Merging clinical and EEG biomarkers in an Elastic-Net regression for disorder of consciousness prognosis prediction

Piergiuseppe Liuzzi, Antonello Grippo, Silvia Campagnini, Maenia Scarpino, Francesca Draghi, Annamaria Romoli, Hakiki Bahia*, Raisa Sterpu, Antonio Maiorelli, Claudio Macchi, Francesca Cecchi, Maria Chiara Carrozza, *Senior Member IEEE* and Andrea Mannini

***Abstract*—Patients with Disorder of Consciousness (DoC) entering Intensive Rehabilitation Units after a severe Acquired Brain Injury have a highly variable evolution of the state of consciousness which is a complex aspect to predict. Besides clinical factors, electroencephalography has clearly shown its potential into the identification of prognostic biomarkers of consciousness recovery. In this retrospective study, with a dataset of 271 patients with DoC, we proposed three different Elastic-Net regressors trained on different datasets to predict the Coma Recovery Scale-Revised value at discharge based on data collected at admission. One dataset was completely EEG-based, one solely clinical data-based and the last was composed by the union of the two. Each model was optimized, validated and tested with a robust nested cross-validation pipeline. The best models resulted in a median absolute test error of 4.54 [IQR = 4.56], 3.39 [IQR = 4.36], 3.16 [IQR = 4.13] for respectively the EEG, clinical and hybrid model. Furthermore, the hybrid model for what concerns overcoming an unresponsive wakefulness state and exiting a DoC results in an AUC of 0.91 and 0.88 respectively. Small but useful improvements are added by the EEG dataset to the clinical model for what concerns overcoming an unresponsive wakefulness state. Data-driven techniques and namely, machine learning models are hereby shown to be capable of supporting the complex decision-making process the practitioners must face.**

***Index Terms*— Electroencephalography, Disorder of Consciousness, Machine Learning, Coma Recovery Scale – Revised, Prognostic Models**

I. Introduction

Severe Acquired Brain Injuries (sABIs) are defined as traumatic, post-anoxic, vascular or other brain damages that cause coma for at least 24 hours. These patients, after the coma phase, can transit to a state of prolonged Disorder of Consciousness (DoC) which includes the Unresponsive Waking State (UWS, previously referred as vegetative state) and the minimal conscious state (MCS). This condition may persist or cease achieving a full recovery emerging from the MCS (E-MCS) [1].

The mechanisms underlying recovery from a DoC are currently unknown and are strongly dependent on etiology, age and injury severity [2]. Furthermore, the patients' rehabilitative paths have a high inter-individual variability. For this reason, prognostication on neurological outcomes in patients with DoC remains a challenging task. A conspicuous number of predictive parameters in rehabilitation of patients with DoC have been reported both from the clinical world [3]–[5] and the instrumental world [6]–[8]. Multiple electroencephalography (EEG) descriptors as reactivity, alpha waves and an antero-posterior gradient (APG) have been shown to be predictive of a possible consciousness recovery [9]–[12]. Many are also brain-computer interfaces application with embedded EEG-based prognostic models [13]–[15]. However, researchers reported some limitations in using the EEG as a prognostic instrument for DoC given a substantial lack of standardized terminology. Recently, the American Clinical Neurophysiology Society (ACNS) attempted a standardization in nomenclature and assessment techniques of EEG biomarkers in patients in critical care. The latter has proven to be a valid starting point in the systematic evaluation of an EEG signal in critical patients [16]–[19].

At the same time, the latest international guidelines for diagnosis of patients with DoC have recommended the use of both clinical and instrumental evaluations to minimize the risk of misdiagnosis [20], [21]. In the context, EEG assessments must show its potential in improving prognosis accuracy and precision.

Data-driven approaches have been proved to be effectively helping practitioners in the clinical decision-making processes as reported in systematic reviews in this field [22]. Namely, when evaluating outcomes in patients with DoC, authors

This article was submitted for review 20th October, 2021. The study was supported by the Italian neuroscience and neurorehabilitation research hospitals network ("Rete IRCCS delle Neuroscienze e della Neuroriabilitazione") which funded the study jointly with the "Ricerca corrente RC2020-RC2021 program" and the 5x1000 funds AF2018: "Data Science in Rehabilitation Medicine" AF2019: "Study and development of biomedical data science and machine learning methods to support the appropriateness and the decision-making process in rehabilitation medicine" by the Italian Ministry of Health.

P. Liuzzi and S. Campagnini are with the BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, PI 56025 IT and with the IRCCS Fondazione Don Carlo Gnocchi, Firenze, FI 50143 IT (pliuzzi@dongnocchi.it, scampagnini@dongnocchi.it)

M.C. Carrozza is with the BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, PI, 56025, IT (chiara.carrozza@santanna.it)

A. Mannini, A. Grippo, M. Scarpino, B. Hakiki, F. Draghi, A. Romoli, R. Sterpu, A. Maiorelli, C. Macchi and F. Cecchi are with the IRCCS Fondazione Don Carlo Gnocchi, Firenze, FI 50143 IT (amannini@dongnocchi.it, agrippo@dongnocchi.it, bhakiki@dongnocchi.it, fdraghi@dongnocchi.it, rsterpu@dongnocchi.it, cmacchi@dongnocchi.it, fcecchi@dongnocchi.it). F. Cecchi and C. Macchi are also with the department of Exp. and Clin. Medicine of the University of Florence.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

targeted specific rehabilitation milestones as visual pursuit [23], command following [24] or decannulation [25]. For what concerns predicting consciousness recovery different multi-modal based machine learning models have implemented, although often lacking of rigorous cross-validation or suffering from low sample size [6], [26].

For this reason and to estimate the extent to which a qualitative, but standardized, inspection of the EEG signal can improve the prognosis of consciousness changes in DoC, we proposed three different Machine Learning (ML) models internally cross-validated and tested using data from 271 sABI patients entering the Intensive Rehabilitation Unit (IRU) with a DoC.

First, we targeted the estimation of the Coma Recovery Scale revised (CRS-R) total score [27] at discharge via an Elastic Net regressor with three different input datasets (one based only on EEG, one based only on clinical evaluation, and one based on the union of the two, namely "hybrid"). Secondly, we evaluated classification accuracies of overcoming boundary values in the CRS-R at discharge, most likely indicating a significant change of consciousness state.

II. Methods

A. Study design and participants

A retrospective observational study was performed including 271 patients who were admitted to IRCCS Fondazione Don Carlo Gnocchi from August 1, 2012 to January 31, 2019. Inclusion criteria were diagnosis of DoC after a sABI, adults (age > 18). Approval from the local Ethical Committee was obtained (N. R17505) and enrollment was done following the Helsinki Declaration. Patients have been included after obtaining a written consent signed by a legal guardian (ethical committee waived the necessity of a consent for retrospective observational studies if unable to contact or reach the patient due to negligible risk).

B. Data collection

1) Clinical data: Data concerning demographical (age, gender), clinical and functional aspects were recorded. Functional evaluations were performed by skilled operators (neurologist, neuropsychologists and speech therapists) at the IRUs admission. Based on the repetitive assessment of at least 3 consecutive CRS-R administrations (in three consecutive days) a clinical diagnosis of consciousness was formulated (UWS, MCS or E- MCS) both at admission and at discharge. The latter was based on the CRS-R subscales following international guidelines *[28]*. The repeated CRS-R administration in three consecutive days, is known to notably reduce the possibility of misdiagnosis *[29]*. Besides, the following clinical scales were added to the dataset: Level of Cognitive Function (LCF, *[30]*), Glasgow Coma Scale (GCS, *[31]*), Food Oral Intake Scale (FOIS, *[32]*) and Functional Independence Measure (FIM, *[33]*). The time between the event and the admission to the IRU was also recorded (time post-onset). Lastly, epileptic seizures during the acute phase have been recorded. All features retained for the subsequent ML model were collected within one week from admission.

