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

Claudio Sandroni a,*, Fabio Cavallaro a, Clifton W. Callaway b, Sonia D’Arrigo a, Tommaso Sanna c, Michael A. Kuiper d, Matteo Biancone a, Giacomo Della Marca e, Alessio Farcomeni f, Jerry P. Nolan g

a Department of Anaesthesiology and Intensive Care, Catholic University School of Medicine, Rome, Italy
b Department of Emergency Medicine, University of Pittsburgh, United States
c Department of Cardiovascular Sciences, Catholic University School of Medicine, Rome, Italy
d Department of Intensive Care, Medical Center Leeuwarden, Leeuwarden, The Netherlands
e Department of Neurology, Catholic University School of Medicine, Rome, Italy
f Department of Public Health and Infectious Diseases, Statistics Section, Sapienza University of Rome, Italy
g Department of Anaesthesia and Intensive Care Medicine, Royal United Hospital, Bath, UK

A R T I C L E   I N F O

Article history:
Received 16 March 2013
Received in revised form 13 June 2013
Accepted 23 June 2013

Keywords:
Coma
Prognostication
Therapeutic hypothermia

A B S T R A C T

Aims and methods: To systematically review the accuracy of early (≤7 days) predictors of poor outcome, defined as death or vegetative state (Cerebral Performance Categories [CPC] 4–5) or death, vegetative state or severe disability (CPC 3–5), in comatose adult survivors from cardiac arrest (CA) treated using therapeutic hypothermia (TH). Electronic databases were searched for eligible studies. Sensitivity, specificity, and false positive rates (FPR) for each predictor were calculated. Quality of evidence (QOE) was evaluated according to the GRADE guidelines.

Results: 37 studies (2403 patients) were included. A bilaterally absent N20 SSEP wave during TH (4 studies; QOE: Moderate) or after rewarming (5 studies; QOE: Low), a nonreactive EEG background (3 studies; QOE: Low) after rewarming, a combination of absent pupillary light and corneal reflexes plus a motor response no better than extension (M ≤ 2) (1 study; QOE: Very low) after rewarming predicted CPC 3–5 with 0% FPR and narrow (<10%) 95% confidence intervals. No consistent threshold for 0% FPR could be identified for blood levels of biomarkers. In 6/8 studies on SSEP, in 1/3 studies on EEG reactivity and in the single study on clinical examination the investigated predictor was used for decisions to withdraw treatment, causing the risk of a self-fulfilling prophecy.

Conclusions: in the first 7 days after CA, a bilaterally absent N20 SSEP wave anytime, a nonreactive EEG after rewarming or a combination of absent oculocutaneous reflexes and M ≤ 2 after rewarming predicted CPC 3–5 with 0% FPR and narrow 95% CIs, but with a high risk of bias.

© 2013 Elsevier Ireland Ltd. All rights reserved.

Contents

1. Introduction .............................................................................................................................................. 1325
2. Materials and methods. .......................................................................................................................... 1326
  2.1. Eligibility criteria .................................................................................................................................... 1326
  2.1.1. Patient population ............................................................................................................................. 1326
  2.1.2. Interventions ..................................................................................................................................... 1326

* Corresponding author: Department of Anaesthesiology and Intensive Care, Catholic University School of Medicine, Largo Gemelli 8, 00168 Rome, Italy.
E-mail address: sandroni@rm.unicatt.it (C. Sandroni).

0300-9572/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.resuscitation.2013.06.020

A Spanish translated version of the abstract of this article appears as Appendix in the final online version at http://dx.doi.org/10.1016/j.resuscitation.2013.06.020.
1. Introduction

Mortality after resuscitation from cardiac arrest is high. Two thirds of initially resuscitated patients die before hospital discharge.1 Many of these deaths are due to post cardiac arrest brain dysfunction2 and in more than one-fourth of those who survive to hospital discharge brain hypoxia-ischaemia results in severe neurological impairment.3

Prediction of neurological outcome is an important component of the management of comatose resuscitated patients. In 2006, the Quality Standards Subcommittee of the American Academy of Neurology (AAN) published a review on available evidence. The document concluded that the presence of myoclonus status epilepticus on day 1, the bilateral absence of the N20 wave of somatosensory evoked potentials (SSEPs) or a blood concentration of neuron specific enolase (NSE) above 33 mcg/L at days 1–3, and absent pupillary and corneal reflexes or a motor response no better than extension (M1–2) at day 3 accurately predicted poor outcome, defined as death or unconsciousness after 1 month or unconsciousness or severe disability after 6 months. However, that review was based on evidence derived from patients not treated with therapeutic hypothermia (TH), which currently represents the standard for the treatment of comatose patients resuscitated from out-of-hospital cardiac arrest.4 Evidence showing that AAN recommendations may not apply to TH-treated patients has been accumulating in the last years. Moreover, this and previous reviews did not comply with the currently recommended guidelines for data reporting in systematic reviews and meta-analysis, such as PRISMA,5 and did not adequately address some important limitations of the included studies, such as the risk of ‘self-fulfilling prophecy’, which is a bias present in most studies on prognostication after cardiac arrest wherein the treating physicians are not blinded to the results of the outcome predictor and use it to make a decision to withdraw treatment.6,7

The aim of the current work is to perform a new systematic review that, in comparison with previous reviews: (a) incorporates any missed studies or studies published more recently; (b) implements an improved approach for the evaluation of main sources of bias and statistical heterogeneity; (c) complies with the most recognised standards for evidence evaluation and data reporting; and, finally (d) addresses prognostication both in patients who have not been treated with TH and in TH-treated patients. In the first part of this review we addressed prognostication in patients who have not been treated with TH.8 The present study deals with
the predictors of neurological outcome in resuscitated comatose patient treated with TH.

We reviewed the evidence using Grading of Recommendations Assessment, Development and Evaluation (GRADE) guidelines, in order to provide grounds for future recommendations, as part of a staged approach.

2. Materials and methods

This is a systematic review and aggregate data meta-analysis of prognostic accuracy studies. Data reporting is consistent with the recommendations included in the PRISMA statement. According to the PICOS template, the review question was formulated as follows: “In adult patients who are comatose following resuscitation from cardiac arrest and who have been treated with TH (P), does the use of predictors based on clinical examination, electrophysiology, serum biomarkers or neuroimaging (I) allow accurate prediction of poor outcome (O)?” Given the review question, the only eligible study design (S) for this review was an observational prognostic accuracy study in which a comparison (C) is made between the respective proportions of poor outcome among the patients having a positive test result and those having a negative test result.

2.1. Eligibility criteria

2.1.1. Patient population

All studies on adult (≥ 16 years) patients who were comatose following resuscitation from cardiac arrest and were treated with TH were considered for inclusion. Patients defined as unconscious, unresponsive, or having a Glasgow Coma Scale score (GCS) ≤ 8 were considered as comatose. Studies including non-comatose patients or patients in hypoxic coma from causes other than cardiac arrest (e.g., respiratory arrest, carbon monoxide intoxication, drowning, hanging) were excluded, except when a subpopulation of cardiac arrest patients could be evaluated separately.

