Available online at www.sciencedirect.com

Resuscitation

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

Clinical paper

Does a combination of ≥2 abnormal tests vs. the ERC-ESICM stepwise algorithm improve prediction of poor neurological outcome after cardiac arrest? A post-hoc analysis of the ProNeCA multicentre study

Maenia Scarpino[a,b], *Francesco Lolli*[c], *Giovanni Lanzo*[a], *Riccardo Carrai*[a,b], *Maddalena Spalletti*[a,b], *Franco Valzania*[d], *Maria Lombardi*[e], *Daniela Audenino*[f], *Maria Grazia Celani*[g], *Alfonso Marrelli*[h], *Sara Contardi*[i], *Adriano Peris*[j], *Aldo Amantini*[b], *Antonello Grippo*[a,b,1], *Claudio Sandroni*[k,l,*]

[a] SODc Neurofisiopatologia, Dipartimento Neuromuscolo-Scheletrico e degli Organi di Senso, AOU Careggi, Florence, Italy
[b] IRCCS, Fondazione Don Carlo Gnocchi, Florence, Italy
[c] Dipartimento di Scienze Biomediche Sperimentali e Cliniche, Università degli Studi di Firenze, Italy
[d] UO Neurofisiopatologia, Arcispedale Santa Maria Nuova, Reggio nell'Emilia, Italy
[e] UO Neurologia, Ospedale San Giuseppe, Empoli, Italy
[f] SC Neurologia, EO Ospedale Galliera, Genova, Italy
[g] UO Neurofisiopatologia, Ospedale Santa Maria della Misericordia, Perugia, Italy
[h] UOC Neurofisiopatologia, Ospedale San Salvatore, L'Aquila, Italy
[i] Neurofisiopatologia Interventiva, Ospedale civile di Baggiovara, Modena, Italy
[j] Cure Intensive del Trauma e delle Gravi Insufficienze D'organo, AOU Careggi, Florence, Italy
[k] Department of Intensive Care, Emergency Medicine and Anaesthesiology, Fondazione Policlinico Universitario A. Gemelli-IRCCS, Rome, Italy
[l] Institute of Anaesthesiology and Intensive Care Medicine, Università Cattolica del Sacro Cuore, Rome, Italy

Abstract

Background: Bilaterally absent pupillary light reflexes (PLR) or N20 waves of short-latency evoked potentials (SSEPs) are recommended by the 2015 ERC-ESICM guidelines as robust, first-line predictors of poor neurological outcome after cardiac arrest. However, recent evidence shows that the false positive rates (FPRs) of these tests may be higher than previously reported. We investigated if testing accuracy is improved when combining PLR/SSEPs with malignant electroencephalogram (EEG), oedema on brain computed tomography (CT), or early status myoclonus (SM).

Methods: Post-hoc analysis of ProNeCA multicentre prognostication study. We compared the prognostic accuracy of the ERC-ESICM prognostication strategy vs. that of a new strategy combining ≥2 abnormal results from any of PLR, SSEPs, EEG, CT and SM. We also investigated if using alternative classifications for abnormal SSEPs (absent-pathological vs. bilaterally-absent N20) or malignant EEG (ACNS-defined suppression or burst-suppression vs. unreactive burst-suppression or status epilepticus) improved test sensitivity.

Results: We assessed 210 adult comatose resuscitated patients of whom 164 (78%) had poor neurological outcome (CPC 3–5) at six months. FPRs and sensitivities of the ≥2 abnormal test strategy vs. the ERC-ESICM algorithm were 0[0–8]% vs. 7 [1–18]% and 49[41–57]% vs. 63[56–71]%,

* Corresponding author at: Department of Intensive Care, Emergency Medicine and Anaesthesiology — Fondazione Policlinico Universitario A. Gemelli, IRCCS, Università Cattolica del Sacro Cuore — Roma, Largo Francesco Vito, 1, 00168 Roma, Italy.
E-mail address: claudio.sandroni@policlinicogemelli.it (C. Sandroni).
[1] Both authors equally contributed to the manuscript.

https://doi.org/10.1016/j.resuscitation.2020.12.003
; Accepted 13 February 2020

respectively (p < .0001). Using alternative SSEP/EEG definitions increased the number of patients with ≥2 concordant test results and the sensitivity of both strategies (67[59−74]% and 54[46−61]% respectively), with no loss of specificity.
Conclusions: In comatose resuscitated patients, a prognostication strategy combining ≥2 among PLR, SSEPs, EEG, CT and SM was more specific than the 2015 ERC-ESICM prognostication algorithm for predicting 6-month poor neurological outcome.
Keywords: Cardiac arrest, Coma, Prognosis

Introduction

Hypoxic-ischaemic brain injury (HIBI) is the major cause of death in patients who are comatose after resuscitation from cardiac arrest. Prediction of poor neurological outcome is essential to avoid futile care in these patients. However, the risk of a falsely pessimistic prediction should also be minimised in order to avoid an inappropriate withdrawal of life-sustaining treatment (WLST).

The 2015 prognostication algorithm recommended by the European Resuscitation Council (ERC) and the European Society of Intensive Care Medicine (ESICM)[1,2] predicts poor neurological outcome based on the presence of at least one most robust predictor, or, in case this criterion is not met, a combination of at least two less robust predictors. Most robust (first-line) predictors include a bilateral absence of either the pupillary light reflex (PLR) and corneal reflexes at ≥72 h from return of spontaneous circulation (ROSC) or the N20 wave of short-latency somatosensory evoked potentials (SSEPs) after rewarming from targeted temperature management (TTM). Less robust (second-line) predictors include the presence of status myoclonus within 48 h after ROSC, high blood levels of neuron specific enolase (NSE) at 48−72 h after ROSC, presence of unreactive burst-suppression or status epilepticus on electroencephalogram (EEG) after rewarming from TTM, or diffuse signs of HIBI on brain computed tomography (CT) or magnetic resonance imaging (MRI). Most robust predictors were identified on the basis of two evidence reviews[3,4] published in 2013 which showed that these indices had the lowest false positive rates (FPRs) for predicting poor neurological outcome. However, these reviews also showed that the certainty of the evidence supporting these indices was low or very low, the main reason being the risk of self-fulfilling prophecy. In fact, in most studies the results of the predictor under investigation were not concealed to the treating team and were often used to support decisions on WLST, potentially overestimating predictor accuracy. A series of multicentre studies[5−7] published after these reviews and a recent systematic review[8] indicate that the FPRs of absent ocular reflexes and N20 SSEP wave were higher than previously reported. This suggests that the distinction between more robust and less robust predictors may not be justified, and that even predictors previously classified as more robust should be used in combination in order to minimise the risk of a falsely pessimistic prediction. However, this might reduce algorithm sensitivity.