2) EEG Recordings: Standard 30-min EEG recordings were performed using a digital machine (Gal NT, EBNeuro) and an EEG prewired head cap, with 19 electrodes (Fp1-Fp2-F7-F8-F3-F4-C3-C4-T3-T4-P3-P4-T5-T6-O1-O2-Fz-Cz-Pz) set according to the 10-20 International Standard System *[34]* adopting previously proposed EEG recording parameters [11]. In particular, recordings were acquired with a sampling rate of 128 Hz and filtered with a low-pass filter (cut-off frequency in the 30-70 Hz), a high pass filter (with time constant 0.1-0.3s) adjusted according to interpretation needs (standard gain set to 7V/mm, sensitivity gain 2-10V/mm) as in Scarpino et al. *[11]*. EEG labeling was performed by the agreement of two expert neurologists according to the ACNS terminology [35]. The descriptors included were frequency bands, presence of an anterior/posterior gradient (APG) in the background activity, reactivity, variability (spontaneous), detectable sleep spindles (stage II) and lastly, epileptic discharges.

Furthermore, we added two predictive scores derived from the ACNS labeling. The first, by Estraneo et al. classifies EEG background activity in five groups: normal, mildly and moderately abnormal, diffuse slowing (symmetric or not symmetric diffuse theta/delta rhythm, $\geq$ 20 μV, with no APG) and low voltage (< 20 μV, theta/delta in most brain regions) [36]. The second is a score by Bagnato et al. [37], going from 3 to 7, composed by assigning a score of 1, 2 or 3 to delta, theta

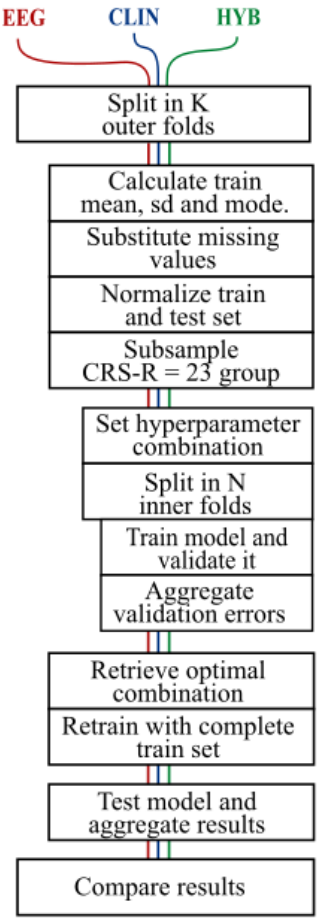


Fig. 1. Nested cross-validation approach used embedding hyperparameters optimization. A subsampling is performed in each outer training set, reducing the instances having discharge CRS-R = 23 to 1/23 and consequently balancing the regression problem. Each indentation corresponds to a for loop in the code.

and alpha frequency plus a score of 1 or 2 for both present/not present reactivity and reduced/normal voltage respectively.

3) Outcome assessment: the CRS-R revised scale was assessed at discharge in absence of drug sedation by skilled operators. The thresholds used to evaluate a meaningful outcome in the discharge consciousness state were set to the maximum value of the CRS-R at discharge in respectively the UWS discharge group ($CRS - R_{dis} = 16$) and the MCS discharge group ($CRS - R_{dis} = 23$). These thresholds were obtained after a sensitivity-specificity analysis on the actual CRS-R discharge values within the discharge clinical state. Results for this preliminary analysis are shown in Appendix A confirming that from $CRS_{dis} = 16$ upwards no UWS is found and that patients reaching $CRS_{dis} = 23$ are certainly E-MCS.

C. Model implementation

1) Statistical analysis: The features used for the CRS-R prediction were first analyzed with univariate statistical analyses on SPSS (Vs 26, Chicago, SPSS Inc.). After testing for normality with the Shapiro-Wilk test, Spearman correlations were computed to verify associations between continuous independent variables and the CRS-R at discharge, whereas Kruskall-Wallis (KW) test was applied for categorical independent variables. Conditioned to KW significance, Dunn-Bonferroni post-hoc tests assessed differences between groups. To compare prediction accuracies across the three different ML models, a Friedman test was conducted between the test absolute errors followed by Dunn-Bonferroni post-hoc tests. In all analyses, a p-value <0.05 was considered statistically significant. Furthermore, to evaluate improvements of the models with respect to the chance level, we defined a Median Guess estimator (MG) which constantly predicts the median of the CRS-R values at discharge. Then, the MG absolute errors are compared via Wilcoxon-Signed Rank Test with the developed ML models' absolute errors.

2) Elastic-Net regression: In order to avoid any sort of train-test contamination and minimize the risk of bias in the performances, all features, independently from univariate statistics analyses results, were retained for the multivariate prediction model. The Elastic-Net (EN) regression model is a regularized method which linearly combines the penalties of the LASSO and Ridge regression *[38]* overcoming respective implementation problems. Ridge adds to linear regression models quadratic regularization via L_2 penalties. On the other hand, Ridge always assigns a non-zero coefficient to all features in the model, consequently failing in eliminating coefficients even if the corresponding independent variable is irrelevant to the prediction.

Conversely, LASSO regression is known to suffer when the dimensionality of the dataset is higher than the number of the available examples or when multicollinear independent variables are present [39], [40], but foster the neglection of specific features.

Hence, Elastic-Net combines feature elimination of LASSO and coefficient reduction from Ridge improving on either LASSO or Ridge modifying the regression parameter estimates as follows:

$$\hat{\beta} = argmin_{\beta}\{||y - X\beta||^2 + \lambda_2||\beta||^2 + \lambda_1||\beta||_1$$

TABLE I
DESCRIPTIVE STATISTICS AND PREDICTORS UNIVARIATE ANALYSIS

	Median [IQR], (% of occurrences)	Test Statistics	p-value
Age	55 [65]	-0.054	0.382
Gender	Female: 33.1, Male: 66.9	1.088(1)	0.297
Clinical parameters			
Etiology	Traumatic: 30.9, Anoxic: 20.8, Ischemic: 17.8, Hemorrhagic 26.4, Other: 4.1	27.818(4)	**< 0.001**
TPO	45 [38]	-0.152	**0.015**
CRS-R adm.	11 [16]	0.651	**< 0.001**
Consciousness state at adm.	UWS: 30.4, MCS: 44.1, EMCS: 25.6	112.498(2)	**< 0.001**
GCS adm.	9 [4]	0.533	**< 0.001**
FOIS adm.	1 [0]	0.237	**< 0.001**
LCF adm.	3 [1]	0.577	**< 0.001**
FIM adm.	18 [10]	0.427	**< 0.001**
EEG parameters			
Seizures in the acute phase	Present: 19, Absent: 81	2.409(1)	0.121
AP gradient	Present: 73.2, Absent: 27.8	51.822(1)	**< 0.001**
Frequency	Delta: 2.6, Theta: 74.0, Alpha: 23.4	21.735(2)	**< 0.001**
Reactivity	Absent: 27.2, Unclear: 18.3, Present: 54.5		
Epileptic activity	No: 74.9, Rare: 19.1, Abundant 1.7, Frequent 4.3	15.409(3)	**0.001**
Sleep spindles	Absent: 85.5, Present & abnormal: 9.8, Present & normal: 4.7	5.409(2)	0.067
Symmetry	Symmetric: 60, Moderately Asymm.: 18.7, Sever. Asymm.: 21.3	0.876(2)	0.645
Variability	Absent: 21.4, Unclear: 9.8, Present: 68.8	77.484(3)	**< 0.001**
Voltage	Suppressed: 11.5, Normal: 86.8, Low voltage: 1.7	32.177(2)	**< 0.001**
Estraneo's score	Norm.: 4.3, Mildly abn.: 47.9, Moderately. abn. 18.4, Diffuse Slowing: 17.9, Suppr. 11.5	65.947(4)	**< 0.001**
Bagnato's score	3: 1.7, 4: 10.3, 5: 33.8, 6: 33.8, 7: 20.5	71.951(4)	**< 0.001**

For continuous variables median and IQR were presented in brackets whilst for categorical independent variables the percentage of occurrence of each label is indicated. The column test statistics indicates the R^2 value of spearman correlations for continuous independent variables and the χ value (degrees of freedom) of the KW test between the variables and the CRS-R continuous value.

