Electroencephalographic Reactivity as Predictor of Neurological Outcome in Postanoxic Coma: A Multicenter Prospective Cohort Study

Marjolein M. Admiraal, MSc,[1] Anne-Fleur van Rootselaar, MD, PhD,[2] Jeannette Hofmeijer, MD, PhD,[3,4] Cornelia W. E. Hoedemaekers, MD, PhD,[5] Christiaan R. van Kaam, MSc,[5] Hanneke M. Keijzer, MSc,[3,6] Michel J. A. M. van Putten, MD, PhD,[4,7] Marcus J. Schultz, MD, PhD,[1,8,9] and Janneke Horn, MD, PhD[1,8]

Objective: Outcome prediction in patients after cardiac arrest (CA) is challenging. Electroencephalographic reactivity (EEG-R) might be a reliable predictor. We aimed to determine the prognostic value of EEG-R using a standardized assessment.
Methods: In a prospective cohort study, a strictly defined EEG-R assessment protocol was executed twice per day in adult patients after CA. EEG-R was classified as present or absent by 3 EEG readers, blinded to patient characteristics. Uncertain reactivity was classified as present. Primary outcome was best Cerebral Performance Category score (CPC) in 6 months after CA, dichotomized as good (CPC = 1–2) or poor (CPC = 3–5). EEG-R was considered reliable for predicting poor outcome if specificity was ≥95%. For good outcome prediction, a specificity of ≥80% was used. Added value of EEG-R was the increase in specificity when combined with EEG background, neurological examination, and somatosensory evoked potentials (SSEPs).
Results: Of 160 patients enrolled, 149 were available for analyses. Absence of EEG-R for poor outcome prediction had a specificity of 82% and a sensitivity of 73%. For good outcome prediction, specificity was 73% and sensitivity 82%. Specificity for poor outcome prediction increased from 98% to 99% when EEG-R was added to a multimodal model. For good outcome prediction, specificity increased from 70% to 89%.
Interpretation: EEG-R testing in itself is not sufficiently reliable for outcome prediction in patients after CA. For poor outcome prediction, it has no substantial added value to EEG background, neurological examination, and SSEPs. For prediction of good outcome, EEG-R seems to have added value.

ANN NEUROL 2019;86:17–27

Cardiac arrest (CA) occurs frequently in Western countries, with an estimated incidence ranging between 85 and 110 per 100,000 person-years.[1] Despite medical advances, overall survival rates remain low.[1] About half of the patients admitted to the intensive care unit (ICU) with postanoxic coma survive.[2–4] The main determinant of outcome is neurological dysfunction, predominantly resulting from hypoxic–ischemic brain injury.[5] To limit futile treatment in

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.25507

Received Dec 7, 2018, and in revised form May 22, 2019. Accepted for publication May 22, 2019.

Address correspondence to Admiraal, Department of Intensive Care, Amsterdam University Medical Centers, University of Amsterdam, Meibergdreef 9, PO Box 22660, 1100 DD Amsterdam, the Netherlands. E-mail: m.m.admiraal@amc.nl

From the [1]Amsterdam University Medical Centers, University of Amsterdam, Department of Intensive Care, Amsterdam Neuroscience, Amsterdam, the Netherlands; [2]Amsterdam University Medical Centers, University of Amsterdam, Department of Neurology/Clinical Neurophysiology, Amsterdam Neuroscience, Amsterdam, the Netherlands; [3]Rijnstate Hospital, Department of Neurology, Arnhem, the Netherlands; [4]Clinical Neurophysiology, TechMed Centre, University of Twente, Enschede, the Netherlands; [5]Radboud University Medical Center, Department of Intensive Care, Nijmegen, the Netherlands; [6]Department of Intensive Care Medicine and Neurology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, the Netherlands; [7]Department of Clinical Neurophysiology, Medisch Spectrum Twente, Enschede, the Netherlands; [8]Amsterdam University Medical Centers, University of Amsterdam, Laboratory for Experimental Intensive Care and Anesthesiology, Amsterdam, the Netherlands; and [9]Mahidol University, Mahidol Oxford Tropical Medicine Research Unit, Bangkok, Thailand

patients without a chance of neurologic recovery, and to avoid unjust withdrawal of life-sustaining treatment (WLST), several diagnostic tests have been advocated.[6] Traditional markers predicting poor outcome are absent brainstem reflexes, bilateral absence of N20 responses on somatosensory evoked potentials (SSEPs), and electroencephalography (EEG) at or after 72 hours post-CA. These tests have high specificity but limited sensitivity for prediction of poor neurological outcome, ranging between 20 and 50%.[6] In recent years, early EEG background patterns at 12 and 24 hours were reported to be reliable prognostic markers for both poor and good outcome.[4,7,8] EEG is widely available, and there is extensive experience in many hospitals.[6,9]

Besides background patterns, the EEG response to external stimulation, that is, EEG reactivity (EEG-R), can also be assessed.[6–8,10] Absence of EEG-R has been reported with a specificity ranging between 70 and 100% for prediction of poor outcome.[7,8,10–14] For good outcome, presence of EEG-R has been shown to result in a specificity between 55 and 95%.[7,8] Unfortunately, protocols for EEG-R are often poorly described and vary widely, and a clear definition for interpretation of EEG-R is not available.[15] The studies performed so far have been mainly single-center studies with limited cohort sizes. Nevertheless, EEG-R in combination with other factors is included in recommendations for neurological prognostication in comatose patients after CA.[6] Given these limitations, we aimed to determine the prognostic value of EEG-R for the prediction of outcome when using a strict and replicable protocol for EEG-R testing. We aimed to determine the prognostic value of EEG-R for poor and for good neurological outcome separately, because clinical consequences are different in both situations. In patients with a high probability of poor outcome, a test confirming this might lead to WLST. In a patient with a high probability of good outcome, substantiating this would justify more invasive therapies.