The primary aim of the present study was to compare the sensitivity and specificity for predicting poor neurological outcome at six months of the stepwise approach recommended in the 2015 ERC/ESICM prognostication algorithm vs. a prognostic strategy combining at least 2 abnormal results of any of the tests recommended in the algorithm without distinguishing between first-line and second-line predictors. The secondary aim was to investigate if the prognostic accuracy of EEG and SSEPs could be improved by using more recent classifications to define the abnormality of these tests.

Materials and methods

The ProNeCA multicentre study

The present investigation is based on a secondary analysis of data from the ProNeCA prospective multicentre study.[7,9,10] ProNeCA (Prognostication of Neurological outcome after Cardiac Arrest; NCT03849911) was conducted in the intensive care units (ICUs) of 13 Italian hospitals and coordinated by the Careggi University Hospital in Florence, Italy. The methodology of ProNeCA has been described in detail elsewhere.[11] Briefly, the study included consecutive adult patients (June 1, 2016 − June 1, 2018) who were admitted to the participating ICUs in a coma (Glasgow Coma Scale [GCS] ≤8) following resuscitation from cardiac arrest. The aim of the ProNeCA study was to assess the accuracy of EEG, SEPs (each recorded at 12−24 and 72 h after CA), and brain CT (recorded within the first 24 h after CA) for predicting poor neurological outcome, defined as severe neurological disability, persistent vegetative state, or death, corresponding to cerebral performance categories [CPC] 3−5. A CPC was assigned at hospital discharge using clinical assessment and at six months using telephone interviews. CPC assessors were blinded to the results of the investigated index tests. In 7/13 centres, all prognostic modalities were performed. In the remaining six centres, brain CT was not recorded. In this study, we only included data from the seven centres where results of all prognostic modalities were available. In addition, in order to ensure compatibility with the ERC-ESICM recommendations, we included only EEG and SSEP evaluations made at 72 h after ROSC.

Index test recording and classification

A routine 30-min EEG was recorded according to the International 10−20 System by board-certified neurologists and EEG technicians and classified according to the ACNS terminology.[12,13] SEP cortical responses of the median nerve were recorded using standard procedures.[14] The N20 and P25 waves were identified as the major negative peak with a latency of approximately 20 ms from the stimulus, and the major positive peak following N20, respectively. The SEP cortical response was defined as absent if no reproducible cortical component could be recorded in presence of the P14 lemniscal wave.[15−19] Brain CT scans (slice thickness 2.5−4.8 mm) were acquired within 24 h from ROSC. Circular (0.6cm^2) regions of interest (ROIs) were identified in the corpus callosum (CC), and − at the basal ganglia level − in the caudate nucleus (CN), putamen (PU), and the posterior limb of the internal capsule (PLIC), bilaterally. The severity of brain oedema was measured as the density ratio (GWR) between the grey and white matter:GWR = (CN + PU)/(CC + PLIC)
as previously described.[20] Based on our previous study, we adopted a GWR threshold of <1.21 to identify patients with poor neurological outcome. We visually detected standard (penlight) PLR on day 3 after ROSC using a penlight. Corneal reflexes were not investigated. We

defined status myoclonus according to the 2015 ERC-ESICM criteria[2] as continuous and generalized myoclonic jerks persisting for at least 30 min. NSE and MRI were not assessed. Index test results were interpreted by the study investigators at each participating hospital. Results of brain CTs were interpreted by investigators blinded to the patients' outcome. Results of the index tests were available to the treating team. They were used to assess the severity of HIBI, inform communication with patients' relatives, provide appropriate treatment of complications (e.g., seizures, for EEG) and rule out causes of brain damage other than HIBI (e.g., subarachnoid haemorrhage, for brain CT).

We compared the sensitivity of combining index tests using the ERC-ESICM algorithm vs. a combination of $\geq$2 abnormal results of any of the investigated tests. We also assessed the sensitivity of these combinations using alternative classifications of SSEPs and EEG. Based on the original ProNeCA study, in addition to the bilaterally absent N20 SSEP wave (AA pattern), we also tested the AP pattern (absent N20 wave on one side, pathological N20 wave on the other side). A pathological SSEP result was defined as N20/P25 amplitude lower than the limit of normality (5th percentile) for each participating centre. Considering EEG, aside from the malignant EEG pattern defined by ERC-ESICM (unreactive burst-suppression or status epilepticus) we tested an alternative definition of malignant EEG as isoelectric, suppression, or burst-suppression, classified according to the ACNS terminology.[12] This definition is used in recent literature.[21–23]

Based on a recent study[24] which showed that the ERC-ESICM algorithm yields no false positive predictions when a motor response no better than abnormal flexion (GCS motor score M $\leq$ 3) is used as a screening criterion for prognostication, we also used M $\leq$ 3 as an entry point of the prognostication strategies we tested. However, since the entry point recommended by the 2015 ERC-ESICM guidelines is currently M $\leq$ 2, we performed a supplementary analysis restricted to patients with M $\leq$ 2.

As a supplementary analysis, we also assessed the predictive value of GWR thresholds above 1.21 both individually and in combination with positive results of other tests, to investigate if this would increase the sensitivity of the model while maintaining 100% specificity.

Finally, as we did in the main ProNeCA study reports[7,9] we calculated the performance of the index tests also for poor outcome defined as CPC 4–5 (persistent vegetative state or death). In fact, although the majority of the most recent studies on neuroprognostication after cardiac arrest used CPC = 3–5 for defining poor neurological outcome after cardiac arrest,[8,25] dichotomising outcomes as CPC 4–5 vs. 1–3 may be of value for some communities or individuals for whom CPC = 3 (awake, with severe neurological disability) may be seen as an acceptable poor outcome.[26]

Patient management

Patients were managed according to local practices at each participating ICU. The choice of the TTM target temperature, i.e., 34 °C vs. 36 °C, as well as the choice of sedatives, analgesics and neuromuscular blocking agents, were at the discretion of the participating centre. However, use of TTM for a minimum of 24 h, avoiding fever (central body temperature below 37.5 °C) until 72 h after CA, and use of short-acting sedative agents were recommended. WLST was not performed in any of the participating centres and treatment was continued for all patients, except in the event of brain death. Patients who had not recovered consciousness after at least one week after suspension of sedation and had bilaterally absent N20 SSEP wave were destined to a long-term care unit. All other patients were destined to a rehabilitation unit. None of the study investigators