TPO: time-post onset; CRS-R: Coma Recovery Scale-Revised, GCS: Glasgow Coma Scale; FOIS: Functional Oral Intake Scale; LCF: Level of Cognitive Functioning; FIM: Functional Independence Measure; UWS: Unresponsive Wakefulness State; MCS: Minimally Conscious State; EMCS: Emergence from MCS.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

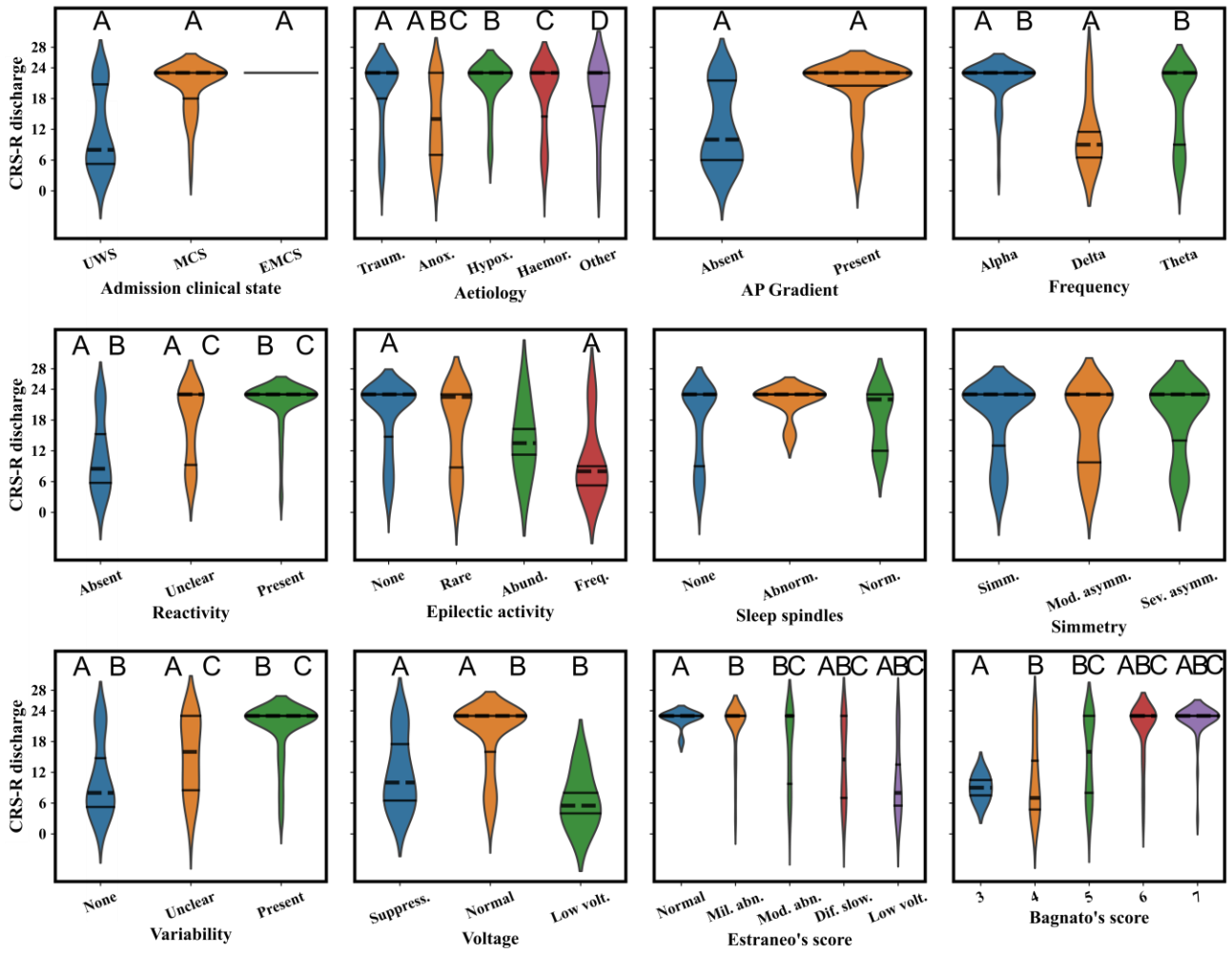


Fig. 2. Violin plots of the CRS-R distribution among groups of categorical variables. Thick continuous black lines indicate CRS-R group median, whilst thinner dotted black lines are the first and third quartile levels. Capital letters above the box indicate significant differences in the post-hoc tests of the Kruskal-Wallis test (Dunn-Bonferroni with Bonferroni correction, adjusted p-value considered significant for $p_{adj} < 0.05$). Specifically, groups sharing the same letter have significantly different discharge CRS-R.

with the special cases $\lambda_2 = 0$, $\lambda_1 \neq 0$ and $\lambda_1 = 0, \lambda_2 \neq 0$ corresponding to the LASSO and Ridge regression respectively, therefore including in the EN model hypothesis space both LASSO and Ridge. This reflects in the *Sklearn* implementation of the parameters estimate equation being defined as:

$$\hat{\beta} = argmin_{\beta}\{\frac{1}{(2 \cdot N_{samples})}||y - X\beta||^2 + \alpha \cdot l1_{ratio}||\beta||_1 + \frac{1}{2}\alpha \cdot (1 - l1_{ratio})||\beta||_2^2\}$$

where $l1_{ratio}$ describes the tendency toward a LASSO regularization ($l1_{ratio} \sim 1$) or the Ridge regularization ($l1_{ratio} \sim 0$).

3) Training, cross-validation, optimization and testing:

The algorithms implementation was carried out separately and individually for each of the three models (EEG, CLIN and HYB) using Python custom code, the *Scikit-Learn* and the *Optuna* [41], [42] libraries. All multiclass categorical features were first converted into dummy variables (one-hot encoded).

A nested-cross validation approach was implemented [23]. In brief, such approach consists in two k-fold cross-validation loops: an outer loop identifies the test set for each of its folds while the inner loop implements the further split of the dataset for training and validation.

The outer loop was designed as a k-fold cross-validation (k = 7), obtaining 7 combinations of train and test set (Fig. 1). Four of them were composed by 231 patients and three by 232 patients. Missing values in each of train and test set were substituted with the mean/mode of the respective train set. Then, the train and test sets were normalized by subtracting the train set mean and dividing by its standard deviation.

Due to the discharge CRS-R ceiling effect, we limited instances in the training sets with $CRS\text{-}R_{dis} = 23$, to an overall frequency of 1/23 with respect to the full dataset.

In each fold of the outer loop, a hyperparameter optimization was performed. The optimized parameters were α and the $l1_{ratio}$ in the ranges [10^{-4} – 10] and [10^{-4} – 1] respectively. The optimization was obtained by an iterative pruning algorithm based on successive halving of the hyperparameter value within the prefixed range for a predefined number of trials. In each trial, hence for each evaluated hyperparameter combination, data were further split in the actual train and validation sets

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

implementing the inner *n*-fold cross-validation (n = 5) of the proposed nested approach. (Fig. 1).