2.1.2. Interventions

Four types of outcome predictors were assessed: clinical examination, electrophysiology, biomarkers, and imaging (see the Part 1 of this review for details). Since we are interested only in the prediction of outcome in comatose patients in the acute phase after resuscitation, we included only studies where the outcome predictor was evaluated within seven days from cardiac arrest.

2.1.3. Outcome

We included only prognostic accuracy studies in which the neurological outcome was described using the five Cerebral Performance Categories (CPC) (see ESM Table E1) or in such a manner that an equivalent CPC could be determined. In those studies, the outcome is usually dichotomized as poor or good according to a predefined CPC threshold, either CPC 4–5 vs. 1–3 or CPC 3–5 vs. CPC 1–2. We accepted both of those definitions for our review, but reported and pooled relevant results separately. When no threshold had been defined, we dichotomized outcomes as CPC 3–5 vs. 1–2.

In order to calculate both the outcome variables and their confidence intervals, we included only studies where the complete contingency table (i.e., the number of true/false negatives and positives for prediction of poor outcome) was reported, or where those variables could be calculated from reported data.

2.1.4. Study type

Only clinical prognostic accuracy studies published in English, French, German, Italian or Spanish as full-text articles on indexed journals were included. Reviews, case reports and studies published in abstract form were excluded. No publication date or publication status restrictions were imposed.

2.2. Search strategy

We searched MEDLINE via PubMed, Scopus and the Cochrane Database of Systematic Reviews using the search strings included into the ESM Table E2.

The automatic alert system of PubMed was activated to identify further studies published during the process of data extraction and analysis. The reference lists of relevant studies were scanned in order to identify other studies of interest. The last search was launched on May 15, 2013.

2.3. Study selection and data extraction

For each of the four types of outcome predictors, two authors performed a blinded and independent eligibility assessment. Disagreements between reviewers were resolved by consensus.

For each study included in the final analysis, the following data were extracted: number of included comatose patients; age and gender; location of cardiac arrest; initial cardiac rhythm; type and timing of the assessed prognostic index; definition of poor outcome; number of patients with poor outcome; number of patients with true/false positive (TP/FP) and true/false negative (TN/FN) test result.

In order to retrieve missing information, the authors of the papers were contacted whenever possible.

2.4. Statistical methods

Statistical methods are described in detail in the extended version of the manuscript (ESM Appendix 1). Briefly, for each included predictor we calculated sensitivity, specificity, false positive rate (FPR) expressed as 1-specificity, positive predictive value (PPV), expressed as TP/(TP + FP), and positive likelihood ratio (LR+), expressed as sensitivity/sensitivity/1-specificity. Pooling of the indices was made by summing up each member of the contingency table when there were two or more studies having similar time points and outcome definitions (i.e., CPC 3–5 or CPC 4–5). Confidence intervals for pooled values of proportions were calculated using the F distribution method, according to Blyth.

2.5. Evidence appraisal

Given the absence of specific GRADE recommendations on prognostic accuracy studies, we adapted the GRADE recommendations for diagnostic accuracy studies to rate the quality of evidence (QOE). Evidence evaluation was performed independently by two authors.

According to the GRADE methodology, the QOE started as high and was graded down based on the following factors: (1) limitations; (2) indirectness; (3) inconsistency; and (4) imprecision. Given the importance of the risk of self-fulfilling prophecy, limitations were graded as serious when the treating team was not blinded to the results of the predictor of poor outcome that was being studied, and very serious when the index was used for decisions of suspension of life sustaining therapies. Methodological details are provided in the extended version of this review, available as ESM, Appendix 1.
3. Results

3.1. Study selection (Fig. 1)

The initial search produced 977 records from PubMed, 392 records from Scopus and 11 records from the Cochrane Database of Systematic Reviews. Thirty-seven additional records were identified through forward search. After duplicate removal and abstract screening, 200 studies were considered for full-text analysis. Among them, 163 were excluded because they did not fulfill inclusion criteria. The remaining 37 studies were included in our review. Excluded studies with reasons for their exclusion are listed in the ESM Table E3.

3.2. Study characteristics

The characteristics of the 37 included studies (total 2403 patients) are summarised in Table 1. In 11/37 studies (30%) poor outcome was defined as vegetative state or death (CPC 4–5), while in the remaining 26/37 studies (70%) poor outcome was defined as severe disability, vegetative state or death (CPC 3–5). Four studies15–18 were based on clinical examination, fourteen19–32 on electrophysiology, five33–37 on biomarkers, two38,39 on neuroimaging, and twelve40–51 on multiple predictors.

Sensitivity, FPR (1-specificity), and the quality of evidence for indices based on clinical examination, electrophysiology, biomarkers and neuroimaging are reported in Tables 2a–2d. Evidence profiles are reported in ESM Tables E4a–d, while results of individual studies are reported in Tables E5a–d.

Predictors having a 0% FPR (100% specificity) and narrow 95% CIs (upper limit of 95% CI < 10%) are reported in Table 3.

3.3. Clinical examination (Table 2a)

3.3.1. Brainstem reflexes

Absence of pupillary light reflex (PLR) on hospital admission17,42 was an inaccurate predictor of poor outcome (FPR 32%; 95% CI 19–48). Conversely, at 72 or later absence of PLR was associated with an almost invariably poor outcome (see Table 2a). On a total of 41 patients with absent PLR after rewarming described in five studies15,41,42,44,49 (QOE: from Moderate to Very low) only one recovered.41

Absent corneal reflex (CR) at 72 h from cardiac arrest or later was described in three studies (QOE: Very low) and it was not
consistently associated with poor outcome. FPR ranged from 0% to 5%.\textsuperscript{11,46,49}

3.3.2. Motor response

In six studies,\textsuperscript{40,41,44,47–49} presence of extensor or absent motor response to pain (M ≤ 2) after rewarming was still compatible with good outcome (FPR 11–12%). In one study\textsuperscript{40} the coexistence of M ≤ 2 plus absent PLR and CR at 72 h from cardiac arrest predicted poor outcome with 0% [0–8] FPR. However, this finding was also used as a criterion for treatment withdrawal (QOE: Very low).

3.3.3. Myoclonus

In one cohort study\textsuperscript{40} and in a case series,\textsuperscript{15} all patients with status myoclonus (defined as spontaneous, repetitive, unrelenting, generalised multifocal myoclonus) died or remained vegetative. Timing of clinical examination in those studies ranged from 1 to 144 h after cardiac arrest. However, in six studies\textsuperscript{50,56,43,46–48} presence of myoclonus or status myoclonus within 72 h from cardiac arrest did not exclude a good (CPC 1–2) neurological outcome (FPR 5% [3–9]). Distribution and EEG patterns of myoclonus varied between studies (see ESM Table E6 for details).