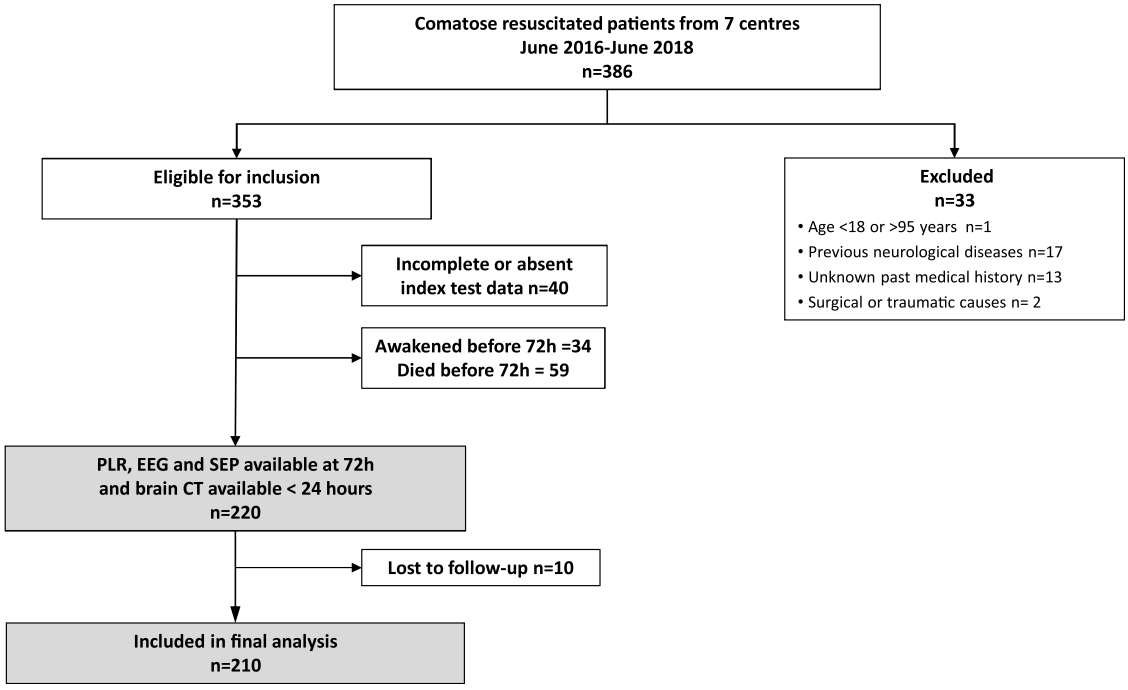

Fig. 1 – Flow-chart of patient inclusion. *Abbreviations* – PLR, Pupil Light Reactivity; CA, Cardiac Arrest; EEG, electroencephalogram, SEP, somatosensory evoked potentials: CT, computed tomography.

were involved in patient management either during or after the patients' ICU stay.

Ethical approval

The protocol was approved by the Regional Ethics Committee of Tuscany (Ref. OSS.15.009). Written informed consent was obtained from the patient's authorized representative prior to the subject's enrolment.

Statistical analysis

Continuous variables were reported as median and inter-quartile range (IQR), whereas categorical variables were reported as numbers and percentages. Normality of baseline distribution was tested using the Shapiro–Wilk test. The Pearson's chi-square and the Mann–Whitney *U* tests were used for comparing categorical and continuous variables, respectively. We calculated sensitivity and FPR (1-specificity) for each individual predictor and for their combinations. We reported the number of combinations of two or more positive results of each of the included predictors in a Venn diagram. We tested sensitivity differences among various combinations of prognostic tests using the McNemar test.

Statistical analysis was performed using Wizard 1.9 version (Evan Miller, USA) and IBM-SPSS Statistics for Windows 25.0 version (IBM Corp., Armonk, NY, USA). Data reporting in this study is compliant with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) Statement, 2015 version.[27] The STARD checklist for this study is included as an ESM.

Results

The ProNeCA cohort included a total of 445 comatose post-CA patients. Of them, 386 came from centres where all prognostic tests were available. After exclusion of 33 non-eligible patients, 353 patients were considered for inclusion. Of these, 34 awakened and 59 died before the 72-h assessment, while 40 patients had incomplete or absent index test data and 10 patients were lost to follow-up. The remaining 210 patients were included in the analysis (Fig. 1). Their characteristics are reported in Table 1. Mean age was 64 (range 18–85) years, and 140 (66.6%) patients were male. Cardiac arrest was witnessed in 183 (87.3%) patients and initial cardiac rhythm was shockable in 95 (45.2%). Overall, 127 (60.5%) patients survived to hospital discharge, of whom 46 (21.9% of the initial cohort) had good neurological outcome at six months.

Individual performance of prognostic tests

Among the 210 patients we included, 197 had a M $\leq$ 2 and 13 (6.2%) had M = 3. Table 2 reports the individual accuracy of the five index tests we assessed in the global cohort. Sensitivity ranged from 3[1–6]% for status myoclonus to 49[41–57]% for PLR. All predictors

Table 1 – Characteristics of the study population.

Variables	Total population n = 210	Favourable outcome n = 46	Unfavourable outcome n = 164	*p* value
Age, years (range)	64 (18–85)	61.7 (18–85)	67.9 (55–75)	<0.0001
Gender, male	140 (66.6)	31 (67.3)	109 (66.4)	0.514
Out-of-hospital	153 (73)	33 (71.7)	120 (73.1)	0.872
Witnessed	183 (87.3)	40 (86.9)	143 (87.1)	0.105
CA duration min (range)	21.3 (9–35)	17.35(9–20)	27.8 (11–35)	0.001
Arrest from non cardiac etiology	53 (25.2)	13 (28.2)	40 (24.3)	0.563
Initial rhythm				0.001
VF/pVT	95 (45.2)	26 (56.5)	69 (42.0)	
PEA	56 (26.6)	11 (23.9)	45 (27.4)	
Asystole	43 (20.5)	4 (8.6)	39 (23.7)	
Unknown	16 (7.6)	3 (6.5)	13 (7.9)	
				0.001
Absent PLR at 72 h	83 (39)	3 (6.5)	80 (48.7)	
GCS at ICU admission (range)	3 (3–8)	3 (3–8)	3 (3–7)	0.06
GCS at 72 h (range)	3 (3–8)	3 (3–8)	3 (3–7)	0.06
TTM				0.06
No	116 (55.2)	22 (47.8)	94 (57.3)	
34 °C	84 (40.1)	23 (50.8)	61 (37.1)	
36 °C	10 (4.7)	3 (6.5)	7 (4.2)	
CPC at 6 months				
CPC 1	21 (10)	21 (10.0)	–	
CPC 2	25 (11.9)	25 (11.9)	–	
CPC 3	8 (3.8)	–	8 (3.8)	
CPC 4	36 (17.1)	–	36 (17.1)	
CPC 5	120 (57.1)	–	120 (57.1)	

Data are presented as number (percentage) or median (interquartile range for GCS score).
Abbreviations: CA, cardiac arrest; CPC, Cerebral Performance Category; GCS, Glasgow Coma Scale; ICU, intensive care unit; PEA, pulseless electrical activity; pVT, pulseless ventricular tachycardia; VF, ventricular fibrillation; TTM, targeted temperature management.