Validation set predictions from the inner 5 folds were computed and aggregated. From here, the validation Median Absolute Error (MAE) of the n models included in the k^{th} outer split was computed according to the following:

$$MAE = \frac{\sum_{i=1}^{M_k} |CRS\_R_{pred} - CRS\_R_{actual}|}{M_k}$$

with M_k being the number of patients in the k^{th} outer split train set before the inner split. The hyperparameter combination minimizing the validation MAE was then chosen for training the final k^{th} model with all M_k samples. This process was repeated for the K outer folds and all models were tested with the respective outer test set. Again, test results were aggregated, and the overall dataset MAE was calculated. Furthermore, R^2 was also included among test evaluation metrics.

4) Regression post-processing

As a final processing step, classification accuracies were computed by discretizing the CRS-R value at discharge. Two thresholds $CRS\text{-}R_{dis} = 16$ and $CRS\text{-}R_{dis} = 23$ were considered, resulting in two different classification problems.

The performance of the three proposed regressions models in both problems was verified by means of ROC curves. The Area under the Curve (AUC) was finally computed using the Simpson integration rule.

III. Results

A. Univariate analysis

Age and gender did not appear to be significant predictors of discharge CRS-R, while etiology ($p < 0.001$, Table 1) resulted to influence the outcome in a negative manner if anoxic (significantly different from all other etiologies in the post-hoc analysis) and in a positive manner if traumatic (Fig. 2). Higher values on all the examined scales at admission were found to be predictive of a more prompt recovery on the CRS-R scale ($p < 0.001$), with the most correlated being the CRS-R ($R^2 = 0.651$) and the GCS ($R^2 = 0.533$). Furthermore, a strong correlation between consciousness state at admission and CRS-R at discharge ($p < 0.001$) was found from the KW test. The post-hoc test with a $p < 0.001$ indicated that patients in EMCS state at admission have a significantly higher discharge total CRS-R score than the MCS and UWS. Same holds for the MCS patients compared to the UWS ones ($p < 0.001$).

For what concerns the EEG variables of the ACNS classification, we confirmed that a theta and delta background frequency lead to a smaller CRS-R total score at discharge with respect to an alpha background. Same holds for the absence of reactivity, with respect to both its clear ($p < 0.001$) and unclear ($p < 0.001$) presence. Furthermore, significant improvements were found in patients with a clear reactivity compared to the ones with an unclear reactivity response ($p < 0.001$).

Patients with unclear ($p < 0.05$) or clear ($p < 0.001$) variability reached a higher CRS-R than patients with no spontaneous variability. The presence of a clear variability w.r.t. an unclear one is associated with a higher outcome value.

A clear inverse relationship between the discharge CRS-R and the amount of epileptic activity recorded was observed (Fig. 2), where a worse outcome is expected as the amount of epileptic activity increases. On the other hand, post-hoc test showed a significance only between the pair no epileptic activity – frequent epileptic activity ($p < 0.001$). Sleep spindles and symmetrical brain organization was found to be uncorrelated with CRS-R at discharge.

A strong correlation was found between voltage and discharge CRS-R ($p < 0.001$). Namely, a suppressed ($p < 0.001$) and a low voltage ($p < 0.01$) resulted to be predictors of a lower CRS-R (Fig. 2). Additionally, the presence of an APG was found to be indicative of better consciousness recovery ($p < 0.001$).

For what concerns derived scores, both Estraneo's and Bagnato's score resulted in strongly significant associations with the CRS-R in the KW test ($p<0.001$). Specifically, for the Estraneo's score, a background activity with low voltage or with a diffuse slowing pattern is a predictor of significantly lower discharge CRS-R if compared to normal and mildly abnormal EEG activity. Furthermore, patients with low voltage EEG were also found to be associated with a worse outcome than patients with moderately abnormal background activity. Conversely, patients with Bagnato's score of 6 or 7 were significantly worse in all the possible pairings and no difference was found between the groups with a score 3, 4 and 5.

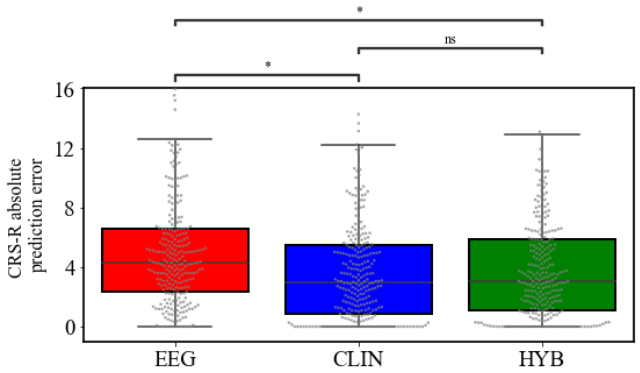


Fig. 3. Box-plot of absolute errors of agglomerated outer testing results for the three models. Superimposed on the boxes, the individual samples error is plotted showing the underlying error distribution of the predictions. Significative improvements ($p < 0.05$) were found between the EEG and both the CLIN and the HYB models. R^2 between actual and predicted values resulted equal to 0.49, 0.71 and 0.73 for respectively the EEG, CLIN and HYB models.

B. CRS-R regression models

All features from Table 1 were included in the model after conversion of categorical variables in dummy variables resulting in 43 features (Fig. 4). In each of the k^{th} Elastic-Net models optimization paths, 250 trials resulted to be sufficient to converge to a constant optimum. Specifically, the α parameter resulted in a median of 0.018, 0.082 and 0.049 respectively for the EEG, CLIN and HYB models.

On the other hand, the $l1_{ratio}$, showed that the use of the clinical features strongly shifts the amount of regularization toward the LASSO approach. In particular, the CLIN and the HYB have respectively a median $l1_{ratio}$ of 0.952 and of 0.539, whilst EEG median $l1_{ratio}$ resulted in 0.069 points.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