3.4. Electrophysiology (Table 2b)

3.4.1. Burst-suppression

Presence of burst-suppression on EEG recorded during TH was not associated with an invariably poor outcome (FPR 5% [1–14]; QOE very low). Conversely, in a single study\textsuperscript{27} no patient with a burst-suppression pattern on continuous amplitude-integrated EEG (aEEG) recorded immediately after rewarming recovered consciousness. In another study,\textsuperscript{20} presence of burst-suppression after rewarming at 48 h from cardiac arrest was 100% specific of poor outcome. Definition of burst-suppression was very inconsistent among studies (see ESM Table E7A).
3.4.2. Seizures and status epilepticus

In three studies, presence of electrographic seizures during TH or after rewarming predicted poor outcome (CPC 3–5) with 0% FPR. However, in a recent study, presence of prolonged seizures, i.e. status epilepticus (SE), on continuous or intermittent EEG monitored for a median of 48 h after cardiac arrest was followed by neurological recovery (CPC 1–2) in two patients. In both of those patients EEG reactivity was maintained. In another study, an electrographic status epilepticus (ESE) evolving from burst-suppression (SB-ESE) was associated with poor outcome in 100% of cases, but an ESE evolving from a continuous aEEG background (C-ESE) was not (FPR 4% [0–121]). Definitions of status epilepticus were inconsistent among studies (see ESM Table E7C).

3.4.3. Flat or low-amplitude EEG

In one study, a flat or low-amplitude (<20 µV) EEG during TH at 24 h from cardiac arrest predicted poor outcome (CPC 3–5) with 0% [0–11] FPR. In another study, however, a flat (<10 µV) aEEG recorded during TH at a median of 8 h from cardiac arrest or immediately after rewarming was often followed by recovery of awareness (FPR 46% [32–59] and 5% [1–15] respectively). In two studies, a BIS (bispectral index) value of 6 or less during TH, corresponding to a flat or low-amplitude EEG, predicted a CPC 3–5 with 0% [0–6] FPR (QOE: Low). However, another recent study did not confirm this result.

3.4.4. Nonreactive EEG

In three studies, two of which from the same group, a nonreactive EEG background after rewarming accurately predicted CPC 3–5 (FPR 0% [0–3]; QOE: Low). However, during TH this pattern was still compatible with neurological recovery (FPR 3% [0–11]). Furthermore, in a large cohort study, three patients with posthypoxic myoclonus and no EEG reactivity within 72 h from cardiac arrest had a good outcome.

3.4.5. EEG grading

Unlike the findings in the first part of our review only one study described the application of an EEG grading in patients treated with TH. In this study, a Grade 3 EEG, defined as the presence of low voltage, burst suppression, status epilepticus, seizures, non-reactive, or alpha/theta coma, accurately predicted a poor outcome (CPC 3–5; FPR 0% [0–9]) when recorded after rewarming, while it showed a 6% [1–20] FPR when recorded during TH.
Table 2b  
GRADE summary of findings for predictors based on electrophysiology.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Index</th>
<th>Reference</th>
<th>Sensitivity % [95% CI]</th>
<th>FPR % [95% CI]</th>
<th>No. of patients</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPC 4–5 vs. 1–3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CPC 3–5 vs. 1–2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During TH induction</td>
<td>BAEP Absence of wave V</td>
<td>Sakurai, 2006[^38]</td>
<td>56 [31–78]</td>
<td>0 [0–31]</td>
<td>26</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cloostermans, 2012[^20]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bouwes, 2012[^15]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rossetti, 2010[^36]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rossetti, 2012[^27]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BAEP: brainstem auditory evoked potentials; BIS: bispectral index; BSR: burst suppression ratio; C-ESE: electrographic status epilepticus from continuous pattern; CI: confidence intervals; CPC: Cerebral Performance Categories; EEG: electroencephalogram; FPR: false positive rate; MLCEP: middle-latency cortical somatosensory evoked potentials; RW: rewarming; SB-ESE: electrographic status epilepticus from burst-suppression; SR: suppression ratio; SSEP: somatosensory evoked potentials; TH: therapeutic hypothermia.

[^a]: Initial EEG during TH, recorded at a median of 8 h from cardiac arrest.
[^b]: Recorded at 24 h from cardiac arrest. Low voltage was defined as EEG activity below 20 µV.
[^c]: It includes low voltage, burst-suppression, seizures, status epilepticus, nonreactive, or alpha/theta coma.
[^d]: It includes electrographic seizures or periodic epileptiform discharges (PEDs) or burst-suppression with sharp bursts.
[^e]: The only patient with burst suppression and good outcome had this sign recorded in the induction phase of TH.
[^f]: The two patients that had a good outcome had non-reactive background during TH.
[^g]: The lowest peak of BIS was detected at a median time of 5 [4–14.5] h from cardiac arrest.
3.4.6. SSEP

In four studies,\textsuperscript{19,20,31,41} two of which from the same group, a bilateral absence of N20 SSEP wave during TH accurately predicted a CPC 3–5 (FPR 0% [0–2]) QOE: Moderate. Absence of N20 SSEP after rewarming was also invariably associated with poor outcome, either defined as CPC 3–5\textsuperscript{13,32,40,47,48} or CPC 4–5,\textsuperscript{21,42,44,49} except in one study,\textsuperscript{23} where one case of full neurological recovery with reappearance of the N20 SSEP wave was observed.

Table 2c
GRACE summary of findings for predictors based on biomarkers.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Index</th>
<th>Cutoff (µg L\textsuperscript{-1})</th>
<th>Reference</th>
<th>Sensitivity % [95% CI]</th>
<th>FPR % [95% CI]</th>
<th>No. of patients</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC 4–5 vs. 1–3</td>
<td>At 48 h</td>
<td>NSE</td>
<td>33.0</td>
<td>Cronberg, 2011\textsuperscript{44}</td>
<td>61 [41–79]</td>
<td>0 [0–39]</td>
<td>34</td>
</tr>
<tr>
<td>CPC 3–5 vs. 1–2</td>
<td>At ROSC</td>
<td>NSE</td>
<td>41.5</td>
<td>Kim, 2012\textsuperscript{25}</td>
<td>32 [15–53]</td>
<td>0 [0–29]</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>At 24 h</td>
<td>NSE</td>
<td>31.2</td>
<td>Tainen, 2003\textsuperscript{37}</td>
<td>20 [3–56]</td>
<td>4 [0–20]</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>41.0</td>
<td>Oksanen, 2009\textsuperscript{34}</td>
<td>18 [8–34]</td>
<td>4 [0–14]</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>49.6</td>
<td>Kim, 2012\textsuperscript{25}</td>
<td>80 [59–93]</td>
<td>0 [0–29]</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>52.4</td>
<td>Wennervirta, 2009\textsuperscript{51}</td>
<td>10 [0–45]</td>
<td>0 [0–14]</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>S-100B</td>
<td>0.18–0.21</td>
<td>Mortberg, 2011\textsuperscript{33}</td>
<td>65 [44–83]</td>
<td>0 [0–7]</td>
<td>66</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td>At 48 h</td>
<td>NSE</td>
<td>4.97</td>
<td>Mortberg, 2011\textsuperscript{33}</td>
<td>50 [25–75]</td>
<td>7 [0–32]</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.0</td>
<td>Tainen, 2003\textsuperscript{37}</td>
<td>22 [3–60]</td>
<td>0 [0–12]</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>44.3</td>
<td>Kim, 2012\textsuperscript{25}</td>
<td>86 [64–97]</td>
<td>0 [0–34]</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>S-100B</td>
<td>81.8</td>
<td>Bouwes, 2012\textsuperscript{21}</td>
<td>18 [13–25]</td>
<td>0 [0–2]</td>
<td>310</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
<td>Tainen, 2003\textsuperscript{37}</td>
<td>22 [3–60]</td>
<td>4 [0–20]</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.30</td>
<td>Stammet, 2013\textsuperscript{50}</td>
<td>21 [9–38]</td>
<td>0 [0–7]</td>
<td>75</td>
</tr>
<tr>
<td>At ≤72 h</td>
<td>NSE</td>
<td>33.0</td>
<td>Samaniego, 2011\textsuperscript{49}</td>
<td>75 [53–90]</td>
<td>22 [6–48]</td>
<td>42</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57.2</td>
<td>Storm, 2012\textsuperscript{26}</td>
<td>46 [26–67]</td>
<td>0 [0–28]</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>78.9</td>
<td>Steffen, 2010\textsuperscript{41}</td>
<td>48 [32–63]</td>
<td>0 [0–6]</td>
<td>97</td>
</tr>
</tbody>
</table>