Table 2 – Individual test accuracy for prediction of poor outcome.

Index test	TP	FP	TN	FN	Sensitivity % (95%CI)	False positive rate % (95%CI)
Pupillary Light Reflex at 72 h	80	3	43	84	49 (41–57)	6 (1–18)
SSEPs at 72 h						
A/A pattern (ERC-ESICM)	71	0	46	93	43 (36–51)	0 (0–8)
A/P pattern (ProNeCA)	85	0	46	79	52 (44–60)	0 (0–8)
EEG at 72 h						
Unreactive burst-suppression or status epilepticus (ERC-ESICM)	23	0	46	141	14 (9–20)	0 (0–8)
ACNS-defined suppression or burst-suppression (ProNeCA)	78	0	46	87	47 (39–55)	0 (0–8)
Brain CT ≤24 h						
GWR <1.21	57	0	46	107	35 (27–43)	0 (0–8)
GWR < 1.25	82	3	43	82	46 (42–58)	7 (2–18)
Status myoclonus ≤72 h	4	0	46	160	3 (1–6)	0 (0–8)

CI: Confidence interval; EEG: electroencephalogram; GWR: grey matter/white matter ratio; SSEPs: short-latency somatosensory evoked potentials. TP, true positive; FP, false positive; TN, true negative; FN, false negative.

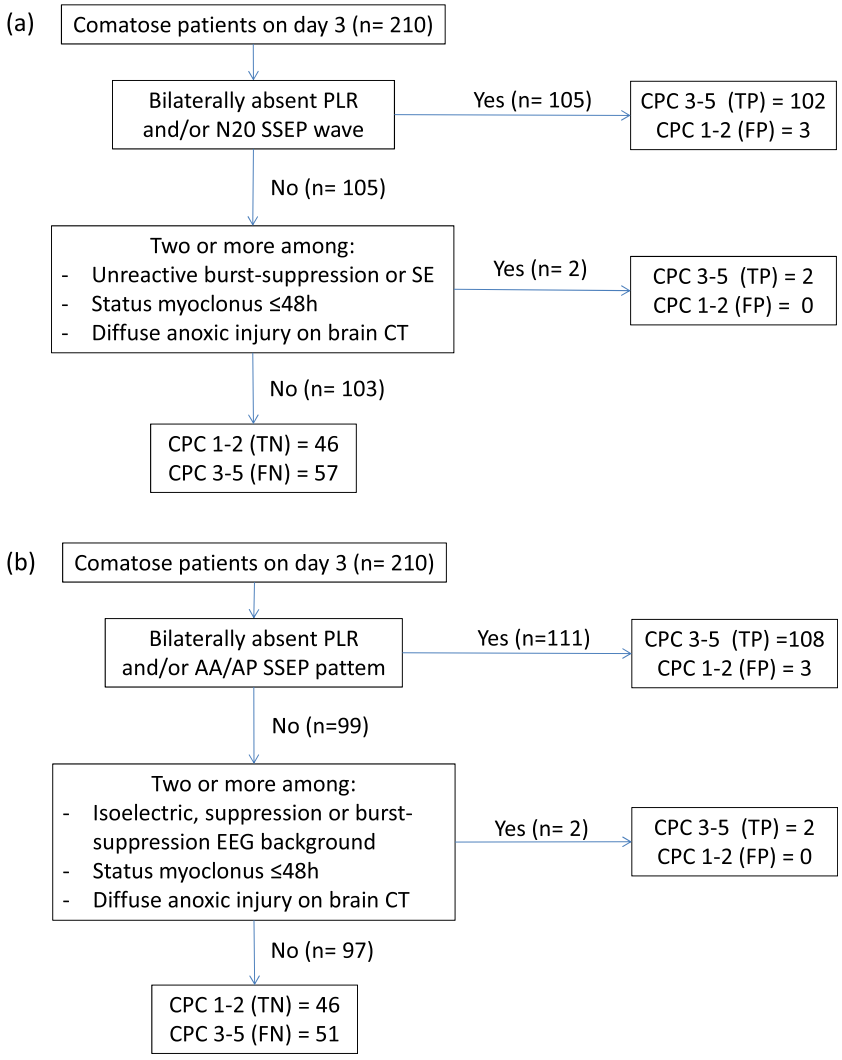

Fig. 2 – a, b — Performance of the ERC-ESICM prognostication algorithm with the original EEG/SSEP classifications (a) and with the modified classifications (b). We report the number of patients meeting the ERC/ESICM criteria for poor outcome, and the actual outcomes. *Abbreviations* — TP = true positives (patients with predicted and actual poor outcome); FP = false positives (patients with predicted poor outcome who had neurological recovery); TN = true negatives (patients not meeting the criteria for poor outcome who had neurological recovery); FN = false negatives (patients not meeting the criteria for poor outcome who had a final poor outcome).

had 0% FPR, except bilaterally absent PLR, for which three false positive test results occurred (FPR 6[1−18]%).

Electrophysiological tests were performed at a mean of 72 ± 8 h from ROSC. Sensitivity for prediction of poor neurological outcome of the malignant EEG defined according to ProNeCA was 47[39−55]% vs.14[9−20]% of the ERC-ESICM criteria for malignant EEG ($p < .0001$). Sensitivity of the SSEP AA/AP pattern was 52[44−60]% vs. 43[36−51]% of the conventional AA pattern ($p = .0002$) (Table 2).

Performance of the ERC-ESICM prognostication algorithm

A total of 105/210 (49.5%) patients had abnormal results of one or both the first-line predictors for poor outcome, i.e., a bilaterally absent PLR or N20 SSEP wave. Of these, 3 patients with bilaterally absent PLR had good neurological outcome at six months. Among the remaining 106 patients who were assessed with second-line predictors, two had ≥2 abnormal test results, and both had poor outcomes (Fig. 2a). The overall sensitivity and FPR of the ERC-ESICM algorithm were 63 [56−71]% and 7[1−18]%, respectively (Table 3).

When the bilaterally absent SSEP and the ERC-ESICM malignant EEG patterns were replaced with the AA/AP pattern and with the malignant EEG defined according to the ProNeCA study, six additional patients with poor outcome were identified at the first step of the algorithm, while no change was observed at the second step (Fig. 2b). The overall algorithm sensitivity consequently raised to 67 [59−74]% ($p = .02$), with no change in FPR (Table 3).