In addition, we investigated the additional prognostic value of EEG-R to other prognostic markers for neurological outcome. We also determined the reliability of the interpretation of EEG-R.

We hypothesized that EEG-R is a reliable tool for prediction of poor neurological outcome in patients after CA.

Patients and Methods

Design

This was an investigator-initiated prospective multicenter cohort study conducted in the ICUs of 2 Dutch university-based hospitals and 1 teaching hospital. The institutional review board of the Amsterdam University Medical Centers, the Netherlands, approved the study protocol and waived the need for written informed consent, as EEG-R assessment and basic follow-up are part of routine monitoring in comatose patients after CA. The trial was registered at trialregister.nl (identifier: NTR6231).

Patients

Consecutive comatose adult patients after CA in whom continuous EEG (cEEG) monitoring was started within 24 hours after CA were eligible for participation. Exclusion criteria included traumatic brain injury, acute stroke, progressive neurodegenerative disease, prearrest modified Rankin scale ≥4, or prearrest life expectancy ≤6 months based on comorbidity.

Standard Care and Monitoring

Patients were treated according to local guidelines for comatose CA patients. This included targeted temperature management (TTM), aiming at 32 or 36°C, for 24 hours at all centers. cEEG monitoring was continued for 3 days unless a patient regained consciousness. Recordings were done with at least 9 electrodes, placed according to the international 10-20 system. Additionally, a ground and reference electrode were placed in the midline. EEGs were recorded with a Viasys Nicolet (CareFusion, Middleton, WI), BrainQuick ICU (Micromed, Mogliano Veneto, Italy), or Nihon Kohden system (VCM Medical, Leusden, the Netherlands). Decisions for withdrawal of life support were based on the Dutch recommendations for prognostication in postanoxic coma.[16] These include neurological examination after clearance of sedative drugs by a neurologist, SSEP testing performed after TTM, and EEG patterns at 72 hours post-CA.[6] Biomarkers are currently not advised in the Dutch guidelines.

EEG Reactivity

For the duration of cEEG monitoring, EEG-R was assessed twice per day, morning and afternoon, according to a standardized protocol of commonly used stimuli,[15] regardless of the EEG background pattern. The protocol consisted of a fixed set of auditory (clapping hands and calling out the patient's name), visual (passive eye opening), tactile (intranasal tickling of the nasal septum with a cotton swap), and noxious stimuli (sternal rub). Each stimulus was performed for 5 seconds and applied 3 times in a row with intervals of 30 seconds. The complete stimulus protocol was executed in all patients. Stimulation was stopped if a patient showed clinical signs of arousal, such as eye opening or localizing, to any of the stimuli (independent of EEG-R findings).

All EEG-R assessments were made centrally. Three experienced EEG readers (A.-F.v.R., J.Hor., and M.M.A.), blinded to the individual patient characteristics and time interval from resuscitation, independently rated EEG-R offline using a custom-made interface in MATLAB R2012b (MathWorks, Natick, MA). EEG-R was considered present if

any of the stimuli induced a change in EEG amplitude or frequency at least twice. Artifacts due to muscle activity or other noise- and stimulus-induced rhythmic or periodic discharges were not classified as present EEG-R. Incomplete EEG-R assessments (>2 of the 15 stimuli missing not caused by arousal of the patient) and assessments obscured by artifacts were excluded from analysis.

The 3 raters also scored whether they were certain or uncertain whether EEG-R was present. In the case of uncertainty, the EEG-R was considered present in the analysis, to avoid unjust pessimistic test results leading to an increased number of false positives. When the raters disagreed on reactivity, the decision as to whether EEG-R was present was based on the majority vote. As a secondary analysis, EEG-R assessments without unanimous agreement were re-evaluated in a consensus meeting.

For analysis, 1 EEG-R assessment was selected per patient. This was the first assessment scored as reactive or, if no reactive assessment was available, the EEG-R assessment where the patient received the lowest dose of sedation. If all EEG-R assessments of a patient were excluded because of incorrect execution of the protocol, artifacts, or technical reasons, the patient was excluded from analysis and we did not include an extra patient.

EEG Background Pattern

Time windows of 5 minutes of EEG recording were selected from the cEEG at 12 and 24 hours after CA and immediately before each stimulus protocol. These were scored by the 3 raters according to validated critical care EEG criteria as defined by the American Clinical Neurophysiology Society (ACNS), separated into 3 categories (A, B, and C; Table 1).[12]

TABLE 1. Electroencephalographic Background Categories

Category	Background Patterns
A	• Suppressed background (<10μV) without discharges • Suppressed background (<10μV) with continuous periodic discharges • Burst-suppression background (with or without discharges) with suppression (<10μV) or attenuation (>10μV, but <50% of background voltage) periods constituting >50% of the recording
B	• Abundant periodic discharges (>50% of recording). • Abundant rhythmic SW (polyspike/spike/sharp and wave; >50% of recording) • Unequivocal electrographic seizure (at least 1) • Discontinuous background with suppression periods (<10μV) constituting >10% of the recording • Low-voltage background (most activity is <20 μV)
C	Absence of all features stated above

Clinical Variables

The following clinical variables were collected: age, sex, location of CA, arrest witnessed, time to return of spontaneous circulation (ROSC), initial heart rhythm, presumed etiology of CA, and the results of SSEP recordings if performed. At the time of each EEG-R assessment, information on temperature, sedative medication, and pupillary and corneal reflexes was collected.

Outcome

Primary outcome measure was the best achieved neurological outcome within 6 months after CA, defined as good (Cerebral Performance Category score [CPC] = 1–2; no or mild cerebral impairment) or poor (CPC = 3–5; moderate cerebral impairment, vegetative state, or death).[17]

Power Calculation

A power calculation was performed for the primary research question: prognostic value of absent EEG-R for prediction of poor outcome. This was based on a prevalence of absent EEG-R of 45% and an expected specificity for the prediction of poor outcome of 94%, both based on pooled existing data.[7,11,18–20] With an alpha error of 0.05, a power of 80%, and 2-tailed testing, we needed 160 patients to reach the expected specificity level.

