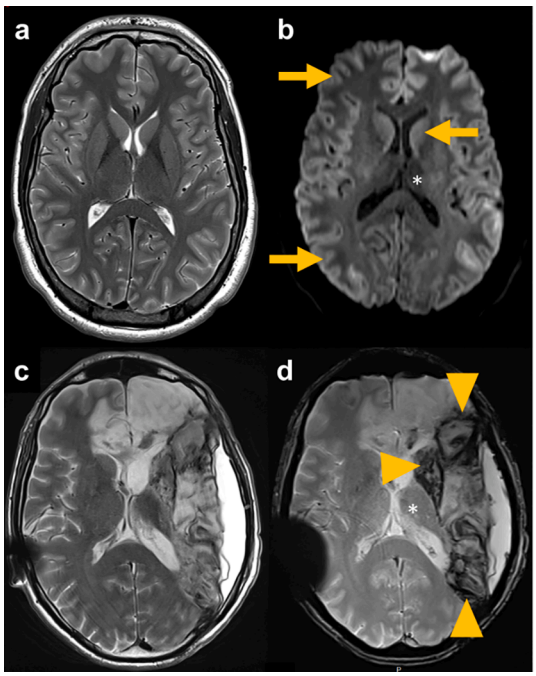

Fig. 2. Diffuse and large brain lesion examples. In a patient admitted after cardiac arrest, axial T2 weighted image (a) and diffusion (b) demonstrated bilateral diffuse (a and b) anoxic injury of the temporal, frontal, parietal lobes and of the basal ganglia (arrows) but not of the thalamus (asterisk). In a patient admitted for a malignant stroke complicated by hemorrhage, axial T2 weighted image (c) and T2 gradient echo (d) demonstrated a large lesion of the left frontal and temporal lobes and of the left basal ganglia involving>30% of their volumes. The left thalamus (asterisk) was not injured.

2.5. Logistic regression

We developed a logistic regression model for the patient's neurological outcome after brain injury (CMD or DOC) using the brain regions with large/diffuse lesions as independent predictors. The outcome was defined as the probability that a patient has no residual consciousness, i.e. has a true DOC. Using a univariable selection approach, we first performed univariate logistic regressions on the development group for each of the 18 brain regions. Statistically significant brain regions ($p < 0.05$) were then further analyzed with a multivariate regression. In favor of a richer predictive model, we used odds ratios (ORs) of all the predictors in the multivariate model to build the MRI score. We chose unadjusted odds ratios to account for an independent contribution of each selected brain region to the neurological outcome. A brain region with a large/diffuse lesion was assigned a weighted score based on its odds ratio in the following manner (Guédon et al., 2018): one point if $1.5 \leq OR < 2.5$, two points if $2.5 \leq OR < 4$, 3 points if $4 \leq OR < 6$, and 4 points if $OR \geq 6$.

We used the Youden's index (Youden, 1950) to estimate the optimal diagnostic cut-off point for the resulting cumulative scores. Patients whose cumulative score was equal or higher than the cut-off point were classified as mrDOC and patients with the score below the cut-off point as mrCMD. We then tested the logistic regression scoring system on the validation group, and evaluated its classification performance with ROC analyses.

2.6. Support vector machine (SVM) with recursive feature elimination

SVM is a machine learning method used to classify multidimensional

data by finding the best hyperplane to maximize the margin between two classes (Cortes and Vapnik, 1995). It is particularly suitable for clinical data as it does not require a linear relationship and independence between variables (Unnikrishnan et al., 2016), and is less sensitive to unbalanced data (Lin and Chen, 2012), as observed in our sample. To develop a predictive model, a set of optimal predictors was selected through recursive feature selection (Kuhn, 2008) using a linear SVM. To prevent model over fitting, we used a 5-fold cross-validation re-sampling with 5 repetitions. In each repetition, the development data set was divided in 5 equal parts. Each part was used once to validate the model, while the remaining 4 served for the model training. With this method, a full model with 18 predictors (brain regions) was entered in the iteration. For each fold, the most relevant subset of predictors was selected through backward elimination, where the subsets were ranked according to their classification accuracy derived from the validation part. The final set of predictors was determined based on the average rank of classification accuracy for all cross-validation folds.

The weights (w) of the predictors in the final model were transformed to construct the MRI scoring system: each weight value was multiplied by 10 and rounded to the nearest whole number. The patients in the validation group were then classified into mrCMD or mrDOC according to the cut-off score, estimated with the Youden's index. The classification performance was evaluated with the ROC curve.

2.7. Classifier performance comparison

We compared the LR and SVM classification performances on the validation dataset as well as on the whole dataset. The statistical differences were tested with the DeLong's test for 2 correlated ROC curves. The differences were considered significant at $p < 0.05$.

2.8. Including age, sex, and brain injury etiology in the predictive models

To account for a possible contribution of age, sex, and brain injury etiology to the classification accuracy, we included these variables to the LR and SVM model construction as described above and compared the accuracy of the extended models.

2.9. Classification based on the CRS-R diagnosis

We verified if a similar classification of brain lesion profiles could also be made based on the CRS-R diagnosis. We grouped the patients into coma/UWS, and MCS groups. As for the MBT-r, the CRS-R classification was used to train the LR and SVM classifiers on the development data set. The performance of the classifiers was inspected with the ROC curve analyses.

3. Results

3.1. Demographic and clinical characteristics of the patients

Overall, 128 patients were included (77 males, mean age = 51.1 ± 17.2 years). 48 patients suffered from traumatic brain injury, 33 from anoxia/ischemic stroke, 38 from brain hemorrhage, and 9 from other etiology (brain tumor, encephalopathy, and meningoencephalitis). The development and validation group did not significantly differ between the male to female ratio ($p = 0.58$), mean age ($p = 0.67$) and brain injury etiology distribution ($p = 0.14$). The mean elapsed time between the brain injury and the MRI exam was longer in the development group ($p = 0.03$). The groups did not differ in the mean elapsed time between the brain injury and the MBT-r evaluation ($p = 0.50$). The proportion of CMD and DOC patients did not differ between groups ($p > 0.99$), nor did the coma/UWS to MCS ratio, as diagnosed by the CRS-R ($p = 0.70$). On average, the DOC patients had a higher number of large/diffuse brain lesions, i.e. more widespread brain injury than the CMD patients in the development (DOC: mean = 6.9 ± 2.8, CMD: mean = 2.9 ± 2.4, $p < 0.001$) and validation groups (DOC: mean = 6.7 ± 4.1, CMD: mean = 2.2 ± 1.6, $p = 0.005$). In contrast, no differences in the number of lesions were found between the coma/UWS and MCS patients as diagnosed by the CRS-R in the development (coma/UWS: mean = 3.7 ± 2.8, MCS: mean = 3.3 ± 2.8, $p = 0.44$) and validation groups (coma/UWS: mean = 3.3 ± 1.2, MCS: mean = 2.0 ± 3.0, $p = 0.21$). More extensive statistical comparisons of the demographic and clinical data are reported in

Table 1
Demographic and clinical data for the development and validation data sets.

