ORIGINAL

Quantitative versus standard pupillary light reflex for early prognostication in comatose cardiac arrest patients: an international prospective multicenter double-blinded study

Mauro Oddo[1]* , Claudio Sandroni[2], Giuseppe Citerio[3,4], John-Paul Miroz[1], Janneke Horn[5], Malin Rundgren[6], Alain Cariou[7,8], Jean-François Payen[9], Christian Storm[10], Pascal Stammet[11] and Fabio Silvio Taccone[12]

Abstract

Purpose: To assess the ability of quantitative pupillometry [using the Neurological Pupil index (NPi)] to predict an unfavorable neurological outcome after cardiac arrest (CA).

Methods: We performed a prospective international multicenter study (10 centers) in adult comatose CA patients. Quantitative NPi and standard manual pupillary light reflex (sPLR)—blinded to clinicians and outcome assessors—were recorded in parallel from day 1 to 3 after CA. Primary study endpoint was to compare the value of NPi versus sPLR to predict 3-month Cerebral Performance Category (CPC), dichotomized as favorable (CPC 1–2: full recovery or moderate disability) versus unfavorable outcome (CPC 3–5: severe disability, vegetative state, or death).

Results: At any time between day 1 and 3, an NPi ≤ 2 ($n = 456$ patients) had a 51% (95% CI 49–53) negative predictive value and a 100% positive predictive value [PPV; 0% (0–2) false-positive rate], with a 100% (98–100) specificity and 32% (27–38) sensitivity for the prediction of unfavorable outcome. Compared with NPi, sPLR had significantly lower PPV and significantly lower specificity ($p < 0.001$ at day 1 and 2; $p = 0.06$ at day 3). The combination of NPi ≤ 2 with bilaterally absent somatosensory evoked potentials (SSEP; $n = 188$ patients) provided higher sensitivity [58% (49–67) vs. 48% (39–57) for SSEP alone], with comparable specificity [100% (94–100)].

Conclusions: Quantitative NPi had excellent ability to predict an unfavorable outcome from day 1 after CA, with no false positives, and significantly higher specificity than standard manual pupillary examination. The addition of NPi to SSEP increased sensitivity of outcome prediction, while maintaining 100% specificity.

Keywords: Pupillometry, Pupillary reactivity, Neurological pupil index, Outcome, Cardiac arrest, Prognostication

*Correspondence: mauro.oddo@chuv.ch
[1] Department of Intensive Care Medicine, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne, CH-1011 Lausanne, Switzerland
Full author information is available at the end of the article
Mauro Oddo and Claudio Sandroni contributed equally to this work.

Springer

Introduction

Approximately one-half of patients admitted to the intensive care unit (ICU) in coma following a cardiac arrest (CA) die or have an unfavorable neurological outcome at hospital discharge [1, 2]. Mortality in comatose resuscitated CA patients predominantly results from the withdrawal of life-sustaining therapies (WLST), based upon the prediction of severe irreversible post-anoxic brain damage [3, 4]. Clinical neurological examination, and in particular the pupillary light reflex (PLR), is a key component of the prognostic assessment of CA patients [5, 6]. In subjects who remain comatose on day 3 after CA, a bilaterally absent PLR is considered as highly specific for poor neurological outcome [7–9], and is currently recommended—in combination with the bilateral absence of N20 wave on short-latency somatosensory evoked potentials (SSEP)—as the most accurate test for prognosticating outcome of coma after CA [6]. However, it must be recognized that the quality of evidence supporting prognostic tests based on clinical examination is limited by both the absence of an objective quantitative assessment of pupillary function and the lack of blinded measurements, i.e. the results of the index test are usually available to clinicians involved in patient care and WLST decisions, thereby causing a self-fulfilling prophecy [10].

Automated infrared pupillometry provides a quantitative measurement of pupillary response to a calibrated light stimulus [11, 12] and improves reliability and validity in comparison with the standard manual methods for the subjective assessment of PLR, especially in the acute neurological setting [13–15]. Recent studies have suggested that quantitative PLR (qPLR; expressed as the percentage pupillary constriction to a calibrated light stimulus) may improve prediction of neurological outcome after CA [16–19]. Most of these studies, however, were single-centered and had both a small sample size and/or an absent or incomplete blinding, and only qPLR was examined. Quantitative pupillometry, however, allows a more comprehensive assessment of pupillary function, in particular when using the Neurological Pupil index (NPi). The NPi is a scalar value (between 0 and 5) which is calculated based on an algorithm that accounts for several measured pupillary variables, including size, percentage constriction, constriction velocity, dilation velocity, and latency [20]. Contrary to the qPLR, the NPi is only minimally influenced by medications (in particular, opioids) and ambient light, and it accounts for individual baseline pupil size [20–23].