Statistical Analysis

The prognostic values of absence of EEG-R for prediction of poor outcome and presence of EEG-R for good outcome were described as specificity, sensitivity, positive predictive value (PPV), and negative predictive value (NPV). Based on the differences in clinical implications of predicting poor or good neurological outcome, thresholds for a reliable marker for poor outcome prediction and good outcome prediction were chosen differently. EEG-R by itself was deemed a sufficiently reliable marker for poor outcome if the specificity was ≥95% with a 95% confidence interval (CI) of <5%.[6] For good outcome, on the other hand, EEG-R was deemed sufficiently reliable if the specificity was ≥80% with a 95% CI of <10%. We calculated prognostic value for both situations. CIs were calculated according to the Clopper–Pearson exact confidence interval method.[21]

As a sensitivity analysis, prognostic value of EEG-R was also assessed in several time windows post-CA: <12 hours, 12 to 24 hours, 24 to 48 hours, and >48 hours post-CA. For each window, the best available score per

patient was used for analysis and prognostic values for both poor and good outcome were calculated. Influence of non-neurological deaths on the prognostic value of EEG-R was assessed by excluding these cases and calculating prognostic value. Influence of sedative medication on EEG-R was assessed by comparing the proportion of patients receiving sedation and the median dose between reactive and unreactive patients.

The additional value of EEG-R to other available variables was assessed as the increase in prognostic value of a baseline multimodal prediction algorithm when EEG-R is added. For prediction of poor outcome, the baseline algorithm was EEG background category A (see Table 1) at 24 hours OR bilaterally absent pupillary reflexes and corneal reflexes on all tests at <72 hours OR bilaterally absent cortical N20 response on SSEP testing. For predicting good outcome, the baseline algorithm was EEG background category C (see Table 1) at 12 hours OR brainstem reflexes present at all tests at <72 hours.

The reliability of EEG-R interpretation was assessed by inter- and intrarater variability. Inter-rater variability among the 3 raters was assessed using intraclass correlation coefficient (ICC) and 95% CI based on an average-measures, absolute-agreement, 2-way random-effects model. For test–retest reliability, the 3 raters scored 30 randomly selected EEG-R protocols 3 times, mixed with other EEG-R assessments and in separate batches to omit recognition. Intrarater variability was assessed by ICC and corresponding 95% CI based on an average-measures, absolute-agreement, 2-way mixed-effects model.

Difference between outcome groups was described using chi-square, Fisher exact test, or Mann–Whitney U test, whichever was appropriate. Probability values <0.05 were considered statistically significant. Statistical analyses were performed using MATLAB 2012b. Missing data were not imputed.

Results

Patients

Patients were enrolled between April 2015 and February 2018. A total of 182 patients were considered eligible, and 160 patients were enrolled (see Fig 1 for details). At the Amsterdam University Medical Centers 129 patients were enrolled, Rijnstate Hospital Arnhem included 17 patients, and University Medical Center Nijmegen included 14 patients. In these 160 patients, 426 EEG-R assessments were executed, of which 51 were excluded because of artifacts, 10 because of technical issues, and 20 because the EEG-R assessment was incomplete. The remaining 345 EEG-R assessments in 149 patients were available for analysis. Examples of present, unsure, and absent EEG-R are shown in Figure 2.

Patient characteristics are presented in Table 2. Seventy patients (47%) had a poor outcome (9 CPC = 3, 2 CPC = 4, 59 CPC = 5), and 79 (53%) had a good

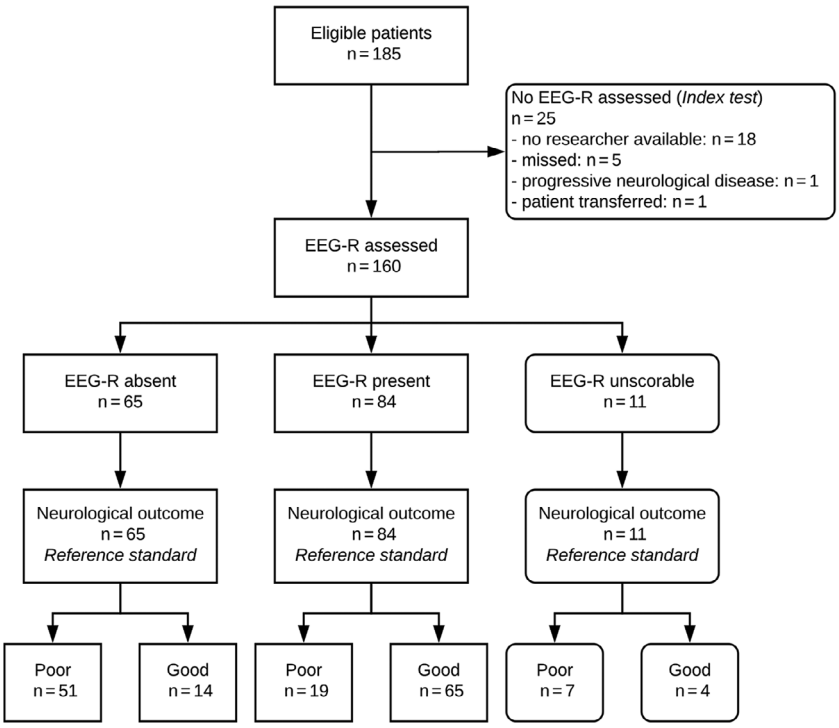

FIGURE 1: Flowchart according to STARD (Standards for Reporting Diagnostic Accuracy Studies) guidelines of available patients for inclusion and analysis. EEG-R = electroencephalographic reactivity.