	Development group		Validation group		p value
N	84		44		
Age (years)					
Range (Mean ± SD)	17 – 83 (51.5 ± 17.8)		20 – 78 (50.3 ± 16.0)		0.667[a]
Sex (male/female)	49/35		28/16		0.576
N (%) of patients with lesioned brain region*	**CMD**	**DOC**	**CMD**	**DOC**	N/A
Frontal lobe left	27 (37%)	8 (73%)	10 (26%)	5 (83%)	
Frontal lobe right	32 (44%)	9 (82%)	12 (32%)	5 (83%)	
Temporal lobe left	20 (38%)	6 (55%)	7 (18%)	3 (50%)	
Temporal lobe right	25 (34%)	8 (73%)	10 (26%)	2 (33%)	
Parietal lobe left	16 (22%)	4 (36%)	4 (11%)	4 (67%)	
Parietal lobe right	12 (16%)	7 (64%)	4 (11%)	4 (67%)	
Occipital lobe left	6 (8%)	2 (18%)	3 (8%)	3 (50%)	
Occipital lobe right	10 (14%)	3 (27%)	3 (8%)	3 (50%)	
Basal ganglia left	12 (16%)	4 (36%)	5 (13%)	2 (33%)	
Basal ganglia right	12 (16%)	6 (55%)	4 (11%)	2 (33%)	
Thalamus left	3 (4%)	1 (9%)	0 (0%)	2 (33%)	
Thalamus right	7 (10%)	4 (36%)	1 (3%)	1 (17%)	
Mesencephalon left	7 (10%)	6 (55%)	3 (8%)	1 (17%)	
Mesencephalon right	8 (11%)	3 (27%)	4 (11%)	1 (17%)	
Pons left	3 (4%)	2 (18%)	5 (13%)	0 (0%)	
Pons right	3 (4%)	1 (9%)	4 (11%)	0 (0%)	
Cerebellum left	4 (5%)	1 (9%)	2 (5%)	1 (17%)	
Cerebellum right	7 (10%)	1 (9%)	1 (3%)	1 (17%)	
Elapsed time (days)					
Brain injury to MRI exam					
Range (Mean ± SD)	0 – 418 (36.4 ± 61.5)		1 – 314 (23.9 ± 46.4)		0.028[c]
Brain injury to MBT-R exam					
Range (Mean ± SD)	3 – 281 (31.5 ± 41.0)		4 – 77 (25.1 ± 15.9)		0.495[c]
MBT-R classification					
CMD	73		38		> 0.999
DOC	11		6		
CRS-R classification					
MCS	29		17		0.704
Coma/UWS	55		27		
Coma/UWS with CMD	45		21		0.551
MCS identified as DOC	1		0		>0.999

*Considered when a diffuse lesion or large lesion affecting at least 30% of the region volume was present. The comparisons were performed using the Fisher's exact test, except for when we used a) Mann-Whitney *U* test: W = 1931.5, b) X^2 test: df = 3, $X^2 = 5.45$, and c) Mann-Whitney *U* test: W = 2287.5, W = 1984.5, respectively. N/A = not applicable.

Table 1. The intra-rater reproducibility, re-assessed on a subset of 64 patients after a period of 6 months, was 0.95 [95% CI: 0.93, 0.96] for the first neuroradiologist and 0.82 [95% CI: 0.79, 0.85] for the second neuroradiologist. The overall inter-rater reproducibility was 0.76 [95% CI: 0.72, 0.80].

3.2. Logistic regression

Separate univariate logistic regressions showed that diffuse or large lesions in the left frontal cortex ($p = 0.04$), right frontal cortex ($p = 0.03$), right temporal cortex ($p = 0.02$), right parietal cortex ($p = 0.002$), right basal ganglia ($p = 0.01$), right thalamus ($p = 0.02$), and left mesencephalon ($p = 0.001$) were statistically significant predictors of a negative outcome (DOC), i.e. the absence of any signs of consciousness as identified by the MBT-r. A multivariate logistic analysis based on these seven statistically significant brain regions showed that lesions in the left mesencephalon (adjusted OR: 29.0 [95% CI: 2.3–361.9], $p = 0.01$) and in the right basal ganglia (adjusted OR: 26.2 [95% CI: 1.8–392.7], $p = 0.02$) are significant independent predictors for DOC as identified by the MBT-r. Detailed statistical results for all brain regions are presented in Table 2.

The odds ratios of the significant brain regions used for the construction of the MRI scoring system are shown in Table 3. The CMD patients (score range: 0 – 17, mean = 5.4 ± 4.9) had a significantly lower score than the DOC patients (score range: 4 – 21, mean = 14.8 ± 5.1; Wilcoxon unpaired test: $W = 79.5$, $p < 0.001$).

The optimal diagnostic cut-off point based on Youden's index corresponded to a score of 10. Patients with a score equal to or higher than the cut-off point were classified as without residual consciousness (DOC), and patients below the cut-off as CMD. The accuracy of the classification was 0.81, sensitivity for predicting DOC 0.91, and specificity 0.80. The classifier also correctly recognized 39 out of 45 patients diagnosed as unconscious (UWS or coma) by the CRS-R, but recognized as CMD by the MBT-R.

When applied to the validation data set, the accuracy of the logistic regression classifier increased to 0.87, the sensitivity decreased to 0.67, and the specificity increased to 0.90. The classifier correctly detected 19 out of 21 CMD patients with the coma/UWS diagnosis by the CRS-R in the validation group. Detailed ROC performance parameters of the training and validation groups are displayed in Table 4.

3.3. Support vector machine (SVM) with recursive feature elimination

Six brain regions (left frontal cortex, left mesencephalon, right parietal cortex, right basal ganglia, right thalamus, and left temporal cortex) were selected through recursive feature selection as the optimal subset of predictors of the MBT-r category (RMSE = 0.34, $R^2 = 0.14$). The MRI scoring system, constructed on the basis of weights of the selected predictors in the final SVM model is shown in Table 3. The CMD patients (score range: 0 – 22, mean = 6.4 ± 6.1) had a significantly lower MRI score than the DOC patients (score range: 5 – 24, mean = 18.5 ± 5.5; Wilcoxon unpaired test: $W = 62.5$, $p < 0.001$).

The Youden's index cut-off point was estimated at a score of 15. The classification accuracy at this cut-off point was 0.86, the sensitivity for predicting the negative outcome 0.91, and specificity 0.85. Classification of patients in the validation group according to the cut-off point resulted in an increase in accuracy to 0.93, whereas sensitivity decreased to 0.67, and specificity increased to 0.97. The classifier was able to correctly classify 40 out of 45 CMD patients with the coma/UWS diagnosis by the CRS-R in the development group, and 20 out of 21 CMD patients in the validation group. The detailed performance parameters of the SVM classifier are shown in Table 4.