Accuracy of combining ≥2 abnormal test results

Two or more abnormal test results were present in 80 patients when using the ERC-ESICM criteria for SSEP and EEG and in 88 patients when using the ProNeCA criteria. The relevant sensitivities were significantly lower when compared to the ERC-ESICM stepwise algorithm (49[41−57]% vs. 63[56−71]% and 54[46−61]% vs. 67[59−74]% respectively; Table 3; $p < .0001$ for both). However, no false positive predictions occurred when combining ≥2 abnormal test results (FPR 0[0−8]%).

The Venn diagram in Fig. 3a shows the concordance of abnormal test results for PLR, brain CT, SSEP, and EEG when the ERC-ESICM criteria were used for these last two tests. Three patients had all four abnormal tests, 14 had three abnormal tests, 63 had two abnormal tests and 49 had only one abnormal test result. The most common concordances were observed among SSEP, PLR and brain CT (n = 11), SSEP and PLR (n = 23), and SSEP and brain CT (n = 13). In addition, status myoclonus was present in two patients with absent PLR, in one patient with abnormal brain CT, and in one patient with abnormal SSEPs (Fig. 2a and ESM Table 1).

When the ProNeCA criteria for SSEP and EEG were used, a greater proportion of abnormal SSEP or EEG results were in concordance with those of the other three tests. The number of patients with all four abnormal test results raised from 3 to 27, and the number of patients with three abnormal tests raised from 14 to 40 (Fig. 3b and ESM Table 1). In particular, 61 additional patients had abnormal EEG results in combination with those of the three other tests, especially with brain CT.

Supplementary analyses

Results of sensitivity analysis for GWR on brain CT (see ESM Table 2 and ESM Fig. 1) showed that a GWR threshold above <1.21 was associated with false positive results. FPR ranged from 3% for a GWR < 1.22−18% for a GWR < 1.26. However, when combining GWR thresholds <1.22 or above with one or more other abnormal test results, the FPR was 0% up to a GWR threshold of <1.25. The sensitivity of this combination was slightly higher at a GWR threshold of <1.25 than at a threshold of <1.21 (56 [48−64]% vs. 54 [46−61]%; see Table 3).

In the supplementary analysis restricted to patients with M ≤ 2 only one false positive result was observed for PLR (ESM Table 3). Consequently, the 2015 ERC-ESICM prognostication algorithm yielded a non-significant lower rate of false positives than when M ≤ 3 was used as an entry criterion (FPR 3[0−13]% vs. 6[1−18]%; ESM Table 4), Sensitivity of both single predictors and combination strategies did not change according to the motor threshold.

When the accuracy of individual predictors was assessed using CPC 4−5 as a criterion for poor neurological outcome, one additional false positive prediction occurred with ACNS-defined suppression or burst-suppression on EEG (ESM Table 5). This did not affect the overall FPR of the prognostication strategies we assessed (either ERC-ESICM or ≥2 positive test results; ESM Table 6). Sensitivity of both single predictors and combination strategies for predicting CPC 4−5 was consistently higher than for predicting CPC 3−5.

Discussion

Our study showed that combining two or more of any prognostic indices recommended by the ERC-ESICM guidelines predicted poor neurological outcome after cardiac arrest with lower sensitivity but

Table 3 – Prognostic accuracy of different combinations of the individual tests investigated in our study.

Prognostication strategy	TP	FP	TN	FN	Sensitivity % (95% CI)	False positive rate % (95%CI)
ERC-ESICM algorithm						
ERC-ESICM EEG/SSEP criteria	104	3	43	60	63 (56−71)	7 (1−18)
ProNeCA EEG/SSEP criteria	110	3	43	54	67 (59−74)	7 (1−18)
≥2 abnormal test results						
ERC-ESICM EEG/SSEP criteria	80	0	46	84	49 (41−57)	0 (0−8)
ProNeCA EEG/SSEP criteria	88	0	46	76	54 (46−61)	0 (0−8)
ProNeCA EEG/SSEP criteria (GWR < 1.25)	92	0	46	72	56 (48−64)	0 (0−8)

Abbreviations: CI: confidence interval; EEG: electroencephalogram; SSEPs: short-latency somatosensory evoked potentials. TP, true positive; FP, false positive; TN, true negative; FN, false negative.

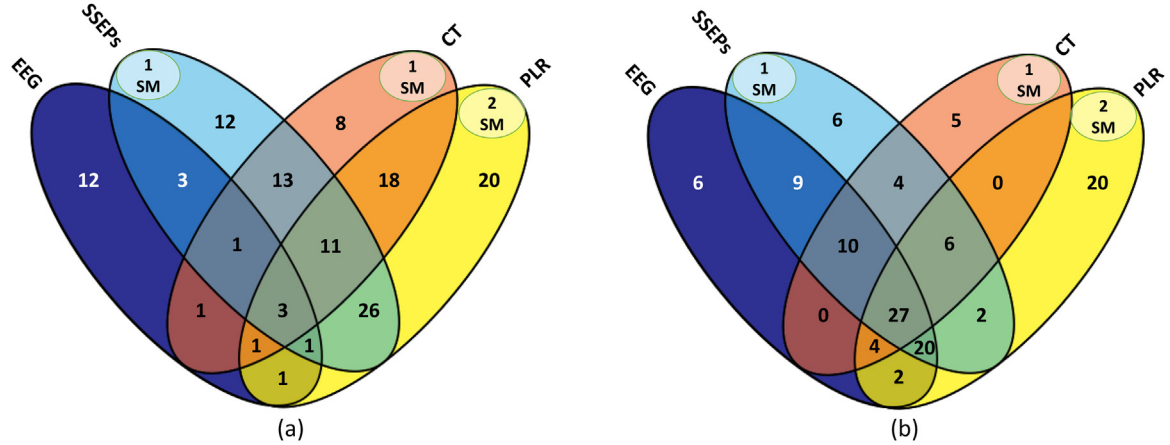

Fig. 3 – a, b — Venn diagram showing the concordance of abnormal tests results when the ERC-ESICM (a) or the ProNeCA (b) classification criteria for SSEP and EEG were used. The numbers within each diagonal oval area correspond to the number of patients having abnormal results of each test or combination of tests. As shown by the comparison between the two figures, when SSEP and EEG were classified according to ProNeCA, the number of patients with ≥2 abnormal test results increased, while the number of patients with single test results decreased. The shift was mainly led by EEG results. *Abbreviations* — EEG = electroencephalogram; SSEPs = somatosensory evoked potentials; CT = computed tomography; PLR = pupillary light response; SM = status myoclonus (inner circles).

higher specificity than the 2015 ERC-ESICM stepwise algorithm. We also showed that the recent criteria for unfavourable SSEP and EEG assessed in our population were as specific as the corresponding ERC-ESICM criteria but had greater sensitivity. In addition, when these alternative criteria were used, both SSEP and EEG results showed a greater concordance with the results of other tests.