We therefore designed a multicenter study in comatose post-CA patients to assess the prognostic accuracy of quantitative pupillometry, performed in a blinded fashion and on an adequately sized population. The primary study endpoint was to examine the value of NPi in comparison to that of standard manual PLR (sPLR). Additionally, we analyzed the value of the NPi when used in combination with SSEP.

Take-home message

A novel tool for the automated measurement of the pupillary function, assessed blindly in a large multicentre international prospective study, already had 100% specificity and positive predictive value to predict unfavourable outcome in comatose post-anoxic patients by day 1 after cardiac arrest.

The Quantitative Neurological Pupil index, unaffected by confounders such as hypothermia and sedation, had significantly higher specificity than standard pupillary light reflex, with no false positives. These findings have potentially important clinical implications and indeed demonstrate that early prognostication, using quantitative devices for the objective assessment of pupillary function, appears accurate and reliable across different cardiac arrest centers.

Patients and methods

Study design

This was an investigator-initiated, prospective, multicenter, international prognostication study conducted in ten European ICUs (ClinicalTrials.gov Identifier: NCT02607878). Centers were selected within the Neurointensive Care section of the European Society of Intensive Care Medicine (ESICM), based on their expertise on prognostication after CA. The study protocol was approved by the Ethical Committee at each institution, and written informed consent to participate in the study was obtained from patients' next of kin or a legal surrogate, according to local ethical recommendations. Accuracy of the collected data was monitored by principal and local investigators. Study design and methodology was in accordance with the STARD guidelines for reporting diagnostic accuracy studies [24].

Role of the funding source

The device manufacturer (Neuroptics®, Irvine, CA, USA) supplied equipment and disposables for the study, but it did not provide any financial support for the study and had no role in the study design, data collection, analysis and interpretation, or writing of the report.

Blinding

The ICU clinicians/nurses who performed pupillometry tests were not involved in patient care and the outcome assessors were blinded to quantitative pupillometry data. Decisions about intensity of care and WLST were based on a local prognostication algorithm at each center (see below) which did not include any of the pupillometry data.

Participants

Patients were adult subjects (> 18 years) with coma (Glasgow Coma Score ≤ 6) following a CA (regardless of the initial arrest rhythm), admitted to the ICU from January 2015 to March 2017. Patients were treated according to the European Resuscitation Council (ERC)–ESICM guidelines for post-resuscitation care [5], including TTM (33 or 36 °C, according to local practices at each center) and sedation (midazolam or propofol)/analgesia (morphine, fentanyl or remifentanil).

Demographic and clinical variables

Demographic and clinical variables included age, gender, initial arrest rhythm [categorized as shockable (ventricular fibrillation) vs. non-shockable (asystole or pulseless electrical activity)], duration of CA [defined as the time from CA to return of spontaneous circulation (ROSC)], admission norepinephrine dose, and admission body temperature, TTM target temperature (categorized as TTM 33 °C or TTM 36 °C), TTM duration, and the total cumulative dose over the first 48 h of sedatives, analgesics, and norepinephrine.

Intervention

Quantitative pupillometry (index test) was performed using the NPi®-200 pupillometer (Neuroptics®). The NPi®-200 uses an infrared camera that integrates a calibrated light stimulation of fixed intensity (1000 Lux) and duration (3.2 s), allowing a rapid and precise measure (0.05 mm limit) of the pupil size and of a series of several dynamic pupillary variables. Based on an integrated algorithm, the NPi®-200 calculates the NPi. Pupil variables (including size, percentage constriction, latency, constriction velocity, and dilation velocity) are all parameters of the NPi algorithm [20]. Each variable from an individual pupil measurement, taken by the pupillometer, is compared against the mean of a reference distribution of healthy subjects for the same variable. Finally, the set of all the standardized differences are combined to fall into a scale set between 0 and 5 (with 0.1 decimal precision). An NPi score ≥ 3 defines normal pupillary light reactivity; for example, a value of 4.7 is considered more reactive than a value of 3.2. An NPi score < 3 denotes an abnormal pupillary light reactivity (i.e., weaker than a normal pupil response, as defined by the multidimensional normative model); for example, a value of 1.5 is more abnormal than a value of 2.8.