Table 2
Results of the univariate logistic regression analysis.

Brain region	Univariate logistic regression analysis	
	Coefficient	p – value
Frontal lobe left	**1.51**	**0.035**
Frontal lobe right	**1.75**	**0.032**
Temporal lobe left	1.16	0.080
Temporal lobe right	**1.63**	**0.023**
Parietal lobe left	0.71	0.301
Parietal lobe right	**2.19**	**0.002**
Occipital lobe left	0.91	0.307
Occipital lobe right	0.86	0.257
Basal ganglia left	1.07	0.129
Basal ganglia right	**1.81**	**0.008**
Thalamus left	0.85	0.481
Thalamus right	**1.68**	**0.023**
Mesencephalon left	**2.43**	**< 0.001**
Mesencephalon right	1.11	0.150
Pons left	1.65	0.093
Pons right	0.85	0.481
Cerebellum left	0.55	0.641
Cerebellum right	−0.06	0.958

Table 3
MRI DOC scoring system for LR and SVM classifiers.

Logistic regression			Support vector machine		
Predictor	OR	Score	Predictor	w	Score
Mesencephalon left	11.31	4	Frontal left	0.77	8
Parietal right	8.90	4	Parietal right	0.66	7
Basal ganglia right	6.10	4	Basal ganglia right	0.64	6
Frontal right	5.77	3	Mesencephalon left	0.45	5
Thalamus right	5.39	3	Thalamus right	0.26	3
Temporal right	5.12	3	Temporal left	0.17	2
Frontal left	4.54	3	—	—	—
Total score		**24**	**Total score**		**31**
Youden's cut-off point		**10**	**Youden's cut-off point**		**15**

OR = odds ratio, LR = logistic regression, SVM = support vector machine

3.4. Classifier performance comparison

The performance of the LR and SVM classifiers did not significantly differ in the validation group ($Z = -1.8$, $p = 0.08$), whereas the SVM classifier showed statistically superior performance in the entire cohort (validation and training set combined, $Z = -2.1$, $p = 0.03$). The ROC curves showing the LR and SVM MRI scores and the optimal cut-off classification performances are presented in Fig. 3. The scatter plot of the two MRI scoring systems is shown in Fig. 4.

3.5. Including age, sex, and brain injury etiology in the predictive models

Univariate logistic regressions showed that age significantly predicted the outcome ($\beta = -0.83$, $z = -3.3$, $p < 0.001$, OR = 0.92 [95% CI: 0.88, 0.97]), whereas sex and etiology were non-significant predictors ($p > 0.5$). Before including age as a factor in the extended logistic model, we discretized the continuous variable into a binary factor using OneR package (von Jouanne-Diedrich, 2017) in order to have comparable scales between all the predictors. As such age was dichotomized with a resulting cut-off point that was optimally aligned with the target outcome (39 years). The odds ratio of binarized age factor (OR = 13.6 [95% CI: 3.14, 58.61) were transformed into points (4 points) and added to the MRI scoring system. The optimal cut-off point for classification based on Youden's index was 14. When validated on the test data set, this classification resulted in accuracy of 0.89 [95% CI: 0.75, 0.96], sensitivity of 0.67 [95% CI: 0.22, 0.96], and specificity of 0.92 [95% CI: 0.79, 0.98]. However, when comparing the extended LR model (AUC = 0.79) to the initial LR model (AUC = 0.78), the increase in the AUC was not statistically significant (DeLong's test: $z = 1$, $p = 0.32$).

Furthermore, we added age, sex, and etiology to the 18 brain regions in the SVM recursive feature selection with backward elimination. The final SVM model was built based on the most relevant subset of

Table 4
Classification performance of LR and SVM classifiers on the development group, validation group, and whole data set when using the optimal cut-off score.

	Logistic regression			Support vector machine		
Performance parameter (95% CI)	Development group	Validation group	Whole data set	Development group	Validation group	Whole data set
AUC	0.85 (0.75–0.95)	0.78 (0.57–0.99)	0.83 (0.73–0.93)	0.88 (0.78–0.98)	0.82 (0.61–1.00)	0.86 (0.76–0.96)
Sensitivity	0.91(0.59–1.00)	0.67(0.22 – 0.96)	0.82(0.57–0.96)	0.91(0.57–1.00)	0.67(0.24–0.94)	0.82(0.56–0.95)
Specificity	0.80(0.68–0.88)	0.90(0.75–0.97)	0.83(0.75–0.89)	0.85(0.74–0.92)	0.97(0.85–1.00)	0.89(0.82–0.94)
Accuracy	0.81(0.71–0.89)	0.87(0.73–0.95)	0.83(0.75–0.89)	0.86(0.76–0.92)	0.93(0.81–0.96)	0.88(0.81–0.93)
PLR	4.42(2.72–7.21)	6.33(2.14–18.8)	4.81(3.02–7.66)	6.03(3.39–10.73)	25.3(3.38–190)	7.62 (4.27–13.58)
NLR	0.11(0.02–0.74)	0.37(0.12–1.16)	0.21(0.08–0.60)	0.11(0.02–0.70)	0.34(0.11–1.06)	0.20(0.07–0.55)
PPV	0.40(0.21–0.61)	0.50(0.16–0.84)	0.42(0.26–0.61)	0.48(0.26–0.70)	0.80(0.30–0.99)	0.54(0.34–0.73)
NPV	0.98(0.91–0.99)	0.94(0.81–0.99)	0.97(0.91–0.99)	0.98(0.90–0.99)	0.95(0.81–0.99)	0.97(0.91–0.99)
CMD with the coma/UWS detected by the classifier	39/45	19/21	58/66	40/45	20/21	60/66

AUC = area under the curve, PLR = Positive likelihood ratio, NLR = Negative likelihood ratio, PPV = Positive predictive value, NPV = Negative predictive value, CMD = Cognitive Motor Dissociation, 95% CI = 95% confidence interval

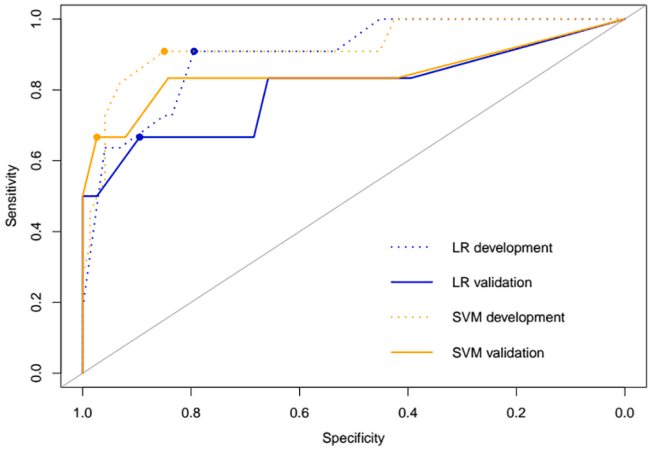

Fig. 3. ROC curve comparison between the MRI scores constructed by the LR and the SVM methods. ROC curves of the MRI score classification performance estimated from the logistic regression (LR) and support vector machine (SVM) classifications. Classification performance is shown for the development group (dotted line) and for the validation group (solid line). The dot on each curve represents the estimated optimal cut-off point.