Abbreviations: CI: confidence intervals; CPC: Cerebral Performance Categories; FPR: false positive rates; NSE: neuron-specific enolase; ROSC: recovery of spontaneous circulation.

Table 2d
GRACE summary of findings for predictors based on imaging.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Index</th>
<th>Reference</th>
<th>Sensitivity % [95% CI]</th>
<th>FPR % [95% CI]</th>
<th>No. of patients</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC 4–5 vs. 1–3</td>
<td>Median 80 h (IQR 55–117)</td>
<td>MRI DWI or FLAIR</td>
<td>Extensive cortical lesion pattern Abnormalities in basal ganglia Abnormalities in brainstem</td>
<td>Mlynash, 2010\textsuperscript{14}</td>
<td>90 [55–100]</td>
<td>9 [0–41]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mlynash, 2010\textsuperscript{14}</td>
<td>80 [44–97]</td>
<td>9 [0–41]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mlynash, 2010\textsuperscript{14}</td>
<td>30 [7–65]</td>
<td>0 [0–24]</td>
</tr>
<tr>
<td></td>
<td>At 49–108 h</td>
<td>MRI DWI (ADC)</td>
<td>ADC &lt;650 x 10\textsuperscript{-6} mm\textsuperscript{2} s in &gt;10% of brain volume</td>
<td>Wijman, 2009\textsuperscript{59}</td>
<td>77 [46–95]</td>
<td>0 [0–28]</td>
</tr>
<tr>
<td>CPC 3–5 vs. 1–2</td>
<td>On admission</td>
<td>CT</td>
<td>Lost grey/white matter interface (CT)</td>
<td>Choi, 2012\textsuperscript{42}</td>
<td>100 [55–100]</td>
<td>0 [0–63]</td>
</tr>
<tr>
<td></td>
<td>Median 46 h (IQR 37–52)</td>
<td>MRI</td>
<td>ADC occipital cortex &lt;616 x 10\textsuperscript{-6} mm\textsuperscript{2} s</td>
<td>Kim, 2012\textsuperscript{45}</td>
<td>91 [75–98]</td>
<td>0 [0–24]</td>
</tr>
<tr>
<td></td>
<td>Median 74 h (IQR 61–86)</td>
<td>MRI DWI</td>
<td>Abnormalities in both cortex and basal ganglia</td>
<td>Cronberg, 2011\textsuperscript{44}</td>
<td>58 [33–80]</td>
<td>0 [0–63]</td>
</tr>
<tr>
<td></td>
<td>At &lt;5 days</td>
<td>MRI DWI</td>
<td>Abnormalities in both cortex and basal ganglia</td>
<td>Choi, 2012\textsuperscript{42}</td>
<td>100 [55–100]</td>
<td>0 [0–63]</td>
</tr>
</tbody>
</table>

Abbreviations: ADC: apparent diffusion coefficient; CI: confidence intervals; CPC: Cerebral Performance Categories; CT: computed tomography; DWI: diffusion weighted imaging; FPR: false positive rates; IQR: interquartile range; MRI: magnetic resonance imaging.
Table 3

<table>
<thead>
<tr>
<th>Timing</th>
<th>Index</th>
<th>Sensitivity % [95% CI]</th>
<th>FPR % [95% CI]</th>
<th>LR+ [95% CI]</th>
<th>No. of patients with positive test</th>
<th>No. of studies</th>
<th>Used for WLST</th>
<th>Quality of evidence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC 4–5 vs. 1–3</td>
<td>During TH</td>
<td>37 [22–54]</td>
<td>0 [0–5]</td>
<td>42 [3–678]</td>
<td>14</td>
<td>1</td>
<td>No</td>
<td>Low</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>After RW</td>
<td>18 [8–34]</td>
<td>0 [0–5]</td>
<td>22 [1–375]</td>
<td>7</td>
<td>1</td>
<td>No</td>
<td>Low</td>
<td>27</td>
</tr>
</tbody>
</table>

CPC 3–5 vs. 1–2

<table>
<thead>
<tr>
<th>Timing</th>
<th>Index</th>
<th>Sensitivity % [95% CI]</th>
<th>FPR % [95% CI]</th>
<th>LR+ [95% CI]</th>
<th>No. of patients with positive test</th>
<th>No. of studies</th>
<th>Used for WLST</th>
<th>Quality of evidence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>During TH (at 24 h)</td>
<td>NSE ≥ 81.8 µg/L</td>
<td>18 [13–25]</td>
<td>0 [0–2]</td>
<td>56 [3–909]</td>
<td>29</td>
<td>1</td>
<td>No</td>
<td>Moderate</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>NSE ≥ 4.3 µg/L</td>
<td>21 [9–38]</td>
<td>0 [0–7]</td>
<td>18 [1–304]</td>
<td>7</td>
<td>1</td>
<td>N/A</td>
<td>Very low</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 [7–26]</td>
<td>0 [0–8]</td>
<td>11 [1–190]</td>
<td>10</td>
<td>1</td>
<td>Yes</td>
<td>Very low</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>62 [53–70]</td>
<td>0 [0–3]</td>
<td>33 [7–163]</td>
<td>76</td>
<td>3</td>
<td>Yes (1/3)</td>
<td>Low</td>
<td>43,47,48</td>
</tr>
</tbody>
</table>


* The pattern was present during TH and immediately after regaining normothermia.

† In studies where poor outcome was defined as CPC 4–5, FPR of bilaterally absent N20 SSEP after RW was 1% [0–7] because one patient had a good outcome.

‡ The number of independent studies documenting N20 SSEP was 8 (1 of 9 studies is reported twice because it included both SSEPs during TH and after rewarming).