Quantitative pupillometry tests were performed on each eye daily from day 1 to 3 after CA and included the NPi and the qPLR (expressed as the percentage of pupillary constriction to the standardized light stimulus). At each time-point, the lowest value (for both NPi and qPLR) from each eye was retained for the analysis.

Standard pupillary light reflex (sPLR) (reference standard) was performed by the nurse or the clinician in charge of the patient in parallel, at the same time-points, using a manual flash lamp. Absent sPLR was defined when pupillary reactivity was not present bilaterally. There was no pre-specified sequential order for NPi or sPLR.

Prognostication and WLST

Outcome prediction was based on local prognostication algorithms that included the sPLR and at least one additional modality among SSEP, electroencephalography (EEG), and serum neuron specific enolase (NSE). Prognostication tests were performed at days 1–3, irrespective of sedation/analgesia. Decisions about continuation of care or WLST were taken based on current prognostication algorithms [5, 6, 25]; however, the choice of the combination of the additional prognosticating tests (SSEP, EEG, NSE) was left to the discretion of each investigator, based on routine practices at each center.

Outcome

Outcome was assessed at 3 months, using the Glasgow–Pittsburgh Cerebral Performance Categories (CPC), dichotomized as favorable (CPC 1 = full recovery and 2 = moderate disability, returned home) versus unfavorable (including CPC 3 = severe disability, at rehabilitation facility, 4 = vegetative state and 5 = death) [26, 27].

Study endpoints

The primary endpoint of this study was to examine the performance of quantitative NPi in predicting 3-month unfavorable outcome and to compare its prognostic value to that of sPLR. ERC-ESICM guidelines on post-resuscitation care recommend using pupillary reflexes and SSEP as the initial step for the prognostication of post-cardiac arrest patients [5, 6]; therefore, as additional endpoint, we evaluated the prognostic value of the NPi when used in combination with SSEP.

Statistical analysis

Descriptive statistics were computed for all study variables and normal distribution was assessed using the Kolmogorov–Smirnov test. Differences between groups were compared using the Wilcoxon test for continuous variables and the Fisher exact test for categorical variables. Data are presented as counts (percentage) or median (interquartile range). Prognostic performance of each predictor was analyzed by calculating specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), false-positive

rate (FPR) and the area under the receiver operating characteristic (ROC) curve. For the combination of NPi with SSEP, in patients where both tests were performed, if the N20 SSEP wave was bilaterally absent and NPI was between 0 and 2, they were identified as having a positive test. Instead, the absence of one or both these characteristics was identified as a negative test. Sensitivity and specificity for prediction of a poor outcome were calculated accordingly.

Statistical differences between AUCs were computed using the method described by De Long et al. [28]. The Mc Nemar test was used to compare specificity and sensitivity for unfavorable outcome of NPi versus sPLR at each time point. For continuous predictor variables, the optimal cut-offs, i.e. those maximizing the Youden index and the specificity, were also identified. According to Jones et al. [29], based on a specificity of 95% for qPLR in previous studies [16–18], a 95% confidence interval of 3%, and a prevalence of poor neurological outcome of 50%, a sample size of at least 400 patients was calculated. Statistical analysis was performed using STATA 12 (STATA®, College Station, TX, USA). Significance was set at $p < 0.05$.

Results

Patient characteristics

The study flow-chart is summarized in Fig. 1. From a total of 477 patients included, 456 patients were successfully analyzed for the primary endpoint. All patients underwent TTM for 24 h [199 (44%) in whom TTM was set at 33 °C, and 257 (56%) who had TTM at 36 °C]. At 3 months, 187 patients (41%) had a favorable outcome, including CPC 1 ($n = 133$) and CPC 2 ($n = 54$). A total of 269 patients (59%) had an unfavorable outcome, including CPC 3 ($n = 19$), CPC 4 ($n = 7$) and CPC 5 ($n = 243$).