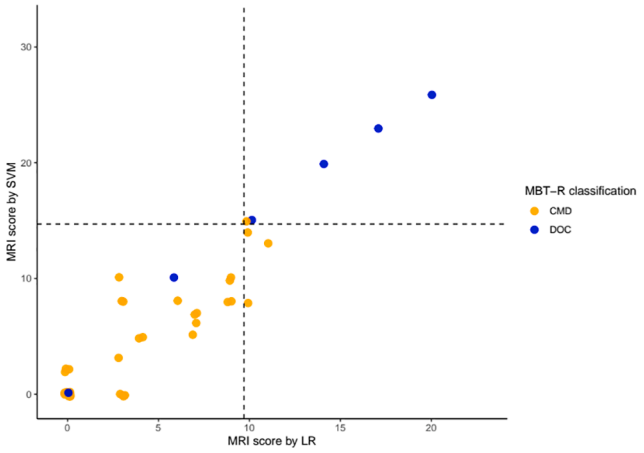

Fig. 4. MRI score and classification by the LR and SVM classifiers on the validation group. Scatter plot of the MRI scores of patients in the validation group. Each data point corresponds to a patient with the MRI score derived from the logistic regression (LR, x-axis) and from the support vector machine (SVM, y-axis) classifier. The dashed lines represent the diagnostic cut-offs for each MRI score. The data points are randomly jittered (by ± 0.5) for a better visual discrimination of the close data points.

predictors selected (RMSE = 0.33, R^2 = 0.21). The final model thus included age (dichotomised, w = 0.85), left mesencephalon (w = 0.73), right parietal lobe (w = 0.21), right basal ganglia (w = 0.62), right thalamus (w = 0.17), and left temporal lobe (w = 0.23). The weights of the final model were transformed into points, which were then used to construct the MRI scoring system, with the Youden index-based cut-off point at 13. Validating this classification on the test set resulted in accuracy of 0.82 [95% CI: 0.67, 0.92], sensitivity of 0.17 [95% CI: 0, 0.64) and specificity of 0.92 [95% CI: 0.79, 0.98). The performance of the alternative SVM predictive model including age as a factor (AUC = 0.54) was significantly worse than the initial SVM model (AUC = 0.82; DeLong's test: z = −2.4, p = 0.02).

3.6. Misclassified cases

We carried out a post-hoc inspection of the misclassified cases in the entire cohort. Three DOC patients incorrectly classified as CMD had low MRI scores, and were characterized with bilateral lesions in the mesencephalon, cerebellum, or frontal lobes. In contrast, misclassified CMD patients had higher MRI scores due to more widespread lesions in mostly supratentorial regions. Eight out of 21 (38%) misclassified CMD patients also had lesions in the mesencephalon, although unilateral. The extent of the lesion in the mesencephalon might be an important indicator of a negative outcome. In fact, when considering the entire cohort, mesencephalon lesions were more frequent in the DOC patients (76%, with 24% being bilateral) than in the CMD patients (33%, with 3% being bilateral).

3.7. Classification based on the CRS-R diagnosis

The univariate logistic regressions showed that none of the brain regions were a significant predictor of the CRS-R diagnosis (all p > 0.15). An SVM method with recursive feature elimination identified a model with only the left basal ganglia as the most relevant predictor (RMSE = 0.59, R^2 = 0.06). Its weight coefficient was however, of a negligible size and with a negative value (-5.5^{e-9}). For this reason, we did not construct a scoring system, as its classification performance would be highly unstable due to a low coefficient value.

4. Discussion

The aim of the present study was to build an MRI score for qualitative evaluation of brain lesions and classification of patients with residual consciousness as defined by the MBT-r. Using LR and SVM classifiers, we

identified a combination of damaged brain structures with a high relevancy to discriminate between patients with and without residual cognition. The identified lesions overlapped between the two classifiers and included the left mesencephalon, right basal ganglia, right thalamus, left frontal lobe, and right parietal lobe. Both classifiers showed an overall good performance, with the SVM having a higher specificity level.

Our analyses showed that LR and SVM methods both retained high classification accuracy when applied to the validation data set. The sensitivity level, i.e., correct classification of the patients who were identified as DOC by the MBT-r, was identical between both classifiers, whereas the SVM model performed better in correctly detecting the patients identified as CMD by the MBT-r (specificity level). The classifiers' performances did not statistically differ when they were tested on the validation dataset; however, the SVM model demonstrated better performance when applied to the entire dataset. This superior performance was due to a lower number of misclassified CMD patients by the SVM method. The two classifiers were comparable in the ability to accurately detect the CMD patients who were diagnosed as unresponsive by the CRS-R, while showing residual cognition with the MBT-r.

The brain regions that were selected as relevant predictors in the LR and SVM models were highly overlapping. Both classifiers included the left mesencephalon, right basal ganglia, right thalamus, right parietal cortex, and left frontal cortex. Large or diffuse lesions in these structures importantly predicted the negative outcome, i.e. the absence of any signs of residual awareness. The selection of these structures agrees with the current understanding of the underlying neuroanatomical changes in DOC. For example, coma is characterized by lesions to the brainstem reticular formation, including the pons and mesencephalon (Posner et al., 2007; Parvizi and Damasio, 2003), which regulate arousal levels; by bilateral injuries to the thalamic nuclei; and by large or diffuse hemispheric grey or white-matter lesions (Rohaut et al., 2019; Parvizi and Damasio, 2003; Laureys et al., 2004). On the other hand, while the brainstem is usually preserved in UWS, maintaining arousal and autonomic functions, the cerebral white and grey matter are extensively damaged (Laureys et al., 2004; Kampfl et al., 1998) underlying the loss of awareness. A distinctive characteristic of UWS also includes a reduced metabolism in the multimodal associative areas, such as the prefrontal cortex and parieto-temporo-occipital junction, which are involved in most of the higher cognitive functions (Laureys et al., 2004; Laureys, 2005). A growing number of studies have also demonstrated a critical role of thalamo-cortical connections (Maxwell et al., 2004; Fernández-Espejo et al., 2011) within the fronto-parietal (Long et al., 2016; Crone et al., 2014) and default mode brain networks (Fernández-Espejo et al., 2012; Vanhaudenhuyse et al., 2009) in the sustenance of awareness. Our findings also align with the study of Morozova et al. (2018)., who developed a diagnostic MRI scoring system differentiating between the chronic UWS and MCS patients based on brain atrophy. The chronic UWS was marked by a degeneration of the brainstem and thalamus, widespread diffuse cortical atrophy, and degeneration of the corpus callosum, pointing to the importance of the preserved integrity of these brain regions in the sustenance of consciousness.