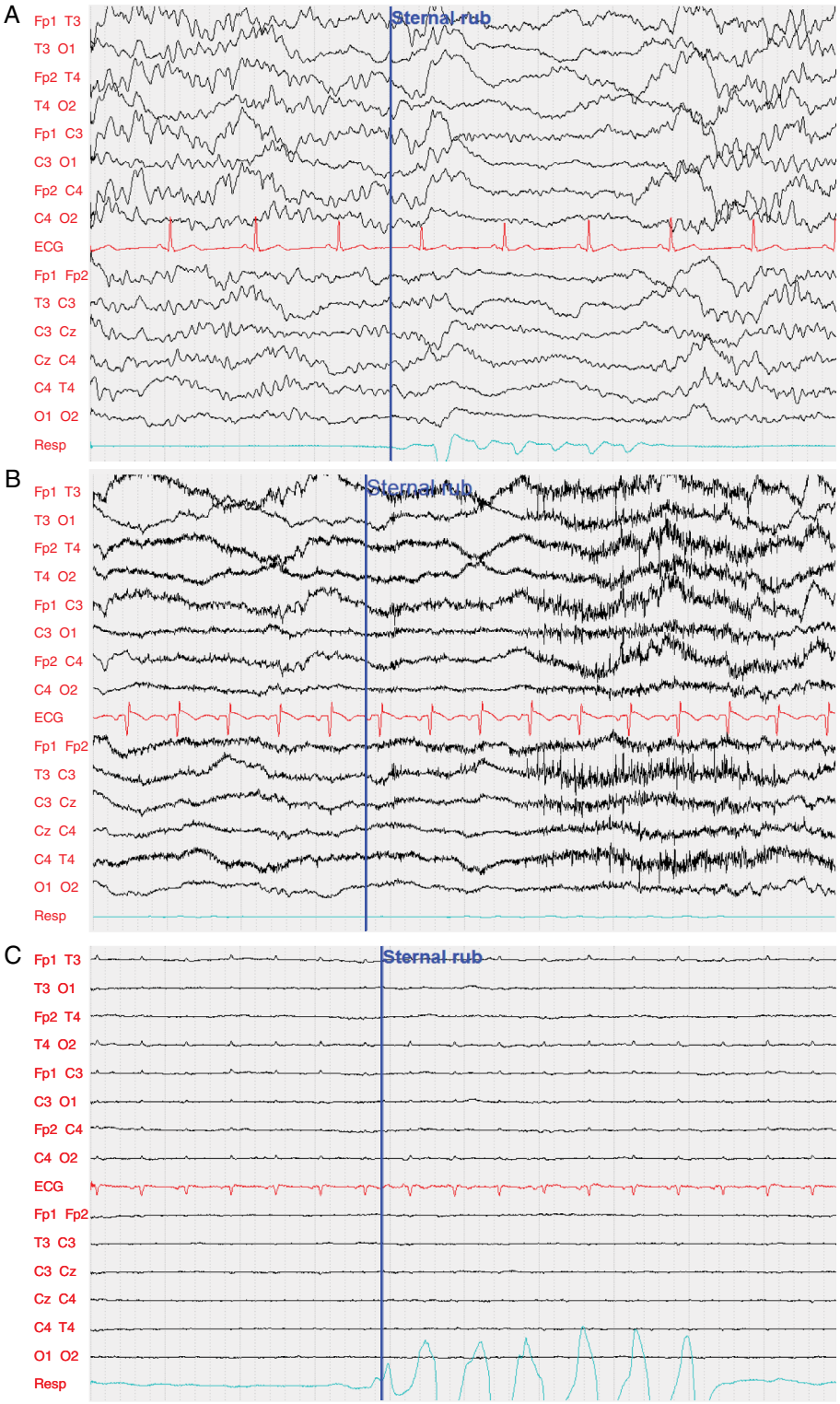

FIGURE 2: (A) Example of an electroencephalogram (EEG) with present EEG reactivity (EEG-R). The patient was stimulated with a sternal rub at the location of the marker. EEG configurations are high pass filter = 0.53Hz, low pass filter = 70Hz, scaling = 50μV/cm. (B) Example of an EEG with uncertain EEG-R. The patient was stimulated with a sternal rub at the location of the marker. EEG configurations are high pass filter = 0.53Hz, low pass filter = 70Hz, scaling = 70μV/cm. (C) Example of an EEG with absent EEG-R. The patient was stimulated with a sternal rub at the marker. EEG configurations are high pass filter = 0.53Hz, low pass filter = 70Hz, scaling = 30μV/cm. The last channel, resp represents data from a movement sensor on the chest for tracking respiratory movements; the sternal rub is clearly seen in this channel. ECG = electrocardiogram. [Color figure can be viewed at www.annalsofneurology.org]

TABLE 2. Patient Characteristics

Characteristic	Good Outcome, n = 79	Poor Outcome, n = 70	*p*
Demographic characteristics			
Age, yr	62 (51–69)	65 (53–74)	0.17
Sex, male	66/79 (84%)	51/70 (73%)	0.16
Characteristics of the cardiac arrest			
OHCA	73/79 (92%)	59/70 (84%)	0.13
Witnessed arrest	58/75 (77%)	52/70 (74%)	0.08
Time to ROSC, min	12.5 (9–17.5)[a]	22 (15–32)[b]	<0.001
Initial rhythm shockable	69/74 (93%)	36/67 (54%)	<0.001
Cardiac etiology	47/70 (67%)	45/61 (74%)	0.45
Clinical characteristics during ICU treatment			
Absent stem reflexes, <72 h	0/79 (0%)	19/70 (27%)	<0.001
Present stem reflexes, <72 h	54/79 (68%)	18/70 (26%)	<0.001
Bilaterally absent SSEP N20	0/3 (0%)	11/35 (31%)	0.54
EEG pattern at 12 h			<0.001
Category A	0/38 (0%)	20/28 (71%)	
Category B	14/38 (37%)	3/28 (11%)	
Category C	24/38 (63%)	5/28 (18%)	
EEG pattern at 24 h			<0.001
Category A	2/75 (3%)	30/66 (45%)	
Category B	10/75 (13%)	14/66 (21%)	
Category C	63/75 (84%)	22/66 (33%)	