The 2015 ERC-ESICM guidelines for prognostication after cardiac arrest suggest a multimodal approach in which the unfavourable prognostic indication from an abnormal test result should be confirmed by the results of at least another prognostic test. However, for both SSEPs and PLR the approach remained substantially unimodal, since an abnormal result of either of these tests in isolation was considered as a sufficient criterion for predicting poor outcome with high likelihood (defined as the upper boundary of 95%CIs for FPR below 5%).[2] This precision of poor outcome prediction was estimated on the basis of two 2013 systematic reviews of prognostic tests informing the guidelines.[3,4] However, more recent studies and a recent review [8] reported a higher rate of false positive results from SSEPs and PLR than that documented in those previous reviews. In two prognostication studies, the FPR of a bilaterally absent N20 SSEP wave was 50%[28] at 24−48 h and 25% at 24−72 h,[29] respectively. In a multicentre study, among 39 patients with good neurological outcome at six months, one had a bilaterally absent N20 SSEP wave (FPR 2.6%).[5]. In another multicentre study, three false positive SSEP results were reported (FPR 3[1−7]%).[30] On post-hoc assessment, the cause of these three false positive results was clearly identified as an incorrect reading of the SSEP record due to excessive noise. This is in line with results of other studies showing that the risk of SSEP misclassification increases when signal/noise ratio is low.[31−34] As far as bilateral absence of PLR is concerned, two multicentre studies [5,6] and a large single-centre study[35] published after the 2015 guidelines reported a 6−7% FPR at 72 h or later for this sign, which is in line with the rate observed in our study. As for SSEPs, the most likely reason for these results could have been due to interference on index assessment. Residual effects from sedatives used during TTM, especially opioids, may reduce pupil size and make PLR less detectable. In one of the studies mentioned above[6] all five patients with falsely positive results of conventional PLR at 72 h had miosis. In these patients, automated quantitative pupillometry revealed that PLR was actually present, but the variation of pupil size in response to light was on average 0.25 mm, too little to be detected to the naked eye. Miosis can also reduce the interrater reliability of conventional PLR.[36] In our study, clinical examination was performed after TTM and off sedation. However, we cannot exclude that residual sedation may have been present and may have interfered with PLR assessment.

When bilaterally absent PLR or N20 SSEP waves were combined with other index tests, no false positive predictions occurred in our population (FPR 0[0−8]%). This is essential, since a high degree of certainty is required when prognosticating for medical decisions regarding life sustaining treatments.[37] However, the 0% FPR we obtained in our cohort through this combination strategy came at the cost of a lower overall sensitivity. In fact, when ≥2 abnormal test results were required, only 49% of the patients who had a poor outcome at six months were identified, as compared to 63% when using the 2015 ERC-ESICM stepwise approach (67% when using the ProNeCA SSEP/EEG criteria). This confirms the results a recent study where any ≥2 pathological findings of the various predictors suggested in the 2015 ERC-ESICM algorithm were combined as a criterion for a likely poor outcome.[24] In that study, a strategy combining ≥2 abnormal findings from clinical examination, electrophysiology, biomarkers and imaging yielded a 34.6% sensitivity vs. 38.7% when using the 2015 ERC-ESICM algorithm. However, while in that study neither of these two strategies had false positive results, in our study only a ≥2 combination strategy ensured a 0% FPR. This advantage was consistent even when alternative criteria for entering the algorithm or for defining poor outcome were used in our supplementary analyses.

Differently from the Moseby-Knappe study mentioned above,[24] increasing the threshold of motor score to ≤3 (i.e., including patients with abnormal flexion) vs. ≤2 as an entry criterion for the prognostication

algorithm did not result in an increase of sensitivity in any of the prognostication strategies we tested. This was because in our cohort most patients with M = 3 and poor outcome were not detected by the predictive tests, which resulted in an increase in false negative predictions (the denominator, in the sensitivity formula) with an only minimal increase in the number of true positives. Notably, the proportion of patients with poor outcome among patients with M = 3 in our study (62%) was the same than in the Moseby-Knappe study, which suggests against an explanation based on differences in case mix between studies. A more likely explanation is a difference in the set of tests used in our study vs. the Moseby-Knappe study. As far as specificity is concerned, using an M ≤ 3 threshold resulted in an increased FPR because of two additional false positive results from PLR were observed. However, this was limited to the 2015 ERC-ESICM prognostication strategy. In fact, no false positives were observed when the strategy of combining ≥2 abnormal test results was used. This further supports our study hypothesis that even predictors previously classified as more robust should be used in combination in order to minimise the risk of a falsely pessimistic prediction.

Another important finding of our study is that replacing the 2015 ERC-ESICM criteria for abnormal SSEP or EEG results with alternative criteria based on more recent literature increased the sensitivity of these tests. While the additional value of the ProNeCA criteria for SSEPs (AA/AP pattern) awaits external validation from further studies, the increased sensitivity we found for malignant EEG (suppression or burst-suppression defined according to ACNS terminology) as compared to the 2015 ERC-ESICM criteria is confirmed by recent investigations. Interestingly, in our study the results of both SSEP and EEG analysed using new criteria had greater concordance with the results of other index tests indicating poor outcome than the corresponding ERC-ESICM criteria. A greater concordance of signals from different tests may increase the clinician's confidence in the correctness of the prediction when decisions about life sustaining treatment are to be made.

Results of our exploratory analysis for GWR thresholds above 1.21 showed that while these higher thresholds are associated with false positive predictions when considered individually, their combination with abnormal results of the other tests maintained a 0% FPR up to a GWR threshold of <1.25. Although the relevant increase in sensitivity was modest, this result indirectly confirms the validity of our approach and suggests that similar combinations might be explored with other predictors based on continuous variables, especially blood biomarkers, for which achieving a consistent threshold for 100% specificity is difficult because of outliers.

An important strength of this study is that no WLST was performed in our population, and life support was continued in all patients regardless of the result of prognostic tests. This minimised the risk of self-fulfilling prophecy for the individual predictors we investigated. In addition, almost 90% of the included patients had results of all prognostic tests available. This allowed a consistent and accurate evaluation of individual test performance, especially sensitivity.