Our findings also demonstrate that the DOC patients showed a more widespread injury compared to the CMD patients. Associations between the extent of injury and the severity of the outcome in the DOC spectrum have been recognized before (Laureys et al., 2004; Laureys, 2005; Bodart et al., 2018). As global conscious experience does not depend on a single isolated brain region but is thought to rely on a preserved functional integrity of a widespread bilateral cortico-thalamo-cortical network (Laureys et al., 2004; Laureys, 2005), an extensive structural and/or functional disconnection might be a hallmark of DOC. In contrast, it has been suggested that CMD is characterized by specific damage to the motor execution system, mainly including the ventrolateral thalamic nucleus and its cortical projections to the motor areas (Bardin et al., 2011; Fernández-Espejo et al., 2015). The cortico-subcortical connectivity is thus supposedly less impaired in CMD patients, presenting a greater capacity of neuroplasticity for functional recovery. This is also in line with our clinical observations showing better cognitive and functional evolution in the CMD patients, in particular after recovery of the lesions in the strategic loci for motor planning (Pincherle et al., 2021; Jöhr et al., 2020).

Mostly, the right-hemisphere brain regions were selected as relevant predictors. We believe this finding is mainly driven by the lesion pattern characteristic of our DOC sample. Nevertheless, partial impairments of self-awareness, such as anosognosia or asomatognosia are often observed after lesions to the right hemisphere, mainly located in the parietal cortex (Berlucchi and Aglioti, 1997). Further studies with a larger proportion of DOC patients are needed to clarify the lesion lateralization found in our results.

In addition, another predictive region, the left frontal cortex, is largely specialized in language functions and supports a vast network for language comprehension and production (Gernsbacher and Kaschak, 2003). This network is distributed in the associative zones of the left perisylvian cortex and neighboring regions, as well as in other frontal regions facilitating motor planning and executive processing. Extensive damage to these crucial language and executive regions severely impacts an individual's ability to communicate the content of consciousness and to interact with his or her environment. Severely decreased metabolism (>50%) in these areas is observed in MCS patients (Laureys et al., 2004), and appeared more reduced in MCS− compared to MCS+ (Majerus et al., 2009). A disconnection of the Broca's area essential for spoken language formation from the rest of the language processing regions was also demonstrated in these patients (Boly et al., 2004). Finally, reduced activation to auditory stimuli reflecting residual neural coding without superior integration is observed in chronic UWS (Boly et al., 2004).

To summarize, the brain regions that we identified here with machine learning algorithms to best differentiate between the CMD and DOC patients, have often been reported in the literature as important for the maintenance of consciousness. In addition, the fact that the MBT-r diagnosis based on behavioral signs is well reflected in the neurostructural differences indicates a very probable neural basis for the MBT-r classification. This points to an important sensitivity of the MBT-r tool. In contrast, when we used the CRS-R diagnosis as a reference, we failed to identify brain lesion differences between the coma/UWS and MCS patients. Furthermore, these two groups could not be distinguished by the number of lesioned brain structures, while the number of observed lesions was significantly higher in the DOC patients when the patients were grouped according to the MBT-r diagnosis. However, we cannot exclude that quantitative lesion volume evaluation would improve the ability to accurately classify the patients based on their CRS-R diagnosis.

The present study has several limitations that merit discussion. The major limitation of the present research is the unbalanced data as the sample size of the DOC patients in our study was relatively low (15%) and it may have biased the classifiers' performance. The small proportion in our study reflects the general low prevalence of the DOC patients in Swiss healthcare institutions (Stretti et al., 2018) due to the state's policy on the use of limitation of life-prolonging therapies. To minimize the bias of data imbalance, we used the SVM method, which was shown to be less sensitive to unbalanced classes (Lin and Chen, 2012). Other additional methods are also available for managing class imbalance, such as for example oversampling the minority class, however they have certain disadvantages such as model overfitting (Weiss et al., 2007).

Furthermore, in order to simplify the clinical use of the MRI score, we constructed our classification models on qualitative brain lesion analyses, where only diffuse lesions or large lesions covering at least 30% of the region were considered. As such, we did not account for small focal lesions, which may also have an impact on the consciousness outcome (Li and Feng, 2009). In addition, the estimation of lesion volumes was performed by visual inspection. Despite a good intra- and inter-rater reliability, the qualitative approach might have introduced additional bias in the current study. This shortcoming could be improved by using

quantitative lesion estimation with automated segmentation and potentially enable a fully automatized MRI scoring procedure. However, due to important brain structural changes in the patients with severe brain injury (atrophy, mass effect, and swelling), application of automated lesion segmentation to this patient's population represents great methodological considerations. Next to it, disorders of consciousness may occur due to heterogeneous brain injury etiologies and thus result in different imaging patterns. Therefore, absolute lesion volume quantification could not be performed in a single manner for every patient without software developed to this end.

Moreover, our analyses could not encompass disturbances in the connectivity between relevant brain regions, which are otherwise important predictors of the DOC outcome (Fernández-Espejo et al., 2012; Bodart et al., 2018). For the same reason, we were not able to identify the pathophysiology underlying the CMD condition as our method of qualitative lesion analysis was not designed to detect fine-grained structural and connectivity changes. Therefore, further studies should aim at investigating the underlying morphological and connectivity differences between CMD and DOC patients using multimodal neuroimaging technology, for example, functional MRI, diffusion tensor imaging, and positron emission tomography.

We tested whether inclusion of non-radiological factors, i.e. age, sex and etiology would contribute to the accuracy of our classifiers. The analyses showed that the age was a relevant predictor of the outcome, as the DOC group had a lower age average. After including age as a factor in the final model, we observed a statistically non-significant increase in the specificity of the LR classifier, whereas the sensitivity of the SVM classifier significantly decreased. As the present study focused on MRI lesion based classification, we have not considered other clinical variables that might be relevant for the detection of patients with residual consciousness. Inclusion of a wide range of factors such as different patient's characteristics, details on brain injury and its severity, neurological exam scores, electrophysiological measures and computed tomography exams, as well as other biomarkers and laboratory variables would permit to develop a more comprehensive, multifaceted, and possibly more accurate predictive model for classification and prognostication of patients with severe brain injury (for reviews see (Lingsma et al., 2010; Sandroni et al., 2020)).

Finally, the use of the present MRI score in a clinical setting should be done with caution, as more data is needed to validate it and confirm its generalizability.