Data are presented as median (interquartile range) or n/n (%). For EEG pattern classification, see Table 1.
[a]n = 76.
[b]n = 63.
EEG = electroencephalographic; ICU = intensive care unit; OHCA = out of hospital cardiac arrest; ROSC = return of spontaneous circulation; SSEP = median nerve somatosensory evoked potential.

outcome (59 CPC = 1 and 20 CPC = 2). Poor outcome was associated with longer time to ROSC, nonshockable initial rhythm, absent brainstem reflexes at <72 hours, and unfavorable EEG patterns at 12 and 24 hours. SSEP recordings were done in 38 patients (26%) at a median time of 56 hours post-CA (interquartile range [IQR] = 43–86 hours). WLST was performed in 46 patients (31%). Clinical signs of arousal leading to premature interruption of the protocol were observed in 4 (1%) EEG-R assessments.

Prognostic Value of EEG-R

Absence of EEG-R predicted poor outcome with a specificity of 82% (95% CI = 74–91) and a sensitivity of 73% (95% CI = 62–83; Tables 3 and 4, A). Nineteen patients had present EEG-R, but had a poor outcome; 3 had CPC = 3 at 6 months and 16 died, 6 from a non-neurological cause, 9 from a neurological cause, and 1 of unknown cause. Of those 9 patients that died from a neurological cause, SSEPs were tested in 7; in 1 patient, the N20 was bilaterally absent. In the other patients, a wait and see policy was applied and treatment was withdrawn when no neurological recovery appeared (median time from CA to death = 5 days, IQR = 3.5–11 days).

For presence of EEG-R predicting good outcome, we found 73% (95% CI = 62–83) specificity and 82% (95% CI = 74–91) sensitivity (see Table 3).

In none of the time windows after CA that we assessed did specificity for prediction of poor outcome reach our predefined limit for sufficient reliability (Table 5). After

TABLE 3. Predictive Value of (Combinations of) Clinical and Neurophysiologic Measures

Measure	Available for Analysis, n	Predicted Outcome	Specificity (95% CI)	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)
EEG reactivity absent	149	Poor	82% (72–90)	73% (61–83)	79% (67–88)	77% (67–86)
EEG cat. A pattern at 24 h	141	Poor	97% (91–100)	46% (33–58)	94% (79–99)	67% (57–76)
Brainstem reflexes absent	149	Poor	100% (95–100)	27% (17–39)	100% (82–100)	61% (52–69)
SSEP N20 absent	38	Poor	100% (29–100)	31% (17–49)	100% (72–100)	11% (2–29)
Baseline: EEG cat. A pattern at 24 h OR brainstem reflexes absent OR SSEP N20 absent	149	Poor	98% (91–100)	54% (42–66)	95% (83–99)	71% (61–79)
Baseline AND EEG reactivity absent	149	Poor	99% (93–100)	51% (39–64)	97% (86–100)	70% (60–78)
EEG reactivity present	149	Good	73% (61–83)	82% (72–90)	77% (67–86)	79% (67–88)
EEG cat. C pattern at 12 h	66	Good	82% (63–94)	63% (46–78)	83% (64–94)	62% (45–78)
Brainstem reflexes present	149	Good	74% (62–84)	68% (57–78)	75% (63–85)	68% (56–78)
Baseline: EEG cat. C pattern at 12 h OR brainstem reflexes present	149	Good	70% (58–80)	79% (68–87)	75% (64–84)	74% (62–84)
Baseline AND EEG reactivity present	149	Good	89% (79–95)	66% (54–76)	87% (75–94)	70% (59–79)

Specificity, sensitivity, PPV, and NPV data are given as percentage (95% CI). For EEG, pattern classification, see Table 1.
cat. = category; CI = confidence interval; EEG = electroencephalographic; NPV = negative predictive value; PPV = positive predictive value; SSEP = somatosensory evoked potential.

excluding non-neurological deaths (n = 13), the specificity for prediction of poor outcome was 82% (95% CI = 72–90) and the sensitivity was 77% (95% CI = 64-87; see Table 4, B). Unreactive patients less often received sedative medication (Table 6). If sedated, dose of propofol or midazolam did not differ between reactive and unreactive patients.

Additional Prognostic Value of EEG-R

Prediction of poor neurological outcome based on EEG background at 24 hours post-CA, brainstem reflexes, or SSEPs resulted in a specificity of 98% (95% CI = 91–100) and the sensitivity of 54% (95% CI = 42–66). If also EEG-R was absent, the specificity was 99% (95% CI = 93–100) and the sensitivity was 51% (95% CI = 39–64; see Table 3 for PPV and NPV).

For prediction of good neurological outcome based on EEG background at 12 hours post-CA or brainstem reflexes, the specificity was 70% (95% CI = 58–80) at 79% (95% CI = 68–87) sensitivity. When present EEG-R was added, the specificity increased to 89% (95% CI = 79–95), and the sensitivity decreased to 66% (95% CI = 54–76; see Table 3 for PPV and NPV).

Reliability of EEG-R Interpretation

Inter-rater reliability was good, with an ICC of 0.85 (95% CI = 0.82–0.88). Intrarater variability was moderate on average: rater A, ICC = 0.68 (95% CI = 0.34–0.85); rater B, ICC = 0.70 (95% CI = 0.37–0.86); rater C, ICC = 0.83 (95% CI = 0.64–0.92). In the 83 (24%) EEG-R assessments that were re-evaluated in a consensus meeting, we found poor agreement between the results of the majority vote and consensus meeting (ICC = 0.40, 95% CI = 0.09–0.61). Based on consensus meeting results, specificity and sensitivity did not change (see Table 4, C).