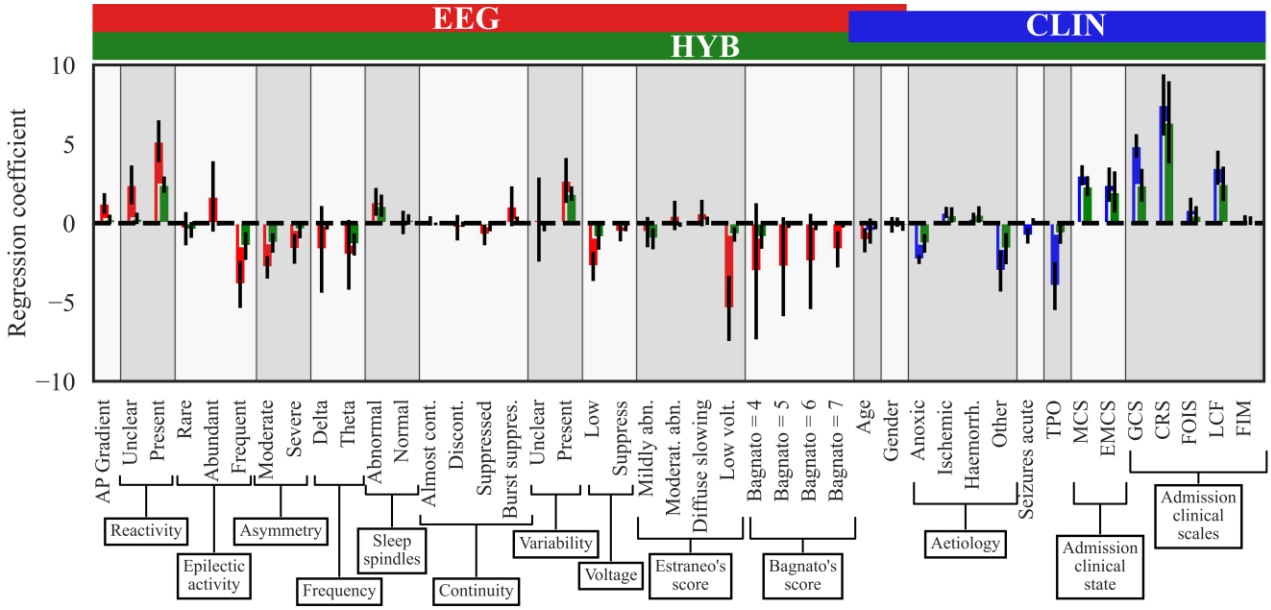


Fig. 4. Elastic-Net regression coefficients for the EEG (red), the CLIN (blue) and the HYB (green) models. The height of each colored bar is the average value of the regression coefficients of the models trained in the outer folds. The black bar indicates the standard deviation of the regression coefficients of the outer folds.

For the validation set, the median between the inner folds MAE values resulted equal to 4.33 points [IQR = 0.66], 3.49 [IQR = 0.67] and 3.18 [IQR = 0.64].

The optimal solutions test MAE resulted equal to 4.6 points [IQR = 11.6] for the EEG, to 3.5 points [IQR = 10.2] for the CLIN and to 3.3 points [IQR = 11.9] for the HYB model. Models absolute errors resulted significantly different in the Friedman Test ($p < 0.01$) with pair-wise Dunn-Bonferroni post-hoc comparisons showed a significant improvement in prediction of both the CLIN ($p < 0.05$) and the HYB ($p < 0.05$) models with respect to the EEG. The MG estimator resulted in a MAE of 6.05 points [IQR = 1.21]. All the developed models showed a significant improvement with respect to the MG estimator ($p<0.001$). Test statistics resulted equal to W = 13.211, W = 7.624 and W = 8.321 for respectively the EEG, CLIN and HYB estimator.

In all three multivariate models, a negative influence on the CRS-R score at discharge was found for an anoxic etiology and an older age whilst an admission state of MCS or E-MCS contributed to an improved outcome with respect to UWS (Fig 4.) The coefficients β of the EEG model showed how the presence of spontaneous variability, reactivity and APG are predictive of a better CRS-R at discharge, coherently with findings from univariate analysis. Also, for frequent epileptic activity and a low background voltage, a strong negative regression coefficient was found. Conversely, in the multivariate EEG, a moderate or severe asymmetry seems associated with a worst CRS-R.

Classes (4-7) of Bagnato's score may be related to a worst outcome (negative regression coefficient), with a lowering of the negative effect with an increasing value on the scale..

Except for the FOIS and FIM, all clinical scales in the CLIN and in the HYB models resulted in a positive regression coefficient, with the CRS-R having the highest importance in both models.

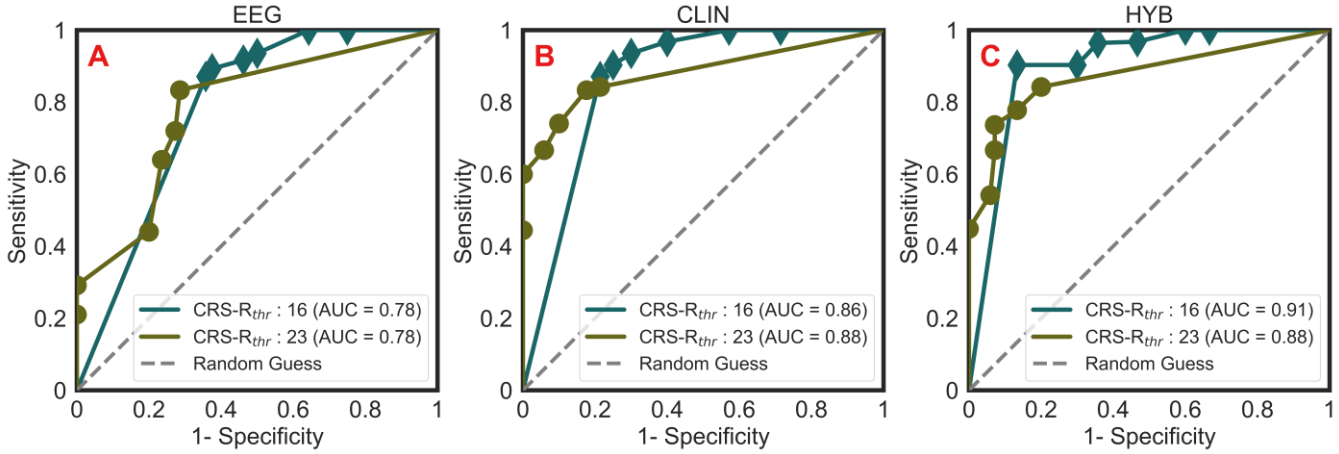


Fig. 5. Receiver-Operating Curve (ROC) of the three models with outcome: *overcoming the CRS-R threshold* (equal or bigger than 16 or 23). Respectively, the EEG, CLIN and HYB model are represented in panel A,B,C.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Generally, even if with a wider distribution of cumulative weights for the models with higher dimensionality, the independent variables relative contributions to the predictions resulted consistent across the three models (Fig. 4).

C. Regression post-processing

The binary classification with respect to the CRS-R_{thr} = 16 threshold showed high sensitivity for all three models, with an AUC of 0.78, 0.86, and 0.91 for respectively the EEG, the CLIN and the HYB model. For what concerns the best model in predicting the emergence from a DoC (CRS-R_{thr} = 23), the HYB algorithm showed an AUC of 0.88, with no improvements on the clinical based classification. For the latter, maximizing sensitivity keeping specificity within an acceptable limit yields a point in the ROC curve with sensitivity 0.85 and specificity 0.78.

IV. Discussions

In literature, many findings assert that prognosticating neurological outcomes in sABI patients is a complex task. Its improvement would facilitate the communication process with relatives and allow for a precise individualized rehabilitation management. The accuracy of the consciousness assessment may be compromised by a number of confounding factors (presence of consciousness fluctuations, paralysis, aphasia...) [44].

For this reason, we should believe that the underlying level of consciousness must be assessed via a multifactorial approach, both evaluating the clinical data and interpreting the internal physiological patterns as recommended by the latest international guidelines. Such multifactorial techniques allow to convey two different sources of information and analyze the interaction between them. Authors analyzed and combined EEG descriptors in predictive scores for consciousness improvement prediction [10], [36], [37], [45]. We extend their previous work in three main directions. Firstly, we use prognostic factors and in general patient characteristics to predict punctually the value of the discharge CRS-R. Secondly, we cross-validate and test our model with a robust technique (nested cross-validation) allowing for inference on a larger population. Lastly, we explicitly provide to the trained models different data sources (clinical only, neurophysiological only and hybrid) and evaluate how an appropriate feature combination can result in an accurate prediction.

Reasonably, a prospective validation of the models is required to confirm the validity of obtained results, even if the nested cross-validation approach simulates a prospective assessment by separating patients used to optimize hyperparameters and patients used to test the algorithms. This point, jointly with the low variability on inner folds validation errors, suggests that the model is robust and generalizes well on new data.