3.5. Biomarkers (Table 2c)

Only one small study44 investigated on the predictive value of biomarkers for poor outcome defined as CPC 4–5. In this study, a NSE above 33 mcg/L at 48 h predicted poor outcome with 0% FPR (95% CI 0–39).

In papers where poor outcome was defined as CPC 3–5, the threshold for 0% FPR varied between 49.6 mcg/mL and 52.4 mcg/mL at 24 h from cardiac arrest.45,51 between 25 mcg/mL and 81.8 mcg/mL at 48 h.45,46,47,51,53,54 and between 57.2 mcg/mL and 78.9 mcg/mL at 72 h.53,54

In two papers33,37 a S–100B value of 0.18 and 0.21 mcg/mL at 24 h predicted poor outcome (CPC 3–5) in 100% of cases, but not at 48 h, when the threshold for 0% FPR was 0.3 mcg/mL.50

3.6. Imaging (Table 2d)

In a single study42 including 20 patients, eight of whom underwent a brain CT scan at 1 h after cardiac arrest, presence of diffuse brain swelling with loss of grey/white interface predicted poor outcome (CPC 3–5) with 0% [0–63] FPR. This study, however, has a limited generalizability, since it included only patients with cardiac arrest caused by drowning.

All studies on MRI were carried out after rewarming. Presence of diffuse ischaemic lesions involving both cortex and deep grey matter nuclei,42,44 detected using diffusion weighted imaging (DWI) MRI within five days from CA predicted poor outcome (CPC 3–5) with 0% FPR, while the presence of ischaemic lesions in either of these structures at a median of 80 h from CA did not.38 The presence of isolated lesions in the brainstem also predicted CPC 3–5 with 0% FPR, but only three patients with this sign were described38 (95% CI 0–24).

Quantitative evaluation of diffusion with MRI using ADC (absolute diffusion coefficient) was described in two studies.39,45 In one study,39 the presence of more than 10% of brain volume with an ADC <650 × 10⁻⁶ mm²/s at 49–108 h from cardiac arrest predicted death or persistent vegetative state with 0% [0–28] FPR. In the other study,45 the area with highest discriminatory power was identified in the occipital cortex, where an ADC <616 × 10⁻⁶ mm²/s at a median of 45.8 h from cardiac arrest predicted death, vegetative state or severe neurological disability with 0% [0–24] FPR (see Table 2d).

3.7. Predictors with highest specificity and narrow CIs (Table 3)

Table 3 includes the list of predictors with 0% FPR and upper 95% CI limit below 10%. For prediction of poor outcome defined as CPC 4–5, the predictors fulfilling these characteristics were the presence of burst-suppression after rewarming (QOE: Low) and an SB-ES during TH and after rewarming (QOE: Low). Both these results came from a single study on aEEG and the number of patients showing this sign was very small.

For prediction of poor outcome defined as CPC 3–5, the list of predictors included (1) a bilaterally absent N20 SSEP wave during TH (4 studies; QOE: moderate) or after rewarming (5 studies; QOE: Low); (2) a nonreactive background on EEG after rewarming (3 studies; QOE: Low); (3) a combination of absent PLR and CR plus a motor score 1–2 after rewarming (1 study; QOE: Very low); and (4) a S–100B > 0.18–0.21 mcg/mL at 24 h, a NSE > 81.8 mcg/mL or a S–100B ≥ 0.3 at 48 h, and an NSE > 78.9 mcg/mL at 72 h from cardiac arrest (5 studies; QOE: from moderate to very low). The presence of electrographic seizures during TH was also associated to a CPC 3–5 with 0% FPR but it was not included in this list, since the presence of status epilepticus during TH did not predict an invariably poor outcome.

In 6/14 studies listed in Table 3 the investigated predictor was used as a criterion for decisions regarding withdrawal of life-sustaining treatment (WLST).
4. Discussion

4.1. Clinical examination

4.1.1. Brainstem reflexes and motor response

In almost all studies included in our review, predictors based on clinical examination were recorded either on admission or at 36 h or more from cardiac arrest, after the end of TH. Use of clinical examination in patients undergoing therapeutic hypothermia is hampered by the interference from hypothermia itself and from sedatives or muscle relaxants used to maintain it. However, even when performed after TH, clinical examination could have been still affected by sedation, firstly because analgesia/sedation is often continued after rewarming for various reasons (e.g., to facilitate mechanical ventilation or to treat seizures or myoclonus), and secondly because even after stopping sedative drugs their residual effect may persist, because hypothermia reduces drug clearance.

In our review, absence of PLR after rewarming was the most accurate predictor among those based on clinical examination. Among 41 patients showing this sign, only one recovered. Moreover, in two studies not included in our review, absent PLR was invariably associated with death or poor neurological outcome.

CR and motor response were less reliable predictors than PLR in our review. This may partly be because both of those signs are likely to be affected by the residual effects of neuromuscular blocking drugs unlike the PLR, which involves the non–striated ciliary muscle as the effector. In one study included in our review, an absent CR and a M ≤ 2 at 72 h or later showed a 5% and 11% FPR, respectively, in patients rewarmed from TH who had received sedative drugs ≤ 12 h before neurological examination, while the FPR was 0% in patients who did not receive sedation. This was not observed for an absent PLR, which in that study predicted poor outcome with 0% FPR in both sedated and non-sedated patients. However, even in patients not treated using TH, the predictive value of CR and motor response was lower than that of PLR.

4.1.2. Myoclonus

Clinical and electrophysiological characteristics of myoclonus varied widely in studies included in our review (see Table E6 for details). Some studies described ‘status myoclonus’ or ‘myoclonus status epilepticus’ (MSE) defined as a spontaneous repetitive unrelenting and generalised myoclonus. Another study defined MSE as the presence of more than 30 min of myoclonic jerks time locked with either bursts in a burst suppression pattern or generalised periodic epileptiform discharges on EEG. Others reported multifocal prolonged myoclonus with no specific definition but similar to the previously described status myoclonus or MSE. The associated EEG activity, when reported, was mainly epileptiform.

According to recent recommendations, generalised myoclonus appearing during the first 24 h after cardiac arrest is usually, but not always, a sign of unfavourable prognosis in patients treated with TH. Indeed, in studies performed in patients not treated with TH presence of myoclonus (mostly generalised and multifocal) in the first 24–48 h was associated almost invariably with a poor outcome, except in a few case reports. In TH-treated patients detection of myoclonus during the first 24 h is difficult, due to the combined presence of neuromuscular blocking and sedation. In the present review only one study focused on myoclonic status epilepticus occurring within the first 24 h. In that study MSE was not consistently generalised and all patients showing that sign had a poor outcome.

No single specific feature of myoclonus was consistently predictive of poor outcome in this review (see Table E6). In particular, multifocal distribution, prolonged duration or absence of EEG reactivity were still compatible with neurological recovery, although the outcome was generally worse with status myoclonus.