Discussion

In this study, we found that EEG-R in itself is not reliable enough for prediction of poor or good neurological outcome in patients after CA. In light of the multimodal approach to prognostication, the question is to what extent

TABLE 4. Contingency Table of EEG Reactivity and Outcome

	Poor Outcome, CPC 3–5	Good Outcome, CPC = 1–2
A		
Unreactive EEG	51	14
Reactive EEG	19	65
B		
Unreactive EEG	44	14
Reactive EEG	13	65
C		
Unreactive EEG	48	9
Reactive EEG	22	70

Outcome was determined by the CPC scale (CPC) and dichotomized as good (CPC = 1–2) and poor (CPC = 3–5) outcome. EEG reactivity was based on majority vote (A), based on majority vote with patients dying of a non-neurological cause excluded (B), and based on consensus meeting results (C).
CPC = Cerebral Performance Category; EEG = electroencephalographic.

does EEG-R provide additional information. EEG-R seems not to be of additional value for prediction of poor outcome; it might be of additional value for prediction of good outcome. We found moderate test–retest reliability and poor agreement between majority vote and consensus meeting results.

This is the first study to assess the prognostic value of standardized assessment of EEG-R in a prospective study, specifically designed for this purpose. The prognostic value of EEG-R has been described before, but mainly in studies with small sample sizes and with conflicting results.[7,11,18,22–28] Two studies are available that describe large cohorts, 357 and 373 patients, with reported specificities for poor outcome of 77% (95% CI = 60–88) and 99% (95% CI = 97–100), respectively.[8,10] We obtained a specificity of 82%. However, these studies were retrospective, protocols were less extensive and structured, and other stimuli were used. In contrast to our study, one study included bilateral nipple pinching.[8] This stimulus was described, albeit in a very small study, as the stimulus most often evoking EEG-R.[29] We deliberately did not include this stimulus, as it is under ethical debate.[29,30] Furthermore, one study described bedside, and therefore unblinded, scoring of EEG-R, which is the opposite of our fully blinded analysis.[8] Finally, we used best CPC within 6 months compared to 3 months or at hospital discharge. Early determination of outcome could lead to an overly pessimistic estimation of outcome, as survivors recover.

Stimulation protocols between studies vary widely, and guidelines do not state which stimuli to use.[6,15,31] In this study, stimulations were always executed in the same order, with increasing intensity of stimulation, and the full stimulus protocol was always executed. Only clinical arousal was a reason to discontinue the protocol. The raters only scored whether they considered a complete stimulation protocol reactive, regardless of the exact stimulus type. In futures studies, best practice for performing stimulations should be investigated.

TABLE 5. Predictive Value of Absence of EEG-R for Prediction of Poor Outcome at Different Time Points after Cardiac Arrest

	TP	FN	FP	TN	Total	Specificity	Sensitivity	PPV	NPV
Unreactive, <12 h	20	2	12	16	50	57% (37–76)	91% (71–99)	63% (52–72)	89% (67–97)
Unreactive, 12–24 h	39	8	16	45	108	74% (61–84)	83% (69–92)	71% (61–79)	85% (75–92)
Unreactive, 24–48 h	33	12	15	25	85	63% (46–77)	73% (58–85)	69% (59–77)	68% (55–78)
Unreactive, >48 h	12	5	1	5	23	83% (36–100)	71% (44–90)	92% (66–99)	50% (31–69)
Unreactive, all time points	51	19	14	65	149	82% (72–90)	73% (61–83)	79% (67–88)	77% (67–86)

EEG-R and outcome are binary; therefore, values for prediction of good outcome are the exact opposite. Sensitivity and specificity are given as percentage (95% confidence interval).
EEG-R = electroencephalographic reactivity; FN = false negative; FP = false positive; NPV = negative predictive value; PPV = positive predictive value; TN = true negative; TP = true positive.

TABLE 6. Sedation Received by Patients during EEG-R Testing

	Reactive, n = 84	Unreactive, n = 65	*p*
Sedated	61/84 (73%)	34/65 (52%)	0.02
Receiving propofol	57/84 (68%)	31/65 (48%)	0.02
Dose propofol, mg/kg/h	3.33 (2.72–4.00)	3.33 (2.50–3.95)	0.60
Receiving midazolam	8/84 (10%)	8/65 (12%)	0.60
Dose midazolam, mg/kg/h	0.04 (0.02–0.07)	0.1 (0.02–0.12)	0.40

Results are given as n/total (percentage of total) or median (interquartile range). For analysis, 1 EEG-R assessment was selected per patient. This was the first assessment scored as reactive or, if no reactive assessment was available, the EEG-R assessment where the patient received the lowest dose of sedation. Four of the reactive patients and 5 of the unreactive patients received both propofol and midazolam.
EEG-R = electroencephalographic reactivity.

Ideally, the specificity of a prognostic test, when used by itself, should be 100%. We decided to aim for at least 95% before denominating it reliable. Although WLST decisions will never be made based on a single predictor, knowing the prognostic value of each predictor in the multimodal prognostication process is valuable. EEG-R was not able to reach our reliability limit, and we therefore do not find it sufficiently reliable to be used by itself. For good outcome prediction, we set a slightly wider limit (at least 80% specificity), because implications are less drastic, but even this limit could not be reached. The sensitivity of EEG-R, on the other hand, is much higher than other clinical tests for prediction of outcome.[6] However, especially for poor outcome, specificity is the main determinant of a reliable marker. EEG-R by itself should therefore be used with great caution.

Although only one clinical assessment (EEG-R) is performed, we report prognostic values of both poor and good. In our opinion, knowing the predictive value of a test for both poor and good outcome facilitates a balanced decision in clinical practice.