Previous studies, already evaluated whether a model with different sources can better identify prognostic factors for the recovery of consciousness [6]. Specifically, Yu et al. classifies with a k-fold cross-validated binary Support Vector Machine trained on 51 patients the presence of consciousness. The latter was diagnosed with the Glasgow Outcome Scale-Expanded (GOS-E) score, only suitable for a dichotomous subdivision of consciousness levels in persistent vegetative state (GOS-E <=2) and conscious (GOS-E>2). Such classification does not allow for the fundamental distinction which must be made between minimally consciousness state and unresponsive waking state [1] and it is not a recommended evaluation tool for DoC diagnosis [46]. Furthermore, the model proposed by Yu et al. suggests that fMRI data and laboratory parameters can be combined successfully reaching an accuracy of 73% (with a non-stratified cohort composed by 34 conscious patients and 17 unconscious), [43]. Song et al. predicts both the CRS-R score and a dichotomized version of the GOS-E at one year from the fMRI and the clinical data extraction with an innerly cross-validated and tested SVM (160 patients) [47]. Their CRS-R test prediction root mean squared error is declared equal to 5.07 points with an R^2 = 0.35. Despite the significant difference in the source of instrumental data, our HYB model achieved a more promising validation median absolute error of 3.185 [IQR = 0.642] and an R^2 = 0.79. Overall, we confirmed previous findings and our hypotheses concerning the positive influence of admission clinical scales as the CRS-R, GCS, LCF, FOIS and FIM onto discharge CRS-R. On the other hand, FOIS and FIM were the only two scales which showed a considerably lower relevance in predicting discharge CRS-R total score in ML models (Fig. 4). Reasonably, this may be justified from the fact that most patients of our dataset were admitted to the IRU with similar FOIS and FIM scale values. Specifically, due to the impossibility from patients with a DoC of intaking food via oral means, the majority of FOIS values stands around 1. Similarly, for motor independence, the same flooring effect was found (~18). Still, due to the non-zero standard deviation of these features we included them in the model.

In our work, previous results were extended by accurately predicting a three-level consciousness stratification on a large enough dataset to allow the nested cross-validation of methods. Achieved results enable us to envision a decision support tool for its clinical use. In this regard, the use of the EEG signal instead of imaging techniques (fMRI) fosters a quicker and less expensive use of the algorithm and reduces the required steps into a Point-of-Care Test. Given the easy montage and the less expensive characteristics, EEG examinations are already performed daily in hospitals, making their cost routinely implemented within budgets.

In this optic, despite EEG data did not significantly improve the prediction obtained by the clinical data, our results show how EEG data only (with the EEG model) can already offer a support to the neurological prognosis of DoC patients with an AUC of 0.78 for both the considered classification problems (Fig. 5). Lastly, we showed how the interactions between neurophysiological patterns and clinical evidence can merge in the hybrid model providing an increase in prediction accuracy based on the CRS-R = 16 threshold. Such increase in accuracy can be attributed to the type of information that EEG provides, thus its expression of a reorganization of a cortical network modulated by thalamo-cortical afferents, necessary condition for the presence of intentional (non-reflexive) behavioral responses. In this optic, such cortical behavior (alpha background frequency and cortical reactivity) result fundamentally more important in transiting from a UWS to MCS state. Furthermore, the evolution of the CRS-R scores is also conditioned by the etiology of sABI and by the lesion topography that may involve selective cognitive deficits

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2022.3178801, IEEE Transactions on Neural Systems and Rehabilitation Engineering

(aphasia, frontal syndromes) that may "mask" the clinical evolution and that cannot be identified by the EEG with consequent falsely optimistic predictions. Henceforth, we speculate that to fruitfully exploit interactions between EEG patterns and clinical variables, models targeting etiology-specific cohorts would maximize the information gain. Lastly, including EEG biomarkers as event-related potentials, somatosensory evoked- and visual evoked- potentials as well as quantitative EEG measures (e.g., functional, source connectivity, etc…) may provide additional prognostic markers to increase significantly the prediction accuracy.

Regarding the outcome selection, we acknowledge that the cumulative CRS-R score, may suffer from missing points in motor sub-domains for some patients [44], hence it can underestimate the actual consciousness level. Furthermore, it is reported how finding precise cut-offs separating consciousness levels is difficult [48]. Still, the CRS-R scale, allows for a good consciousness stratification, remaining the actual gold standard [49]. Hence, to evaluate in the most precise way the consciousness level, consciousness states as UWS, MCS and E-MCS have to be considered. For these reasons, we propose here a classification of patients by a variable threshold on CRS-R, reporting specific results for two specific thresholds. Such binarized classification is performed using boundaries (16 and 23) that in our case resulted from the observation of available data to achieve 100% specificity in detecting UWS patients and 100% sensitivity concerning E-MCS detection. However, given that the model estimates the continuous value of CRS-R, different values in terms of sensitivity-specificity couple for the identification of the consciousness state can be selected in order to fit different clinical requests or hypotheses (Appendix B). Still, the retrospective nature of the data has to be reported within the study limitations. Consciousness assessment at the time of the neurological diagnosis was formulated based on the individual CRS-R subscales. Nevertheless, due to the retrospective nature of the study, it was not possible to retrieve from patients' health records the subscales values since it was not often annotated within the available information. This will be tackled by a prospective study currently ongoing (Clin. Trial. Gov. N. NCT04495192) via the collection of the individual sub-domains of the CRS-R scale, allowing for the development of prediction models targeting individual subscales scores and for a finer definition of the mutual influence between the EEG signal and the clinical status in predicting DoC neurological outcomes. Predicting individual sub-items would investigate how and to what extent clinical functions and brain patterns influence individual consciousness domains. In particular, whether EEG signals can be useful for the prediction of lowest values on each subscale (representative of reflexive activity) or higher values on each subscale (representative of cognitively-driven activity). Additionally, the prospective and new experimental data will allow for a deeper analysis of the role of joint clinical and EEG features for prognostic purposes, as well as a confirmation of the hypotheses raised for the explanation of the results obtained in this study.

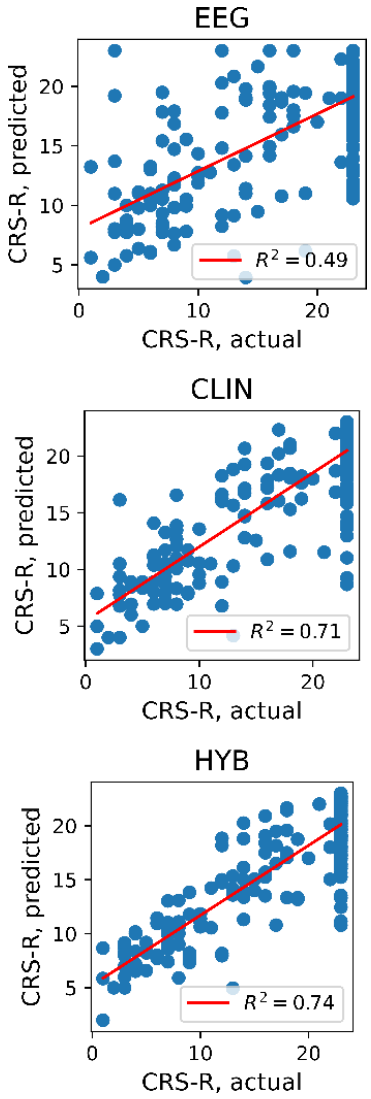
V. Conclusion

The study investigates predictors of consciousness improvements via the use of CRS-R at discharge and ML methods comparing three approaches: with a clinical dataset only, an EEG-based dataset only and with a combination of the two. Setting threshold on predicted CRS-R, overcoming an unresponsive state is successfully predicted with an AUC of 0.91 and exiting from a DoC is assessed with an AUC of 0.88. EEG resulted to be slightly improving the prediction of overcoming an UWS, whilst no improvements are seen in differentiating patients emerging from a MCS. Results will be confirmed with a prospective validation and compared to the 'skilled guess' of a pool of experienced doctors in future works to come.