4.2. Electrophysiology

Using EEG as a predictor in resuscitated comatose patients may have some limitations. Like clinical examination, the EEG is prone to interference from both sedation and hypothermia itself in patients treated with TH after cardiac arrest. Moreover, interpreting the EEG after resuscitation requires the analysis of a large and continuous flow of data during several hours. The techniques of EEG data acquisition used by the authors of studies in our review varied (see also Table E7A). Some have automatically sampled 5 min of EEG every hour for the first 24 h, others analysed a mix of continuous and discontinuous (30-min) EEG, others analysed the median values of EEG-derived variables over an entire 24-h period. Inevitably, sampling and use of signal processing (e.g., BIS) implies a certain loss of information. Finally, the EEG pattern immediately after resuscitation is not stable, but it shows a complex combination of various patterns whose evolution is not entirely known. For this reason, the predictive value of EEG can be influenced by timing of recording.

4.2.1. Burst suppression

Similarly to that observed in patients not treated with TH, the presence of a burst suppression pattern on EEG was not invariably predictive of a poor outcome. In patients with favourable outcome, burst suppression may occur during TH as a transient pattern, which usually disappears shortly after rewarming. This evolution from burst suppression towards a continuous pattern during neurological recovery early after the arrest was already described in the pre-hypothermia era.

The definition of burst-suppression was inconsistent among studies included in our review (see ESM Table E7A for details) and it may have influenced the relevant observed predictive value of this index. In fact, the study which used the most restrictive definition had 37% sensitivity and 0% FPR, another with a less restrictive definition had 55% sensitivity and 4% FPR, and the one with the least restrictive definition had 79% sensitivity and 7% FPR. A more definitive evaluation of the predictive value of burst suppression will need a consistent definition of discontinuous EEG backgrounds. A recent consensus document from the American Clinical Neurophysiology Society could help to standardise the EEG terminology in critical care patients.

4.2.2. Low-amplitude or flat EEG

During TH, presence of a flat or low-amplitude EEG (corresponding either to a flat amplitude-integrated EEG or to a BIS ≤ 630) recorded during TH was not 100% accurate in predicting poor outcome. Similarly to burst suppression pattern, low EEG amplitude can be observed in the first hours after resuscitation in patients with subsequent EEG and clinical recovery. In TH-treated patients, the combined effect of hypothermia and profound sedation can contribute to this reversible EEG depression.

At normothermia, a flat EEG was associated with 0% FPR in one study, while in another study three patients having this sign subsequently regained consciousness (FPR 5% [1–15]). However, all these three false positives had a flat EEG recorded immediately after rewarming and stopping of sedative drugs, whose effects may have still been present at the time of recording.

In summary, the presence of a flat or low-amplitude EEG during TH or after rewarming is not consistently associated to a poor outcome. Its predictive value may be affected by factors like timing of recording and interference from sedatives and body temperature.
4.2.3. Epileptiform activity and status epilepticus

Epileptiform activity is common following cardiac arrest. Moreover, the simplest manifestation is presence of epileptiform discharges (EDs) which consist of spikes, polyspikes or sharp waves followed or not by a slow wave. EDs may occur independently and randomly or periodically (PEDs or PDS). Electrographic seizures (ESz) consist of EDs occurring repetitively and continuously for at least 10 s. The ESM Table E7B provides a comparison of ESz definitions used in papers included in this review.

A prolonged (>30 min) continuous or recurrent series of electrographic seizures is generally electrographic status epilepticus (ESE), although definitions vary among studies in terms of minimal duration and type of electrical activity (see ESM Table E7C for details). Despite those differences, presence of status epilepticus on EEG recorded either during TH or after rewarming was almost invariably associated to poor outcome in our review. An important exception was an ESE evolving from a continuous EEG background (C-ESE). In a single study, two of ten patients with C-ESE recovered consciousness (CPC 2–3) as opposed to none of sixteen patients with ESE evolving from burst-suppression (SB-ESE).

In a case series from Rossetti et al. not included in this review, six comatose patients treated with TH and showing status epileptics from 2 to 9 days after resuscitation from cardiac arrest regained consciousness, and four had a CPC 1–2 at six months. In that paper, status epilepticus was defined as rhythmic focal or generalised EDs, or periodic or rhythmic evolving waves lasting >5 min. Unlike other patients within their original cohorts, all of these patients had intact brain stem reflexes and showed EEG reactivity. All these results suggest that status epilepticus in postanoxic patients is a heterogeneous condition which may include different variants with possibly different prognosis, and that other clinical and EEG signs should be evaluated to improve its interpretation.

Although there is a clear association between seizures or status epilepticus and poor neurological outcome in resuscitated comatose patients, it is not clear whether epileptiform activity in those patients is just a marker of irreversible postanoxic brain injury or it contributes to poor outcome by causing direct or indirect neuronal damage. In this last case, seizure detection and treatment may have a potential additional clinical benefit.

4.2.4. EEG reactivity

EEG reactivity is tested by assessing reproducible changes in amplitude or frequency of EEG background following external stimuli, such as tactile or nociceptive stimulation, auditory stimuli (clapping, voice sounds) or eye opening. Absence of EEG reactivity both during TH and after rewarming predicted poor outcome with 100% specificity in two studies from the same group. In another study, an unreactive EEG recorded within 14 days after cardiac arrest in 13 resuscitated patients was associated with an invariably poor outcome. However, a recent cohort study did not confirm these results for patients undergoing TH. Moreover, in a case series three resuscitated comatose patients with absent EEG reactivity and clinical myoclonus had a good outcome. Evidence about EEG reactivity in postanoxic coma is still limited and warrants further investigation. Limitations of this index include being both operator dependent and non-quantitative, and lacking standardisation. The criteria for eliciting and recording EEG reactivity have been outlined in a recent consensus document.

4.2.5. EEG grading

Only one study reported EEG grading for prognostication. The classification that study adopted differed from the general scheme of the classical EEG grading systems (see ESM Table E8–Comparison of EEG Grading systems in Part 1 of this review). The EEG patterns classified as Grade 3 in that study showed a 0% FPR after rewarming. Those patterns were described separately in other studies included in this review where, in at least one case, they did not confirm 0% FPR. Moreover, one of those EEG patterns, alpha-theta coma, did not demonstrate 100% specificity in patients not treated using normothermia (see Table 2a in the first part of this review).

4.2.6. N20 SSEP wave

The bilateral absence of N20 wave of short-latency somatosensory evoked potentials was among the most accurate predictors of poor outcome in patients treated with TH. In studies included in our review, no patient whose N20 SSEP wave was bilaterally absent during TH recovered, and among 538 patients studied with SSEPs after rewarming there was only one case of false positive result. SSEPs were early predictors of poor outcome, a feature demonstrated in patients not treated with TH as well. In comatose patients treated with TH, SSEPs compare favourably with clinical examination and EEG, being resistant to the effects of both sedative drugs and mild hypothermia. However, in the ICU environment evoked potentials are prone to electrical interference coming from muscular artefacts and electrical equipment, which is the most important cause of interobserver variation in the interpretation of the median nerve SSEP in post-anoxic comatose patients.