Our study has also limitations. Firstly, the ProNeCA study did not investigate corneal reflexes, biomarkers and MRI, so that their combination with other predictors could not be assessed, and the sensitivity of the prognostication strategies we investigated might have been underestimated. Unfortunately, an incomplete assessment of the entire set of recommended tests is common in retrospective studies aimed to validate the ERC-ESICM algorithm,[38,39] and specifically designed prospective studies would be necessary for this purpose. Secondly, we collected PLR as part of the study protocol, but assessment of their accuracy was not included in the original study aims and their measurement was not standardised. This may have overestimated their false positive rate. Thirdly, the treating team was not blinded to the results of the index test under investigation. Consequently, we cannot exclude that this may have affected patient management and created a self-fulfilling prophecy bias. However, because of the absence of WLST in our population, we presume that the impact of this bias was much less than in other prognostication studies. Finally, although the calculation method for GWR we adopted to assess HIBI using brain CT is the most described in medical literature, there is no universal consensus on the most accurate method of calculating GWR. This may limit the reproducibility of our results.

Conclusions

Results of our study suggest that in the majority of patients with poor neurological outcome after cardiac arrest, at least two of the following unfavourable prognostic signs recommended by the current ERC-ESICM guidelines (bilaterally absent PLR or N20 SSEP wave, presence of status myoclonus, malignant EEG, and signs of HIBI on brain CT) coexist in the same patient. Using a combination of abnormal results from two of any of these tests avoided false positive predictions in our study but also reduced sensitivity when compared to the 2015 ERC-ESICM stepwise approach. Using more recent definitions of abnormal test results for both SSEPs and EEG improved their sensitivity while maintaining 0% FPR. Using M ≤ 3 as an entry point to the prognostication algorithm as an alternative to M ≤ 2 had 0% FPR when a combination of ≥2 abnormal tests was used, but did not increase sensitivity. Further studies are needed to investigate the added value of combining results of biomarkers of neuronal injury and MRI.

Funding

None.

Conflict of interest statement

Claudio Sandroni is member of the Editorial Board of *Resuscitation*.

The remaining authors have no conflict of interest to disclose.

Appendix A. Members of the ProNeCA study group

Aldo Amantini, IRCCS Fondazione Don Carlo Gnocchi, Florence; Daniela Audenino, SC Neurologia, EO Ospedale Galliera, Genova; Chiara Bandinelli, UO Neurologia, Ospedale San Giuseppe, Empoli; Pasquale Bernardo, Interventistica Cardiologica Strutturale, AOU Careggi, Firenze; Teresa Anna Cantisani, UO Neurofisiopatologia, Ospedale Santa Maria della Misericordia, Perugia; Riccardo Carrai, SODc Neurofisiopatologia, AOU Careggi e IRCCS Fondazione Don Carlo Gnocchi, Florence, Florence; Maria Grazia Celani, UO Neurofisiopatologia, Ospedale Santa Maria della Misericordia, Perugia; Roberta Ciuffini, UOC Neurofisiopatologia, Ospedale San Salvatore, L'Aquila; Sara Contardi, Neurofisiopatologia Interventiva, Ospedale civile di Baggiovara, Modena; Antonello Grippo, SODc

Neurofisiopatologia, AOU Careggi e IRCCS, Fondazione Don Carlo Gnocchi, Florence; Giovanni Lanzo, SODc Neurofisiopatologia, AOU Careggi, Florence; Francesco Lolli, Dipartimento di Scienze Biomediche Sperimentali e Cliniche, Università degli Studi di Firenze; Maria Lombardi, UO Neurologia, Ospedale San Giuseppe, Empoli; Alfonso Marrelli, UOC Neurofisiopatologia, Ospedale San Salvatore, L'Aquila; Andrea Marudi, Neurofisiopatologia Interventiva, Ospedale civile di Baggiovara, Modena; Giulia Masi, SODc Neurofisiopatologia, AOU Careggi, Firenze Chiara Minardi, UO Neurologia Ospedale Bufalini, Cesena; Marco Moretti, UO Neuroradiologia, AOU Careggi, Florence; Giuseppe Olivo, Terapia Intensiva Cardiochirurgica, AOU Careggi; Adriano Peris, Unità di Terapia Intensiva, AOU Careggi, Florence; Klaudio Rikani, SC Neurologia, EO Ospedale Galliera, Genova; Rossella Sabadini, UO Neurofisiopatologia, Arcispedale Santa Maria Nuova, Reggio nell'Emilia; Claudio Sandroni, Università Cattolica del Sacro Cuore – Fondazione Policlinico Universitario "Agostino Gemelli" -IRCCS – Rome; Maenia Scarpino, SODc Neurofisiopatologia, AOU Careggi, e IRCCS, Fondazione Don Carlo Gnocchi, Florence; Maddalena Spalletti, SODc Neurofisiopatologia, AOU Careggi, Florence; Franco Valzania, UO Neurofisiopatologia, Arcispedale Santa Maria Nuova, Reggio nell'Emilia.

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.resuscitation.2020.12.003.