Our findings confirms that ML algorithms, which already proven to improve decision accuracy in many fields, may support the neurological prognosis in DoC patients and the communication with the patients' relatives.

Appendix A

Scatter-plot of actual (x-axis) versus predicted (y-axis) CRS-R values for the three models



APPENDIX B

Sensitivity-Specificity couples indicating the overcoming a specific clinical group (UWS, MCS) on our dataset. These are obtained by evaluating sensitivity and specificity between the actual clinical state at discharge and the dichotomized CRS-R at different threshold. Maximal sensitivity for overcoming an unresponsive state is obtained for a CRS-R > 16 whilst maximal sensitivity for exiting a DoC is found for CRS-R > 23 (indicated in bold).

CRS-R	Overcoming UWS		Overcoming MCS	
threshold	Sensitivity	Specificity	Sensitivity	Specificity
1	0,01	1,00	0,00	1,00
2	0,06	1,00	0,02	1,00
3	0,10	1,00	0,04	1,00
4	0,30	1,00	0,12	1,00
5	0,51	1,00	0,21	1,00
6	0,65	1,00	0,27	1,00
7	0,76	0,99	0,33	1,00
8	0,90	0,97	0,40	1,00
9	0,96	0,91	0,48	1,00
10	0,98	0,89	0,51	1,00
11	0,98	0,85	0,55	1,00
12	0,98	0,82	0,57	1,00
13	0,98	0,77	0,62	1,00
14	0,98	0,68	0,70	1,00
15	0,99	0,67	0,78	1,00
16	1,00	0,66	0,79	1,00
17	1,00	0,57	0,81	1,00
18	1,00	0,49	0,88	1,00
19	1,00	0,40	0,97	1,00
20	1,00	0,39	0,98	1,00
21	1,00	0,38	0,99	1,00
22	1,00	0,37	0,99	1,00
23	1,00	0,37	1,00	1,00

REFERENCES

[1] R. Hirschberg e J. T. Giacino, «The Vegetative and Minimally Conscious States: Diagnosis, Prognosis and Treatment», *Neurol. Clin.*, vol. 29, n. 4, pagg. 773–786, nov. 2011, doi: 10.1016/j.ncl.2011.07.009.

[2] J. Whyte *et al.*, «Predictors of short-term outcome in brain-injured patients with disorders of consciousness», in *Progress in Brain Research*, vol. 177, Elsevier, 2009, pagg. 63–72. doi: 10.1016/S0079-6123(09)17706-3.

[3] B. Hakiki *et al.*, «Decannulation and improvement of responsiveness in patients with disorders of consciousness», *Neuropsychol. Rehabil.*, pagg. 1–17, ott. 2020, doi: 10.1080/09602011.2020.1833944.

[4] M. Song *et al.*, «Prognostic models for prolonged disorders of consciousness: an integrative review», *Cell. Mol. Life Sci.*, vol. 77, n. 20, pagg. 3945–3961, ott. 2020, doi: 10.1007/s00018-020-03512-z.

[5] E. Portaccio *et al.*, «Score on Coma Recovery Scale-Revised at admission predicts outcome at discharge in intensive rehabilitation after severe brain injury», *Brain Inj.*, vol. 32, pagg. 1–5, feb. 2018, doi: 10.1080/02699052.2018.1440420.

[6] Y. Yu *et al.*, «A multi-domain prognostic model of disorder of consciousness using resting-state fMRI and laboratory parameters», *Brain Imaging Behav.*, ott. 2020, doi: 10.1007/s11682-020-00390-8.

[7] Q. Liu *et al.*, «Spectrum Analysis of EEG Signals Using CNN to Model Patient's Consciousness Level Based on Anesthesiologists' Experience», *IEEE Access*, vol. 7, pagg. 53731–53742, 2019, doi: 10.1109/ACCESS.2019.2912273.

[8] S. Aziz, S. Ibraheem, A. Malik, F. Aamir, M. U. Khan, e U. Shehzad, «Electrooculugram based Communication System for People with Locked-in-Syndrome», in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Istanbul, Turkey, giu. 2020, pagg. 1–6. doi: 10.1109/ICECCE49384.2020.9179323.

[9] B. Wutzl *et al.*, «Narrative Review: Quantitative EEG in Disorders of Consciousness», *Brain Sci.*, vol. 11, n. 6, pag. 697, mag. 2021, doi: 10.3390/brainsci11060697.

[10] M. Scarpino, F. Lolli, G. Lanzo, e A. Grippo, «What is the role of post acute EEG in prediction of late neurological outcome in severe disorders of consciousness?», *Future Neurol.*, vol. 15, n. 1, pag. FNL36, feb. 2020, doi: 10.2217/fnl-2019-0017.

[11] M. Scarpino *et al.*, «Prognostic value of post-acute EEG in severe disorders of consciousness, using American Clinical Neurophysiology Society terminology», *Neurophysiol. Clin.*, vol. 49, n. 4, pagg. 317–327, set. 2019, doi: 10.1016/j.neucli.2019.07.001.

[12] O. Meiron, J. Barron, J. David, e E. Jaul, «Neural reactivity parameters of awareness predetermine one-year survival in patients with disorders of consciousness», *Brain Inj.*, vol. 35, n. 4, pagg. 453–459, mar. 2021, doi: 10.1080/02699052.2021.1879398.

[13] R. Spataro *et al.*, «Preserved somatosensory discrimination predicts consciousness recovery in unresponsive wakefulness syndrome», *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 129, n. 6, pagg. 1130–1136, giu. 2018, doi: 10.1016/j.clinph.2018.02.131.

[14] R. Xu, R. Spataro, B. Z. Allison, e C. Guger, «Brain-Computer Interfaces in Acute and Subacute Disorders of Consciousness», *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 39, n. 1, pagg. 32–39, gen. 2022, doi: 10.1097/WNP.0000000000000810.

[15] J. Pan *et al.*, «Prognosis for patients with cognitive motor dissociation identified by brain-computer interface», *Brain*, vol. 143, n. 4, pagg. 1177–1189, apr. 2020, doi: 10.1093/brain/awaa026.

[16] L. J. Hirsch *et al.*, «American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology: 2021 Version», *J. Clin. Neurophysiol.*, vol. 38, n. 1, pagg. 1–29, gen. 2021, doi: 10.1097/WNP.0000000000000806.

[17] M. C. Hermans, M. B. Westover, M. J. A. M. van Putten, L. J. Hirsch, e N. Gaspard, «Quantification of EEG reactivity in comatose patients», *Clin. Neurophysiol.*, vol. 127, n. 1, pagg. 571–580, gen. 2016, doi: 10.1016/j.clinph.2015.06.024.

[18] A. F. Struck *et al.*, «Association of an Electroencephalography-Based Risk Score With Seizure Probability in Hospitalized Patients», *JAMA Neurol.*, vol. 74, n. 12, pag. 1419, dic. 2017, doi: 10.1001/jamaneurol.2017.2459.