The limitations of the present study could be potentially addressed in the future research by using deep learning methods. In contrast to conventional supervised machine learning algorithms, deep learning does not require structured or labeled data in order to parse or classify the data. Instead, the unstructured data is passed through multiple interconnected layers of networks that are able to discover inherent patterns within the data and generate complex, higher-level representations of the initial input (Mazurowski et al., 2019; McBee et al., 2018). Thus in the context of the current study, the advantage of deep learning over traditional machine learning methods would be the ability to use raw imaging data, automatically quantify it based on the variety of radiomic features, and thus avoid qualitative lesion assessment and feature selection decision-making. The techniques of data augmentation within deep learning (Shorten and Khoshgoftaar, 2019) could also potentially reduce the class imbalance bias due to low prevalence of unresponsive DOC patients. Moreover, the classification could be performed over a great number of variables, including non-imaging clinical data as well. However, the limitation of this approach is the need of very large datasets for training and validation, which necessitates multicenter studies and collaborations. As such, application of deep learning methods for diagnosing and prognostication of patients with disorders of consciousness represents a great challenge for the future research.

4.1. Conclusion

In conclusion, our findings suggest that the MRI score is a beneficial complement to the behavioral assessment of patients with residual consciousness by identifying underlying imaging patterns that may help to confirm putative damages to the motor efferent system and preserved cortico-subcortical connectivity in CMD as opposed to the widespread cortico-thalamic damage often observed in true DOC. The advantages of the MRI score are that it only requires conventional MRI images without preprocessing; it is simple to use and time-efficient, and demonstrated good inter- and intra-rater reproducibility. Moreover, the score was constructed by training the classification algorithms on the diagnosis of the new neurobehavioral tool, MBT-r, which showed a better sensitivity than the CRS-R for detection of conscious behavior and accuracy to predict a patient's recovery (Pignat et al., 2016; Pincherle et al., 2019).

Understanding the neural biomarkers of covert consciousness is essential for the improvement of diagnostic accuracy in DOC and consequent planning of immediate access to neuro-rehabilitative interventions. Finally, since the recognition of neural imprints of residual cognition in behaviorally unresponsive patients has a direct implication on life-death decisions, the medico-ethical importance of establishing reliable and clinically applicable measures remains a large motivator of this research.

CRediT authorship contribution statement

Polona Pozeg: Conceptualization, Methodology, Visualization, Writing - original draft. **Jane Jöhr:** Conceptualization, Investigation, Data curation. **Alessandro Pincherle:** Investigation. **Guillaume Marie:** Investigation. **Philippe Ryvlin:** . **Reto Meuli:** Conceptualization, Supervision. **Patric Hagmann:** Conceptualization. **Karin Diserens:** Conceptualization, Investigation, Supervision, Funding acquisition. **Vincent Dunet:** Conceptualization, Investigation, Methodology, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Dr. Melanie Price Hirt for proofreading this manuscript. This study was funded by the Swiss National Science Foundation [grant number: FNS 320030_189129].

References

Andrews, K., Murphy, L., Munday, R., et al., 1996. Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. BMJ. 313(7048):13 LP - 16. doi:10.1136/bmj.313.7048.13.

Bardin, J.C., Fins, J.J., Katz, D.I., et al., 2011. Dissociations between behavioural and functional magnetic resonance imaging-based evaluations of cognitive function after brain injury. Brain 134(3):769-782. doi:10.1093/brain/awr005.

Berlucchi, G., Aglioti, S., 1997. The body in the brain: neural bases of corporeal awareness. Trends Neurosci. 20 (12), 560–564. https://doi.org/10.1016/S0166-2236(97)01136-3.

Bodart, Olivier, Amico, Enrico, Gómez, Francisco, Casali, Adenauer G., Wannez, Sarah, Heine, Lizette, Thibaut, Aurore, Annen, Jitka, Boly, Melanie, Casarotto, Silvia, Rosanova, Mario, Massimini, Marcello, Laureys, Steven, Gosseries, Olivia, 2018. Global structural integrity and effective connectivity in patients with disorders of consciousness. Brain Stimul. 11 (2), 358–365. https://doi.org/10.1016/j.brs.2017.11.006.

Boly, Mélanie, Faymonville, Marie-Elisabeth, Peigneux, Philippe, Lambermont, Bernard, Damas, Pierre, Del Fiore, Guy, Degueldre, Christian, Franck, Georges, Luxen, André, Lamy, Maurice, Moonen, Gustave, Maquet, Pierre, Laureys, Steven, 2004. Auditory processing in severely brain injured patients: differences between the minimally conscious state and the persistent vegetative state. Arch. Neurol. 61 (2), 233. https://doi.org/10.1001/archneur.61.2.233.

Broderick, J.P., Brott, T.G., Duldner, J.E., Tomsick, T., Huster, G., 1993. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. Stroke 24 (7), 987–993. https://doi.org/10.1161/01.STR.24.7.987.

Owen, A.M., 2015. Chapter 18 - Using functional magnetic resonance imaging and electroencephalography to detect consciousness after severe brain injury. in: Grafman, J., Salazar, AMBT-H of CN, eds. Traumatic Brain Injury, Part I. Vol 127. Elsevier; 2015:277-293. doi:10.1016/B978-0-444-52892-6.00018-0.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach Learn. 20 (3), 273–297. https://doi.org/10.1007/BF00994018.

Crone, Julia Sophia, Soddu, Andrea, Höller, Yvonne, Vanhaudenhuyse, Audrey, Schurz, Matthias, Bergmann, Jürgen, Schmid, Elisabeth, Trinka, Eugen, Laureys, Steven, Kronbichler, Martin, 2014. Altered network properties of the fronto-parietal network and the thalamus in impaired consciousness. NeuroImage Clin. 4, 240–248. https://doi.org/10.1016/j.nicl.2013.12.005.

Di, H., Boly, M., Weng, X., Ledoux, D., Laureys, S., 2008. Neuroimaging activation studies in the vegetative state: predictors of recovery? Clin. Med. 8 (5), 502–507. https://doi.org/10.7861/clinmedicine.8-5-502.

Dikmen, S.S., Machamer, J.E., Powell, J.M., et al., 2003. Outcome 3 to 5 years after moderate to severe traumatic brain injury[1]. Arch. Phys. Med. Rehabil. 84 (10), 1449–1457. https://doi.org/10.1016/S0003-9993(03)00287-9.

Dimitriadou, E., Hornik, K., Leisch, F., et al., 2009. E1071: Misc Functions of the Department of Statistics (E1071), TU Wien. In: R Package Version 1.5-24. Vol 1.

Edlow, B.L., Chatelle, C., Spencer, C.A., et al., 2017. Early detection of consciousness in patients with acute severe traumatic brain injury. Brain 140(9):2399-2414. doi: 10.1093/brain/awx176.