For this reason, utmost care should be taken to maximise signal/noise ratio when recording SSEPs in these patients. When noise makes unequivocal detection of N20 impossible, the most correct classification for this result should be ‘indeterminate’ rather than ‘absent N20’.

Like other predictors, SSEPs are also prone to the risk of ‘self-fulfilling prophecy’. In our review, among all predictors, absence of N20 SSEP wave was the one most commonly used for treatment decisions (see ESM Table E8). One observational study found that results of SSEPs are more likely to influence physicians’ and families’ decision to withdraw life-sustaining therapies than those of clinical examination or EEG.

4.3. Biomarkers

Serum biomarkers NSE and S-100B have important theoretical advantages, such as ease of sampling, quantitative results, and likely independence from the effects of sedative drugs. However, similarly to our previous findings in patients not treated using TH, their thresholds for prediction of poor outcome with 0% FPR in patients treated with TH varied largely. For NSE, that threshold ranged from 44.3 to 81.8 μg/l at 48 h, and from 57.2 to 78.9 μg/l at 72 h, but in a very recent cohort study, values close to 100 μg/l at 48 h were still compatible with good outcome. S-100B values for 0% FPR ranged from 0.18 at 24 h and 0.30 at 48 h (see Table 2c).

One major reason for this variability may be the presence of extracerebral sources of biomarkers. NSE values are markedly increased in the presence of haemolysis because red blood cells contain NSE. Other less common sources of NSE include neuroendocrine tumours and small-cell lung carcinoma. S-100B is contained in muscle and adipose tissue, therefore its levels could be increased by a thoracic trauma caused by prolonged CPR. Another source of variability for biomarkers’ values is the use of heterogeneous immunoenzyme measurement techniques. A third issue in the interpretation of biomarker results could be our incomplete knowledge of the kinetics of their blood concentrations in the first few days after cardiac arrest. It is likely that the best threshold for outcome prediction with those biomarkers actually varies over time. Finally, serum concentrations of biomarkers are per se continuous variables, which limits their applicability for predicting a dichotomous outcome as CPC, especially when a threshold for 0% FPR is required.
4.4. Imaging

The main CT finding of anoxic-ischaemic cerebral insult is brain swelling, which appears as a reduction of ventricles and sulci and an attenuation of the grey–white matter interface. In a single study included in our review, the presence of cerebral oedema on brain CT scan, quantitatively measured as a CT density ratio of grey to white matter below 1.22 measured at the basal ganglia level, was associated with death or vegetative state at discharge in 100% of cases. This study included only 20 patients and had a limited generalizability since it was made on victims of cardiac arrest caused by drowning. Its findings will need confirmation from larger cohorts.

MRI can be used to detect the presence of ischaemic brain lesions in resuscitated comatose patients. In a paper included in our review, presence of moderate to severe abnormalities on DWI or FLAIR sequences in either cortical or deep grey matter at a median of 80 h (3.5 days) from cardiac arrest correctly predicted poor outcome in all but three patients. Specificity increased to 100% when patients evaluated within 14 h were excluded. The authors also performed quantitative MRI measurements using ADC and found that poor outcome patients exhibited a nadir in ADC values at 3–5 days after cardiac arrest, which therefore appeared to be the optimal time window for prognostication. Indeed, another study from the same group showed that a global reduction of ADC values predicted poor outcome with 100% specificity within that time window. Two other studies included in our review showed that presence of extensive DWI changes in both cortical and deep grey matter nuclei within 5 days from cardiac arrest were 100% specific for poor outcome.

In summary, detection of diffuse brain cytotoxic oedema using MRI DWI at 3–5 days after cardiac arrest can accurately predict poor neurological outcome after cardiac arrest. Being a qualitative technique, DWI is prone to interobserver variability, but it can be standardised using quantitative methods like ADC. ADC measurements, however, at present require an off-line data analysis with dedicated software and are not universally available. Brain MRI in resuscitated comatose patients has mainly been used as a research rather than a prognostication tool and it has not attained a widespread use yet. Relatively long measurement times and lack of bedside availability may limit MRI use in the most unstable resuscitated patients.

4.5. Self-fulfilling prophecy

Predictors of poor outcome in comatose patients resuscitated from cardiac arrest are prone to self-fulfilling prophecy. In our review, only 4/37 studies (11%) – two of which from the same group – reported blinding of the treating team from the results of the investigated predictor. In two of these studies, results of the predictor (absence of N2O SSEP wave) recorded during TH were not disclosed, but if patients remained comatose after rewarming, a second SSEP was performed and results were disclosed to the treating team, who used this information for treatment decisions.

A treatment suspension policy was reported in 22/37 studies (54%). Nine of those studies the treatment suspension policy was based, at least in part, on one or more of the investigated predictors (see ESM Table E8). In two studies, the 2005 AAN guidelines were used for treatment decisions, although these guidelines were based on evidence from patients not treated with TH. Treatment limitations were applied at a minimum of 3 days or less from cardiac arrest in 8 studies and from 3 to 7 days in 6 studies, while in the remaining 23 studies the minimal duration of life support measures was not reported. One study reported suspension of support as early as day 2, still during TH, in three patients upon family request.

One recent observational study also documented that 18/49 (37%) patients resuscitated from cardiac arrest and treated using TH were given a poor prognosis before the end of TH.

At present, several authors agree that in patients treated with TH the time to prognostication should be delayed beyond 72 h after rewarming, especially with respect to clinical examination. However, there is no definite evidence on how long this should be in order to avoid missing cases of late recovery. In two studies included in our review, recovery of consciousness occurred up to 6 and 25 days from cardiac arrest, respectively. In a recent case report on post-cardiac arrest myoclonus, it occurred at 14 days.

Prevention of self-fulfilling prophecy bias would require blinding of test results to the treating team and providing sufficiently prolonged life support in patients who do not recover consciousness after resuscitation and rewarming. Both those tasks are difficult to accomplish. Some predictors, such as results of clinical examination, cannot be concealed to the treating team. Others, such as an EEG, should not be concealed as they can reveal the presence of potentially treatable complications, like seizures. In some institutions, having a dedicated investigator not involved in patient management who will ensure blinding of collected data may not be feasible. On the other hand, indefinite supportive care in potentially hopeless patients raises both ethical and financial concerns. A potential source of data, however, could be represented by those familial or society contexts where withdrawal of life support is not accepted. In 1998 a cross-sectional interview in outpatient practices of three university hospitals in USA revealed that even in a theoretical scenario where there was no hope of waking up, 15–20% of patients would still choose to have aggressive supportive care maintained. Data from studies on patients resuscitated from cardiac arrest in Seattle, Washington, show that about 20% of those still unconscious at 2–3 days have full supportive care maintained indefinitely. A long-term assessment of patients with predicted poor prognosis in a prospective study may therefore be feasible, at a cost of a likely slow patient recruitment.

Another way of limiting the risk of falsely pessimistic predictions would be using a multimodal approach. Combining predictors of poor outcome seems to be the most logical solution to reduce the risk of false positives even though this may reduce sensitivities. Unfortunately, only a few studies to date have evaluated this multimodal approach, mainly deriving results post hoc without a prospective validation.