[19] M. Scarpino *et al.*, «Neurophysiological and neuroradiological test for early poor outcome (Cerebral Performance Categories 3–5) prediction after cardiac arrest: Prospective multicentre prognostication data», *Data Brief*, vol. 27, pag. 104755, dic. 2019, doi: 10.1016/j.dib.2019.104755.
[20] J. T. Giacino *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*, vol. 91, n. 10, pagg. 450–460, set. 2018, doi: 10.1212/WNL.0000000000005926.
[21] D. Kondziella *et al.*, «European Academy of Neurology guideline on the diagnosis of coma and other disorders of consciousness», *Eur. J. Neurol.*, vol. 27, n. 5, pagg. 741–756, mag. 2020, doi: 10.1111/ene.14151.
[22] F. Shamout, T. Zhu, e D. A. Clifton, «Machine Learning for Clinical Outcome Prediction», *IEEE Rev. Biomed. Eng.*, vol. 14, pagg. 116–126, 2021, doi: 10.1109/RBME.2020.3007816.
[23] F. Riganello, M. D. Cortese, G. Dolce, e W. G. Sannita, «Visual pursuit response in the severe disorder of consciousness: modulation by the central autonomic system and a predictive model», *BMC Neurol.*, vol. 13, n. 1, pag. 164, dic. 2013, doi: 10.1186/1471-2377-13-164.
[24] G. Rivera-Lillo, E. A. Stamatakis, T. A. Bekinschtein, D. K. Menon, e S. Chennu, «Delta band activity contributes to the identification of command following in disorder of consciousness», *Sci. Rep.*, vol. 11, n. 1, pag. 16267, dic. 2021, doi: 10.1038/s41598-021-95818-6.
[25] A. Mannini *et al.*, «Data-driven prediction of decannulation probability and timing in patients with severe acquired brain injury», *Comput. Methods Programs Biomed.*, vol. 209, pag. 106345, set. 2021, doi: 10.1016/j.cmpb.2021.106345.
[26] F. Riganello, M. D. Cortese, G. Dolce, L. F. Lucca, e W. G. Sannita, «The Autonomic System Functional State Predicts Responsiveness in Disorder of Consciousness», *J. Neurotrauma*, vol. 32, n. 14, pagg. 1071–1077, lug. 2015, doi: 10.1089/neu.2014.3539.
[27] A. Estraneo *et al.*, «A multicentre study of intentional behavioural responses measured using the Coma Recovery Scale–Revised in patients with minimally conscious state», *Clin. Rehabil.*, vol. 29, n. 8, pagg. 803–808, ago. 2015, doi: 10.1177/0269215514556002.
[28] J. T. Giacino, K. Kalmar, e J. Whyte, «The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility11No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the authors or upon any organization with which the authors are associated.», *Arch. Phys. Med. Rehabil.*, vol. 85, n. 12, pagg. 2020–2029, 2004, doi: https://doi.org/10.1016/j.apmr.2004.02.033.
[29] E. Portaccio *et al.*, «Improvement on the Coma Recovery Scale-Revised During the First Four Weeks of Hospital Stay Predicts Outcome at Discharge in Intensive Rehabilitation After Severe Brain Injury», *Arch. Phys. Med. Rehabil.*, vol. 99, n. 5, pagg. 914–919, mag. 2018, doi: 10.1016/j.apmr.2018.01.015.
[30] G. Galeoto, S. Turriziani, e A. Berardi, «Levels of Cognitive Functioning Assessment Scale: Italian cross-cultural adaptation and validation», *Ann. Ig. Med. Prev. E Comunità*, n. 1, pagg. 16–26, gen. 2020, doi: 10.7416/ai.2020.2326.
[31] G. Matis e T. Birbilis, «The Glasgow Coma Scale – a brief review Past, present, future», pag. 16.
[32] M. A. Crary, G. D. C. Mann, e M. E. Groher, «Initial Psychometric Assessment of a Functional Oral Intake Scale for Dysphagia in Stroke Patients», *Arch. Phys. Med. Rehabil.*, vol. 86, n. 8, pagg. 1516–1520, ago. 2005, doi: 10.1016/j.apmr.2004.11.049.
[33] G. Carl, H. Byron, L. John, H. Allen, e W. Benjamin, «American Journal of Physical Medicine & Rehab», *Perform. Profiles Funct. Independence Meas.*, vol. 72, pagg. 84–89, 1993.
[34] A. Morley e L. Hill, «10-20 system EEG Placement», pag. 34.
[35] L. J. Hirsch *et al.*, «American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology: 2021 Version», *J. Clin. Neurophysiol.*, vol. 38, n. 1, pagg. 1–29, gen. 2021, doi: 10.1097/WNP.0000000000000806.
[36] A. Estraneo *et al.*, «Standard EEG in diagnostic process of prolonged disorders of consciousness», *Clin. Neurophysiol.*, vol. 127, n. 6, pagg. 2379–2385, giu. 2016, doi: 10.1016/j.clinph.2016.03.021.
[37] S. Bagnato, C. Boccagni, A. Sant'Angelo, C. Prestandrea, R. Mazzilli, e G. Galardi, «EEG predictors of outcome in patients with disorders of consciousness admitted for intensive rehabilitation», *Clin. Neurophysiol.*, vol. 126, n. 5, pagg. 959–966, mag. 2015, doi: 10.1016/j.clinph.2014.08.005.
[38] H. Zou e T. Hastie, «Regularization and variable selection via the elastic net», *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 67, n. 2, pagg. 301–320, apr. 2005, doi: 10.1111/j.1467-9868.2005.00503.x.
[39] J. Fan e R. Li, «Statistical Challenges with High Dimensionality: Feature Selection in Knowledge Discovery», *arXiv:math/0602133*, feb. 2006, Consultato: 13 settembre 2021. [Online]. Disponibile su: http://arxiv.org/abs/math/0602133
[40] M. Yamada, K. Takeuchi, T. Iwata, J. Shawe-Taylor, e S. Kaski, «Localized Lasso for High-Dimensional Regression», pag. 9.
[41] F. Pedregosa *et al.*, «Scikit-learn: Machine Learning in Python», *Mach. Learn. PYTHON*, pag. 6.
[42] T. Akiba, S. Sano, T. Yanase, T. Ohta, e M. Koyama, «Optuna: A Next-generation Hyperparameter Optimization Framework», *ArXiv190710902 Cs Stat*, lug. 2019, Consultato: 21 giugno 2021. [Online]. Disponibile su: http://arxiv.org/abs/1907.10902
[43] G. C. Cawley e N. L. C. Talbot, «On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation», pag. 29.
[44] J. Wang, X. Hu, Z. Hu, Z. Sun, S. Laureys, e H. Di, «The misdiagnosis of prolonged disorders of consciousness by a clinical consensus compared with repeated coma-recovery scale-revised assessment», *BMC Neurol.*, vol. 20, n. 1, pag. 343, dic. 2020, doi: 10.1186/s12883-020-01924-9.

[45] M. Scarpino *et al.*, «EEG and Coma Recovery Scale-Revised prediction of neurological outcome in Disorder of Consciousness patients», *Acta Neurol. Scand.*, vol. 142, n. 3, pagg. 221–228, set. 2020, doi: 10.1111/ane.13247.
[46] R. Seel, M. Sherer, e J. Whyte, «Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research.», *Arch. Phys. Med. Rehabil.*, vol. 91, pagg. 1795–1813, 2010.
[47] M. Song *et al.*, «Prognostication of chronic disorders of consciousness using brain functional networks and clinical characteristics», *eLife*, vol. 7, pag. e36173, ago. 2018, doi: 10.7554/eLife.36173.
[48] J. Annen *et al.*, «Diagnostic accuracy of the CRS-R index in patients with disorders of consciousness», *Brain Inj.*, vol. 33, n. 11, pagg. 1409–1412, set. 2019, doi: 10.1080/02699052.2019.1644376.
[49] R. T. Seel *et al.*, «Assessment Scales for Disorders of Consciousness: Evidence-Based Recommendations for Clinical Practice and Research», *Arch. Phys. Med. Rehabil.*, vol. 91, n. 12, pagg. 1795–1813, dic. 2010, doi: 10.1016/j.apmr.2010.07.218.