Fernández-Espejo, Davinia, Bekinschtein, Tristan, Monti, Martin M., Pickard, John D., Junque, Carme, Coleman, Martin R., Owen, Adrian M., 2011. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. Neuroimage 54 (1), 103–112. https://doi.org/10.1016/j.neuroimage.2010.08.035.

Fernández-Espejo, D., Rossit, S., Owen, A.M., 2015. A thalamocortical mechanism for the absence of overt motor behavior in covertly aware patients. JAMA Neurol. 72 (12), 1442–1450. https://doi.org/10.1001/jamaneurol.2015.2614.

Fernández-Espejo, Davinia, Soddu, Andrea, Cruse, Damian, Palacios, Eva M., Junque, Carme, Vanhaudenhuyse, Audrey, Rivas, Eva, Newcombe, Virginia, Menon, David K., Pickard, John D., Laureys, Steven, Owen, Adrian M., 2012. A role for the default mode network in the bases of disorders of consciousness. Ann. Neurol. 72 (3), 335–343. https://doi.org/10.1002/ana.23635.

Gernsbacher, Morton Ann, Kaschak, Michael P., 2003. Neuroimaging studies of language production and comprehension. Annu. Rev. Psychol. 54 (1), 91–114. https://doi.org/10.1146/annurev.psych.54.101601.145128.

Giacino JT, Ashwal S, Childs N, et al. The minimally conscious state. Neurology. 2002;58 (3):349 LP - 353. doi:10.1212/WNL.58.3.349.

Giacino, J.T., Kalmar, K., Whyte, J., 2004. The JFK coma recovery scale-revised: measurement characteristics and diagnostic utility[1]. Arch. Phys. Med. Rehabil. 85 (12), 2020–2029. https://doi.org/10.1016/j.apmr.2004.02.033.

Gill-Thwaites, H., 2006. Lotteries, loopholes and luck: misdiagnosis in the vegetative state patient. Brain Inj. 20 (13–14), 1321–1328. https://doi.org/10.1080/02699050601081802.

Guédon, A., Blauwblomme, T., Boulouis, G., Jousset, C., Meyer, P., Kossorotof, M., Bourgeois, M., Puget, S., Zerah, M., Oppenheim, C., Meder, J.-F., Boddaert, N., Brunelle, F., Sainte-Rose, C., Naggara, O., 2018. Predictors of outcome in patients with pediatric intracerebral hemorrhage: development and validation of a modified score. Radiology 286 (2), 651–658. https://doi.org/10.1148/radiol.2017170152.

Gwet, K.L., 2008. Computing inter-rater reliability and its variance in the presence of high agreement. Br. J. Math. Stat. Psychol. 61 (1), 29–48. https://doi.org/10.1348/000711006X126600.

Jöhr, J., Halimi, F., Pasquier, J., et al., 2020. Recovery in cognitive motor dissociation after severe brain injury: A cohort study. PLoS One. 15(2):e0228474. DOI:10.1371/journal.pone.0228474.

Kampfl, Andreas, Schmutzhard, Erich, Franz, Gerhard, Pfausler, Bettina, Haring, Hans-Peter, Ulmer, Hanno, Felber, Stefan, Golaszewski, Stefan, Aichner, Franz, 1998. Prediction of recovery from post-traumatic vegetative state with cerebral magnetic-resonance imaging. Lancet 351 (9118), 1763–1767. https://doi.org/10.1016/S0140-6736(97)10301-4.

Kondziella, D., Friberg, C.K., Frokjaer, V.G., et al., 2016. Preserved consciousness in vegetative and minimal conscious states: systematic review and meta-analysis. J Neurol Neurosurg & Psychiatry. 87(5):485 LP - 492. doi:10.1136/jnnp-2015-310958.

Königs, M., Beurskens, E.A., Snoep, L., Scherder, E.J., Oosterlaan, J., 2018. Effects of timing and intensity of neurorehabilitation on functional outcome after traumatic brain injury: a systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 99 (6), 1149–1159.e1. https://doi.org/10.1016/j.apmr.2018.01.013.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26.

Laureys, S., 2005. The neural correlate of (un)awareness: lessons from the vegetative state. Trends Cogn. Sci. 9 (12), 556–559. https://doi.org/10.1016/j.tics.2005.10.010.

Laureys, S., Owen, A.M., Schiff, N.D., 2004. Brain function in coma, vegetative state, and related disorders. Lancet Neurol. 3 (9), 537–546. https://doi.org/10.1016/S1474-4422(04)00852-X.

Laureys, S., Perrin, F., Faymonville, M -E., Schnakers, C., Boly, M., Bartsch, V., Majerus, S., Moonen, G., Maquet, P., 2004. Cerebral processing in the minimally conscious state. Neurology 63 (5), 916–918. https://doi.org/10.1212/01.WNL.0000137421.30792.9B.

Li, X.-Y., Feng, D.-F., 2009. Diffuse axonal injury: novel insights into detection and treatment. J. Clin. Neurosci. 16 (5), 614–619. https://doi.org/10.1016/j.jocn.2008.08.005.

Lin, W.-J., Chen, J.J., 2012. Class-imbalanced classifiers for high-dimensional data. Brief Bioinform. 14 (1), 13–26. https://doi.org/10.1093/bib/bbs006.

Lingsma, Hester F, Roozenbeek, Bob, Steyerberg, Ewout W, Murray, Gordon D, Maas, Andrew IR, 2010. Early prognosis in traumatic brain injury: from prophecies to predictions. Lancet Neurol. 9 (5), 543–554. https://doi.org/10.1016/S1474-4422(10)70065-X.

Long, Jinyi, Xie, Qiuyou, Ma, Qing, Urbin, M.A., Liu, Liqing, Weng, Ling, Huang, Xiaoqi, Yu, Ronghao, Li, Yuanqing, Huang, Ruiwang, 2016. Distinct interactions between fronto-parietal and default mode networks in impaired consciousness. Sci. Rep. 6 (1) https://doi.org/10.1038/srep38866.

Majerus, S., Bruno, M.-A., Schnakers, C., et al., 2009. The problem of aphasia in the assessment of consciousness in brain-damaged patients. in: Laureys, S., Schiff, N.D., Owen AMBT-P in BR, (eds.) Coma Science: Clinical and Ethical Implications. Vol 177. Elsevier; 2009:49-61. DOI:10.1016/S0079-6123(09)17705-1.

Maxwell, W.L., Pennington, K., MacKinnon, M.A., et al., 2004. Differential responses in three thalamic nuclei in moderately disabled, severely disabled and vegetative patients after blunt head injury. Brain 127 (11), 2470–2478. https://doi.org/10.1093/brain/awh294.

Mazurowski, Maciej A., Buda, Mateusz, Saha, Ashirbani, Bashir, Mustafa R., 2019. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. J. Magn. Reson. Imaging 49 (4), 939–954. https://doi.org/10.1002/jmri.v49.410.1002/jmri.26534.