Because prognostication after CA is often executed multimodally, we also assessed the prognostic value of EEG-R in addition to clinical examination, assessment of SSEPs, and EEG background patterns. Including EEG-R did not result in a clinically relevant increase of the specificity for prediction of outcome. Therefore, we cannot conclude that EEG-R is of additional value. For prediction of good outcome, however, specificity increased substantially, and therefore EEG-R might be of added value. The increase in specificity was at the cost of a decrease in sensitivity.

In our cohort, several patients died in the ICU from a non-neurological cause. This could have decreased our sensitivity for prediction of poor outcome, as EEG-R is primarily used for prediction of poor neurological outcome. When these patients were analyzed as having a good outcome, sensitivity and specificity did not change, and therefore this had no major influence on our results. The reliability of the visual assessment of EEG-R was previously described as questionable.[32,33] We thoroughly assessed the reliability and investigated inter-rater agreement, but also test–retest reliability and comparison between majority vote and consensus meeting results. The moderate test–retest reliability and poor agreement between majority vote and consensus meeting are not in favor of using EEG-R in prognostication after CA. Part of the problem might be the poorly described definition of EEG-R. The ACNS defined EEG-R as "Change in cerebral activity to stimulation."[31] This can be any change in amplitude and/or frequency, including attenuation of activity. The ACNS definition is open to multiple interpretations, and there are no cutoffs in terms of frequency or amplitude change for reactivity. Also, these guidelines do not state the time window in which a change in the EEG should be seen to qualify as reactive. In our study, we assumed a reactive EEG to show changes within several seconds after stimulation. We evaluated EEG-R per stimulus and did not assess the change in EEG before stimulations were started compared to after all stimuli were executed. This approach has been used before, for example, in studies by the group of Rossetti.[8] In line with the ACNS definition, we excluded muscle artifact from EEG-R. We note that a reaction to stimulation is potentially not limited to a change in the EEG or the frequency band we investigated. This should be evaluated in future studies.

The variability of the human interpretation of EEG-R could be overcome by quantitative analysis; however, no algorithm is currently available for clinical application.[13,25,32,34] A first step toward a clearer definition could be determining thresholds for changes based on EEG-R assessments where the decision between reactive and unreactive was unanimous. For this purpose, assessments scored as uncertain could be left out.

Another feature of the EEG that has been studied in the context of outcome prediction in comatose patients is spontaneous EEG variability.[35] In comatose patients with varying etiologies, typically, less spontaneous variability is associated with worse outcomes.[35–37] Possibly, spontaneous variability of the EEG, complementary to the stimulus-induced EEG-R, increases diagnostic accuracy.

Strengths and Limitations

Major strengths of this study are the prospective design, properly performed a priori power calculation, and blinding of the EEG readers to outcome, patients' clinical status, and other information such as resuscitation details. Furthermore, our follow-up was long, allowing patients to recover. Primary outcome was best CPC as opposed to CPC at a certain time point to prevent an overly pessimistic outcome estimation. We limited the influence of self-fulfilling prophecy by not notifying the treating physicians of the results of EEG-R testing. The early EEG background as measured by the cEEG was not part of our multimodal decision protocol on prognostication, nor was EEG-R. WLST was performed in a third of patients with poor outcome, but the moment of these decisions was not registered. A bilateral absence of the cortical N20 with the SSEPs resulted in WLST. If the N20 was not bilaterally absent, a wait and see policy was used.

Our study has several limitations. Not all 160 patients deemed necessary could be analyzed, as data from 11 patients could not be used. Both the expected specificity as used in our power calculation and the threshold for a reliable predictor could not have been reached if additional patients were included and 160 analyzable patients were reached.

One center was over-represented in the study. This could limit generalizability of our results. Another limitation is that video recordings were not included in the EEG registration. In almost one-quarter of the EEG-R assessments, raters were uncertain of reactivity for instance in establishing actual EEG-R or muscle or movement artifacts. Possibly, video recordings and/or administration of muscle relaxants could improve reliable assessment. According to our protocol for EEG-R testing, stimulations could be stopped when the patient shows clinical signs of arousal. This could unblind EEG readers and therefore be a potential source of bias. In our study, clinical signs of arousal led to stopping the stimulations in only 1% of the sample and is therefore unlikely to have imposed a bias on the results.

Conclusion

Our findings suggest that EEG-R by itself is not a reliable predictor for either poor or good outcome in comatose patients after CA. In addition to clinical examination, SSEPs, and EEG background pattern, we found no additional value of EEG-R. For prediction of good neurological outcome, EEG-R increased specificity substantially. However, reliable assessment of EEG-R remains an issue. Although our data do not support the use of EEG-R for poor outcome prediction in patients after CA, it may have added value for prediction of good outcome.

Author Contributions

All authors contributed to the conception and design of the study; M.M.A., A.-F.v.R., J.Hor., J.Hof., C.W.E.H., C.R.v.K., and H.M.K. contributed to data acquisition and the analysis of the data; M.M.A., A.-F.v.R., and J.Hor. contributed to drafting the text and preparing the figures.

Potential Conflicts of Interest

Nothing to report.