REFERENCES

1. Nolan JP, Soar J, Cariou A, et al. European Resuscitation Council and European Society of Intensive Care Medicine Guidelines for post-resuscitation care 2015: section 5 of the European Resuscitation Council Guidelines for Resuscitation 2015. Resuscitation 2015;95:202–22.
2. Sandroni C, Cariou A, Cavallaro F, et al. Prognostication in comatose survivors of cardiac arrest: an advisory statement from the European Resuscitation Council and the European Society of Intensive Care Medicine. Resuscitation 2014;85:1779–89.
3. Sandroni C, Cavallaro F, Callaway CW, et al. Predictors of poor neurological outcome in adult comatose survivors of cardiac arrest: a systematic review and meta-analysis. Part 2: patients treated with therapeutic hypothermia. Resuscitation 2013;84:1324–38.
4. Sandroni C, Cavallaro F, Callaway CW, et al. Predictors of poor neurological outcome in adult comatose survivors of cardiac arrest: a systematic review and meta-analysis. Part 1: patients not treated with therapeutic hypothermia. Resuscitation 2013;84:1310–23.
5. Dragancea I, Horn J, Kuiper M, et al. Neurological prognostication after cardiac arrest and targeted temperature management 33 degrees C versus 36 degrees C: results from a randomised controlled clinical trial. Resuscitation 2015;93:164–70.
6. Oddo M, Sandroni C, Citerio G, et al. Quantitative versus standard pupillary light reflex for early prognostication in comatose cardiac arrest patients: an international prospective multicenter double-blinded study. Intensive Care Med 2018;44:2102–11.
7. Scarpino M, Carrai R, Lolli F, et al. Neurophysiology for predicting good and poor neurological outcome at 12 and 72 h after cardiac arrest: The ProNeCA multicentre prospective study. Resuscitation 2020;147:95–103.
8. Sandroni C, D'Arrigo S, Cacciola S, et al. Prediction of poor neurological outcome in comatose survivors of cardiac arrest: a systematic review. Intensive Care Med 2020;46:1803–51.
9. Scarpino M, Lolli F, Lanzo G, et al. Neurophysiology and neuroimaging accurately predict poor neurological outcome within 24 hours after cardiac arrest: The ProNeCA prospective multicentre prognostication study. Resuscitation 2019;143:115–23.
10. Scarpino M, Lolli F, Lanzo G, et al. Neurophysiological and neuroradiological test for early poor outcome (Cerebral Performance Categories 3–5) prediction after cardiac arrest: prospective multicentre prognostication data. Data in Brief. 2019.
11. Scarpino M, Carrai R, Lolli F, et al. EEG and SEP evaluation for good and poor neurological prognosis after cardiac arrest: a prospective multicenter cohort trial (ProNeCA). Future Neurol 201914: FNL 16.
12. Hirsch LJ, Laroche SM, Gaspard N, et al. American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology: 2012 version. J Clin Neurophysiol 2013;30:1–27.
13. Spalletti M, Carrai R, Scarpino M, et al. Single electroencephalographic patterns as specific and time-dependent indicators of good and poor outcome after cardiac arrest. Clin Neurophysiol 2016;127:2610–7.
14. Scarpino M, Lanzo G, Carrai R, et al. Predictive patterns of sensory evoked potentials in comatose brain injured patients evolving to brain death. Neurophysiol Clin 2017;47:19–29.
15. Grippo A, Carrai R, Scarpino M, et al. Neurophysiological prediction of neurological good and poor outcome in post-anoxic coma. Acta Neurol Scand 2017;135:641–8.
16. Carrai R, Grippo A, Scarpino M, et al. Time-dependent and independent neurophysiological indicators of prognosis in post-anoxic coma subjects treated by therapeutic hypothermia. Minerva Anestesiologica 2016;82:940–9.
17. Scarpino M, Lanzo G, Lolli F, et al. Is brain computed tomography combined with somatosensory evoked potentials useful in the prediction of brain death after cardiac arrest? Neurophysiol Clin 2017;47:327–35.
18. Carrai R, Scarpino M, Lolli F, et al. Early-SEPs' amplitude reduction is reliable for poor-outcome prediction after cardiac arrest? Acta Neurol Scand 2019;139:158–65.
19. Scarpino M, Grippo A, Lanzo G, Lolli F. The burden of clinical neurophysiology for the neurological prognosis of coma. Future Neurol 2018;13:127–9.
20. Gentsch A, Storm C, Leithner C, et al. Outcome prediction in patients after cardiac arrest: a simplified method for determination of gray-white matter ratio in cranial computed tomography. Clin Neuroradiol 2015;25:49–54.
21. Backman S, Cronberg T, Friberg H, et al. Highly malignant routine EEG predicts poor prognosis after cardiac arrest in the Target Temperature Management trial. Resuscitation 2018;131:24–8.
22. Lamartine Monteiro M, Taccone FS, Depondt C, et al. The prognostic value of 48-h continuous EEG during therapeutic hypothermia after cardiac arrest. Neurocrit Care 2016;24:153–62.
23. Ruijter BJ, Tjepkema-Cloostermans MC, Tromp SC, et al. Early electroencephalography for outcome prediction of postanoxic coma: a prospective cohort study. Ann Neurol 2019;86:203–14.
24. Moseby-Knappe M, Westhall E, Backman S, et al. Performance of a guideline-recommended algorithm for prognostication of poor neurological outcome after cardiac arrest. Intensive Care Med 2020;46:1852–62.
25. Sandroni C, Nolan JP. Neuroprognostication after cardiac arrest in Europe: new timings and standards. Resuscitation 2015;90:A4–5.
26. Geocadin RG, Callaway CW, Fink EL, et al. Standards for studies of neurological prognostication in comatose survivors of cardiac arrest: a scientific statement from the American Heart Association. Circulation 2019;140:e517–42.
27. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
28. Leao RN, Avila P, Cavaco R, Germano N, Bento L. Therapeutic hypothermia after cardiac arrest: outcome predictors. Rev Bras Ter Intensiva 2015;27:322–32.
29. Dhakal LP, Sen A, Stanko CM, et al. Early absent pupillary light reflexes after cardiac arrest in patients treated with therapeutic hypothermia. Ther Hypothermia Temp Manag 2016;6:116–21.

30. Bouwes A, Binnekade JM, Kuiper MA, et al. Prognosis of coma after therapeutic hypothermia: a prospective cohort study. Ann Neurol 2012;71:206–12.
31. Pfeifer R, Weitzel S, Gunther A, et al. Investigation of the inter-observer variability effect on the prognostic value of somatosensory evoked potentials of the median nerve (SSEP) in cardiac arrest survivors using an SSEP classification. Resuscitation 2013;84:1375–81.
32. Hakimi K, Kinney G, Kraft G, Micklesen P, Robinson L. Reliability in interpretation of median somatosensory evoked potentials in the setting of coma: factors and implications. Neurocrit Care 2009;11:353–61.
33. Zandbergen EG, Hijdra A, de Haan RJ, et al. Interobserver variation in the interpretation of SSEPs in anoxic-ischaemic coma. Clin Neurophysiol 2006;117:1529–35.
34. Celani MG, Carrai R, Cantisani TA, et al. Is there inter-observer variation in the interpretation of SSEPs in comatose cardiac arrest survivors? Further considerations following the Italian multicenter ProNeCa study. Resuscitation 2020;155:207–10.
35. Zhou SE, Maciel CB, Ormseth CH, Beekman R, Gilmore EJ, Greer DM. Distinct predictive values of current neuroprognostic guidelines in post-cardiac arrest patients. Resuscitation 2019;139:343–50.
36. Olson DM, Stutzman S, Saju C, Wilson M, Zhao W, Aiyagari V. Interrater reliability of pupillary assessments. Neurocrit Care 2016;24:251–7.
37. Steinberg A, Callaway CW, Arnold RM, et al. Prognostication after cardiac arrest: results of an international, multi-professional survey. Resuscitation 2019;138:190–7.
38. Sandroni C, Grippo A, Nolan JP. ERC-ESICM guidelines for prognostication after cardiac arrest: time for an update. Intensive Care Med 2020;46:1901–3.
39. Bongiovanni F, Romagnosi F, Barbella G, et al. Standardized EEG analysis to reduce the uncertainty of outcome prognostication after cardiac arrest. Intensive Care Med 2020;46:963–72.