McBee, Morgan P., Awan, Omer A., Colucci, Andrew T., Ghobadi, Comeron W., Kadom, Nadja, Kansagra, Akash P., Tridandapani, Srini, Auffermann, William F., 2018. Deep learning in radiology. Acad. Radiol. 25 (11), 1472–1480. https://doi.org/10.1016/j.acra.2018.02.018.

Monti, M.M., Vanhaudenhuyse, A., Coleman, M.R., Boly, M., Pickard, J.D., Tshibanda, L., Owen, A.M., Laureys, S., 2010. Willful modulation of brain activity in disorders of consciousness. N. Engl. J. Med. 362 (7), 579–589. https://doi.org/10.1056/NEJMoa0905370.

Morozova, Sofya, Kremneva, Elena, Sergeev, Dmitry, Sinitsyn, Dmitry, Legostaeva, Lyudmila, Iazeva, Elizaveta, Krotenkova, Marina, Ryabinkina, Yulia, Suponeva, Natalia, Piradov, Michael, 2018. Conventional structural magnetic resonance imaging in differentiating chronic disorders of consciousness. Brain Sci. 8 (8), 144. https://doi.org/10.3390/brainsci8080144.

Owen, A.M., Coleman, M.R., Boly, M., et al., 2006. Detecting Awareness in the Vegetative State. Science (80-). 313(5792):1402 LP - 1402. doi:10.1126/science.1130197.

Parvizi, J., Damasio, A.R., 2003. Neuroanatomical correlates of brainstem coma. Brain. 126 (7), 1524–1536. https://doi.org/10.1093/brain/awg166.

Pignat, J.-M., Mauron, E., Jöhr, J., et al., 2016. Outcome Prediction of Consciousness Disorders in the Acute Stage Based on a Complementary Motor Behavioural Tool. PLoS One. 11(6):e0156882-e0156882. doi:10.1371/journal.pone.0156882.

Pincherle, A., Jöhr, J., Chatelle, C., Pignat, J.-M., Du Pasquier, R., Ryvlin, P., Oddo, M., Diserens, K., 2019. Motor behavior unmasks residual cognition in disorders of consciousness. Ann. Neurol. 85 (3), 443–447. https://doi.org/10.1002/ana.25417.

Pincherle, A., Rossi, F., Jöhr, J., Dunet, V., Ryvlin, P., Oddo, M., Schiff, N., Diserens, K., 2021. Early discrimination of cognitive motor dissociation from disorders of consciousness: pitfalls and clues. J. Neurol. 268 (1), 178–188. https://doi.org/10.1007/s00415-020-10125-w.

Posner, J.B., Saper, C.B., Schiff, N.D., et al., 2007. Plum and Posner's Diagnosis of Stupor and Coma, 4th ed. Oxford University Press, New York.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. 12 (1) https://doi.org/10.1186/1471-2105-12-77.

Rohaut, B., Doyle, K.W., Reynolds, A.S., Igwe, K., Couch, C., Matory, A., Rizvi, B., Roh, D., Velazquez, A., Megjhani, M., Park, S., Agarwal, S., Mauro, C.M., Li, G., Eliseyev, A., Perlbarg, V., Connolly, S., Brickman, A.M., Claassen, J., 2019. Deep structural brain lesions associated with consciousness impairment early after hemorrhagic stroke. Sci. Rep. 9 (1) https://doi.org/10.1038/s41598-019-41042-2.

Sandroni, Claudio, D'Arrigo, Sonia, Cacciola, Sofia, Hoedemaekers, Cornelia W.E., Kamps, Marlijn J.A., Oddo, Mauro, Taccone, Fabio S., Di Rocco, Arianna, Meijer, Frederick J.A., Westhall, Erik, Antonelli, Massimo, Soar, Jasmeet, Nolan, Jerry P., Cronberg, Tobias, 2020. Prediction of poor neurological outcome in comatose survivors of cardiac arrest: a systematic review. Intensive Care Med. 46 (10), 1803–1851. https://doi.org/10.1007/s00134-020-06198-w.

Schiff, N.D., 2015. Cognitive motor dissociation following severe brain injuries. JAMA Neurol. 72 (12), 1413–1415. https://doi.org/10.1001/jamaneurol.2015.2899.

Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., Moonen, G., Laureys, S., 2009. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. BMC Neurol. 9 (1) https://doi.org/10.1186/1471-2377-9-35.

Seel, R.T., Sherer, M., Whyte, J., Katz, D.I., Giacino, J.T., Rosenbaum, A.M., Hammond, F.M., Kalmar, K., Pape, T.-B., Zafonte, R., Biester, R.C., Kaelin, D., Kean, J., Zasler, N., 2010. Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. Arch. Phys. Med. Rehabil. 91 (12), 1795–1813. https://doi.org/10.1016/j.apmr.2010.07.218.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6 (1), 60. https://doi.org/10.1186/s40537-019-0197-0.

Stender, J., Gosseries, O., Bruno, M.-A., Charland-Verville, V., Vanhaudenhuyse, A., Demertzi, A., Chatelle, C., Thonnard, M., Thibaut, A., Heine, L., Soddu, A., Boly, M., Schnakers, C., Gjedde, A., Laureys, S., 2014. Diagnostic precision of PET imaging and

functional MRI in disorders of consciousness: a clinical validation study. Lancet 384 (9942), 514–522. https://doi.org/10.1016/S0140-6736(14)60042-8.

Stretti, Federica, Klinzing, Stephanie, Ehlers, Ulrike, Steiger, Peter, Schuepbach, Reto, Krones, Tanja, Brandi, Giovanna, 2018. Low Level of vegetative state after traumatic brain injury in a Swiss academic hospital. Anesth. Analg. 127 (3), 698–703. https://doi.org/10.1213/ANE.0000000000003375.

Unnikrishnan, P., Kumar, D.K., Poosapadi Arjunan, S., Kumar, H., Mitchell, P., Kawasaki, R., 2016. Development of health parameter model for risk prediction of CVD Using SVM. In: Arbeev, K.G. (Ed.), Comput Math Methods Med., pp. 1–7. https://doi.org/10.1155/2016/3016245.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L.J.-F., et al., 2009. Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. Brain 133(1):161-171. doi:10.1093/brain/awp313.

Jouanne-Diedrich, H. von. OneR: One Rule Machine Learning Classification Algorithm with Enhancements. 2017.

Weiss, G., McCarthy, K., Zabar, B., 2007. Cost-Sensitive Learning vs. Sampling: which is best for handling unbalanced classes with unequal error costs? DMIN.

Wongpakaran, N., Wongpakaran, T., Wedding, D., et al., 2013. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med. Res. Method. 13 (1), 61. https://doi.org/10.1186/1471-2288-13-61.

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3 (1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.