References

1. Berdowski J, Berg RA, Tijssen JGP, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: systematic review of 67 prospective studies. Resuscitation 2010;81:1479–1487.
2. Nielsen N, Wetterslev J, Cronberg T, et al. Targeted temperature management at 33°C versus 36°C after cardiac arrest. N Engl J Med 2013;369:2197–2206.
3. Bouwes A, Binnekade JM, Kuiper MA, et al. Prognosis of coma after therapeutic hypothermia: a prospective cohort study. Ann Neurol 2012;71:206–212.
4. Hofmeijer J, Beernink TMJ, Bosch FH, et al. Early EEG contributes to multimodal outcome prediction of postanoxic coma. Neurology 2015;85:137–143.
5. Lemiale V, Dumas F, Mongardon N, et al. Intensive care unit mortality after cardiac arrest: the relative contribution of shock and brain injury in a large cohort. Intensive Care Med 2013;39:1972–1980.
6. 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. Intensive Care Med 2014;40:1816–1831.
7. Sivaraju A, Gilmore EJ, Wira CR, et al. Prognostication of post-cardiac arrest coma: early clinical and electroencephalographic predictors of outcome. Intensive Care Med 2015;41:1264–1272.
8. Rossetti AO, Tovar Quiroga DF, Juan E, et al. Electroencephalography predicts poor and good outcomes after cardiac arrest: a two-center study. Crit Care Med 2017;45:e674–e682.
9. Rossetti AO, Rabinstein AA, Oddo M. Neurological prognostication of outcome in patients in coma after cardiac arrest. Lancet Neurol 2016;15:597–609.
10. Amorim E, Rittenberger JC, Zheng JJ, et al. Continuous EEG monitoring enhances multimodal outcome prediction in hypoxic-ischemic brain injury. Resuscitation 2016;109:121–126.
11. Rossetti AO, Oddo M, Logroscino G, Kaplan PW. Prognostication after cardiac arrest and hypothermia: a prospective study. Ann Neurol 2010;67:301–307.
12. Westhall E, Rossetti AO, van Rootselaar A-F, et al. Standardized EEG interpretation accurately predicts prognosis after cardiac arrest. Neurology 2016;86:1482–1490.

13. Duez CHV, Ebbesen MQ, Benedek K, et al. Large inter-rater variability on EEG-reactivity is improved by a novel quantitative method. Clin Neurophysiol 2018;129:724–730.

14. Liu G, Su Y, Liu Y, et al. Predicting outcome in comatose patients: the role of eeg reactivity to quantifiable electrical stimuli. Evid Based Complement Alternat Med 2016;2016:8273716.

15. Admiraal MM, Van Rootselaar AF, Horn J. Electroencephalographic reactivity testing in unconscious patients: a systematic review of methods and definitions. Eur J Neurol 2017;24:245–254.

16. Neurologie NVV, Care NVVI. Richtlijn prognose van post-anoxisch coma. 2011. Available at: https://nvic.nl/sites/nvic.nl/files/Richtlijnen%20aanmaken/Richtlijn%20Postanoxisch%20coma.pdf. Accessed June 25, 2018.

17. Booth CM, Boone RH, Tomlinson G, Detsky AS. Is this patient dead, vegetative, or severely neurologically impaired? Assessing outcome for comatose survivors of cardiac arrest. JAMA 2004;291:870–879.

18. Oddo M, Rossetti AO. Early multimodal outcome prediction after cardiac arrest in patients treated with hypothermia. Crit Care Med 2014;42:1340–1347.

19. Thenayan EAL, Savard M, Sharpe MD, et al. Electroencephalogram for prognosis after cardiac arrest. J Crit Care 2010;25:300–304.

20. Crepeau AZ, Rabinstein AA, Fugate JE, et al. Continuous EEG in therapeutic hypothermia after cardiac arrest: prognostic and clinical value. Neurology 2013;80:339–344.

21. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 1934;26:404.

22. Fantaneanu TA, Sarkis R, Avery K, et al. Delayed deterioration of EEG background rhythm post-cardiac arrest. Neurocrit Care 2016;26:411–419.

23. Tsetsou S, Novy J, Pfeiffer C, et al. Multimodal outcome prognostication after cardiac arrest and targeted temperature management: analysis at 36 °C. Neurocrit Care 2018;28:104–109.

24. Juan E, Novy J, Suys T, et al. Clinical evolution after a non-reactive hypothermic EEG following cardiac arrest. Neurocrit Care 2014;22:403–408.

25. Noirhomme Q, Lehembre R, Lugo ZDR, et al. Automated analysis of background EEG and reactivity during therapeutic hypothermia in comatose patients after cardiac arrest. Clin EEG Neurosci 2014;45:6–13.

26. Tsetsou S, Oddo M, Rossetti AO. Clinical outcome after a reactive hypothermic EEG following cardiac arrest. Neurocrit Care 2013;19:283–286.

27. Rossetti AO, Carrera E, Oddo M. Early EEG correlates of neuronal injury after brain anoxia. Neurology 2012;78:796–802.

28. Rossetti AO, Urbano LA, Delodder F, et al. Prognostic value of continuous EEG monitoring during therapeutic hypothermia after cardiac arrest. Crit Care 2010;14:R173.

29. Tsetsou S, Novy J, Oddo M, Rossetti AO. EEG reactivity to pain in comatose patients: importance of the stimulus type. Resuscitation 2015;97:34–37.

30. Admiraal MM, Van Rootselaar AF, Horn J. International consensus on EEG reactivity testing after cardiac arrest: towards standardization. Resuscitation 2018;131:36–41.

31. 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.

32. Hermans MC, Westover MB, van Putten MJAM, et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol 2015;127:571–580.

33. Westhall E, Rosén I, Rossetti AO, et al. Interrater variability of EEG interpretation in comatose cardiac arrest patients. Clin Neurophysiol 2015;126:2397–2404.

34. Liu G, Su Y, Jiang M, et al. Electroencephalography reactivity for prognostication of post-anoxic coma after cardiopulmonary resuscitation: a comparison of quantitative analysis and visual analysis. Neurosci Lett 2016;626:74–78.

35. Efthymiou E, Renzel R, Baumann CR, et al. Predictive value of EEG in postanoxic encephalopathy: a quantitative model-based approach. Resuscitation 2017;119:27–32.

36. Vespa PM, Boscardin WJ, Hovda DA, et al. Early and persistent impaired percent alpha variability on continuous electroencephalography monitoring as predictive of poor outcome after traumatic brain injury. J Neurosurg 2002;97:84–92.

37. Admiraal MM, Gilmore EJ, van Putten MJAM, et al. Disruption of brain-heart coupling in sepsis. J Clin Neurophysiol 2017;34:413–420.