4.6. Study limitations

Our review has several limitations. Firstly, the lack of specific GRADE guidelines for evaluation of prognostic accuracy studies required us to adapt the GRADE guidelines for diagnostic studies. Some of our choices, such as assigning a serious limitation to studies that lacked blinding, may be considered arbitrary. However, lack of blinding and the consequent risk of self-fulfilling prophecy have long been recognised as a major limitation of prognostic accuracy studies in post-cardiac arrest patients. Studies included in our review did not have a consistent timing of outcome measurement, since we chose to include studies regardless of the length of their follow-up. In our review, prediction of poor outcome refers to a period ranging from discharge from intensive care unit to six months, even if all but one study on the two most robust predictors in this review – SSEPs and EEG reactivity – measured the outcome at least one month after cardiac arrest. In studies where outcome evaluation was undertaken too early the number of patients with poor outcome could have been overestimated, since patients assigned CPC 3 may further improve in the first few months after cardiac arrest. Thirdly, most predictors were documented in only one or two studies and their reproducibility needs to be verified in further studies. A publication bias cannot therefore be excluded for these predictors. Fourthly, we included both studies
with CPC 3–5 and studies with CPC 4–5 in our review and described those outcomes separately. This may be seen as an inconsistency. However, there is no general consensus on what represents a poor neurological outcome after resuscitation. While the majority (70%) of studies included in this review defined poor outcome as CPC 3–5, the opposite was true for studies performed on patients not treated with TH, where poor outcome was defined as CPC 4–5 in 76% of cases. Ideally, both outcomes should have been reported for each predictor. However, this would have required access to original patient data, which was beyond the scope of this aggregated data meta-analysis.

5. Conclusions

Our analysis identified a series of early predictors of poor neurological outcome with 100% specificity and narrow CIs in comatose patients resuscitated from cardiac arrest and treated using TH. For poor outcome defined as CPC 4–5 these included presence of either burst-suppression or electrographic status epilepticus evolving from a burst-suppression anytime (QOE: Low). However, these predictors were described in a small number of patients in a single study, which severely limits the relevance of these findings. Moreover, their use appear problematic, due to the inconsistent definitions of both burst-suppression and status epilepticus (see Table E7).

For poor outcome defined as CPC 3–5, these included a bilaterally absent N20 SSEP wave during TH (4 studies; QOE: moderate) or after rewarming (5 studies; QOE: Low), a nonreactive EEG background after rewarming (3 studies; QOE: Low) and a combination of absent pupillary light and corneal reflexes plus a motor response no better than extension (M ≤ 2) (1 study; QOE: very low) after rewarming. An S-100B > 0.18–0.21 mcg/L at 24 h, a NSE > 81.8 mcg/L or a S-100B ≥ 0.3 at 48 h, and an NSE > 78.9 mcg/L at 72 h also predicted a CPC 3–5 with narrow 95% CIs (5 studies; QOE: from moderate to very low) but no consistent threshold for 0% FPR could be identified. All these predictors had important limitations, the most important being the lack of blinding in included studies and the frequent use of the investigated predictor to support decisions of WLST, with a consequent risk of self-fulfilling prophecy. This occurred in the single study on the combination of absent PLR and CR plus M ≤ 2, in 1/3 studies documenting a nonreactive EEG background and in 6/8 studies documenting absence of N20 SSEP wave. Despite its limitations, bilateral absence of the N20 SSEP wave appears as the most reproducible predictor with 0% FPR.

Although not attaining 0% FPR, status myoclonus and status epilepticus showed a high specificity and can be considered as important predictors, but were both inconsistently described in literature and a common definition based on consensus is warranted. PLR after rewarming also showed a FPR close to 0% but its sensitivity was low. Given the paucity of robust univariate predictors with 100% specificity, predicting outcome on the basis of a single index appears unsafe. An integrated approach using a combination of predictors along with a careful evaluation of all available clinical information at present is probably the best strategy for early prognostication after cardiac arrest.

Author contributions

Study design: Claudio Sandroni, Fabio Cavallaro, Jerry Nolan.

Data search and collection, grading: Matteo Biancone, Clifton Callaway, Fabio Cavallaro, Sonia D’Arrigo, Michael Kuiper, Claudio Sandroni, Tommaso Sanna.

Drafting of the manuscript: Claudio Sandroni, Fabio Cavallaro, Jerry Nolan (Manuscript’s body text); Matteo Biancone, Sonia D’Arrigo, Giacomo Della Marca (Tables).

Revision of the manuscript: Clifton Callaway, Giacomo Della Marca, Michael Kuiper.

Data analysis: Fabio Cavallaro, Claudio Sandroni.

Statistical revision: Alessio Farcomeni.

Conflict of interest statement

Claudio Sandroni, Fabio Cavallaro, Clifton Callaway, Sonia D’Arrigo, Tommaso Sanna, Michael Kuiper, Matteo Biancone Giacomo Della Marca and Alessio Farcomeni have no conflicts of interest to declare.

Jerry Nolan is Editor-in-Chief of Resuscitation.

Acknowledgements

We gratefully thank the following investigators for having provided original data from their studies:

1. Dr. Amy Z. Crepeau, Department of Neurology, Mayo Clinic, Rochester, MN;
2. Dr. Stephane Legriel, Intensive Care Department, CH Versailles – Site André Mignot, Le Chesnay, France;
3. Dr. Pascal Stamm, Department of Anaesthesia and Intensive Care, Centre Hospitalier, Luxembourg;
4. Dr. Michael Mlynash, Stanford Stroke Center, Department of Neurology and Neurological Sciences, Stanford University, Palo Alto, CA, USA;
5. Dr. Kyu Nam Park, Department of Emergency Medicine, Seoul St Mary’s Hospital, Seoul Korea;
6. Dr. Andrea Rossetti, Department of Clinical Neurosciences, University Hospital and Faculty of Biology and Medicine, Lausanne, Switzerland;
7. Dr. G. Bryan Young, FRCPC, Department of Clinical Neurological Sciences, London Health Sciences Centre, London, Ontario, Canada.

We also gratefully thank Dr. Lorenzo Compieta for providing statistical advice during data analysis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.resuscitation.2013.06.020.

References


Survivors

Ron

Deeks

specific

from

and

Schmitt

Med

Jennett

Approximate

bilateral

hypothermia

arrest

with

GL,

van

B.

Kawai

M,

et

al.

cardiac

Care

2006;70:52–8.

51.

57.

46.

43.

because

return

arrest—recommendations

for

post-cardiac arrest—recommendations

for

laboratory

and

Leproust

L.

C.

Pettila

enolase

after

arterial

because

arrest

in

patients

having

diagnosis

of


Costanzo M, Zetterberg H, Blennow K, Rosengren L, Robertsson S. S-100B is superior to NSE, BDNF and GFAP in predicting outcome of resuscitation from cardiac arrest with hypothermia treatment. Resuscitation 2011;82: 207–16.