The general patient characteristics, according to the two outcome subgroups, are reported in Table 1. Patients with an unfavorable outcome were significantly older and had significantly higher rates of non-shockable CA rhythms and longer CA duration than those with a favorable outcome. The proportion of TTM at 33 versus 36 °C, TTM duration, admission norepinephrine, and the 48-h cumulative doses of sedatives, analgesics and norepinephrine did not differ significantly between the two outcome groups.

Prognostic value of Neurological Pupil index (NPi) and quantitative pupillary light reflex (qPLR)

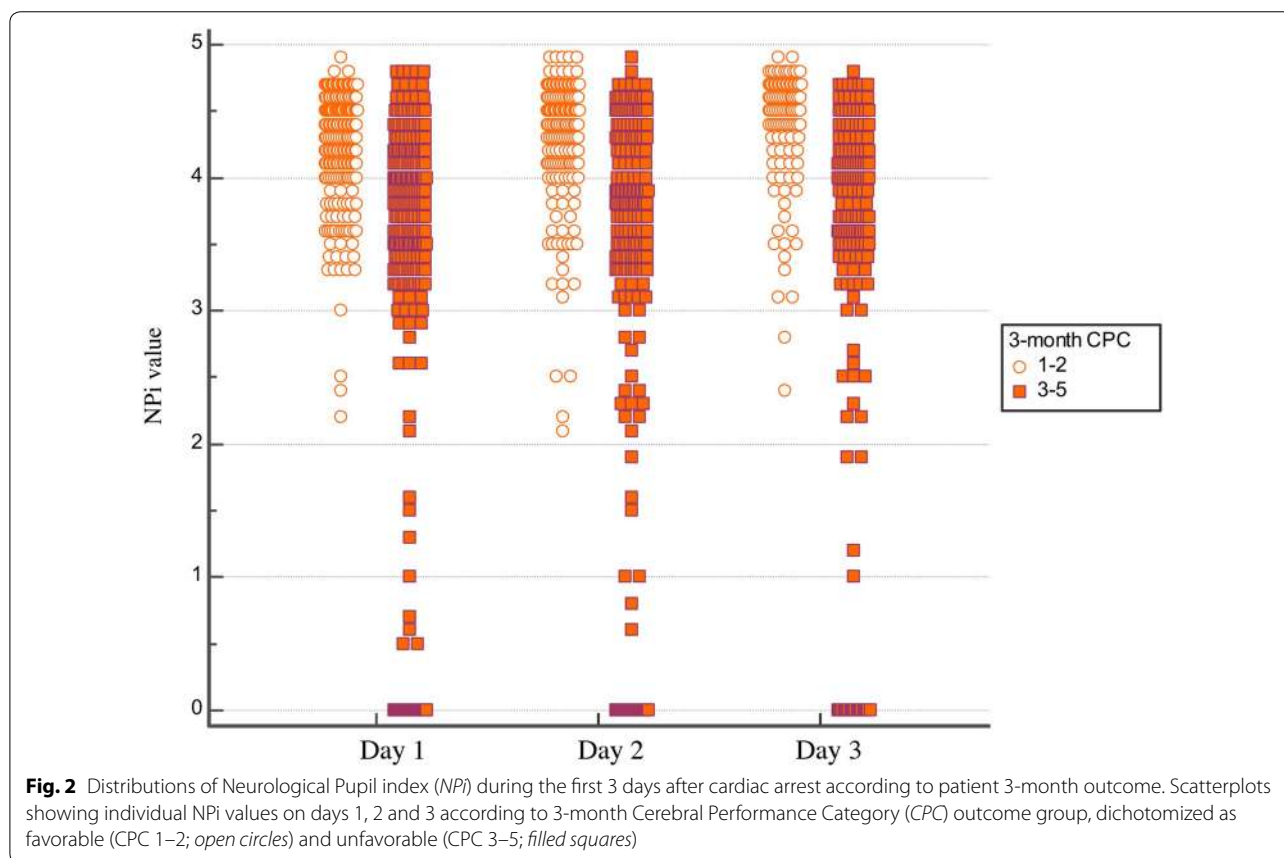At each time point, median NPi and qPLR values were significantly higher in the favorable versus the
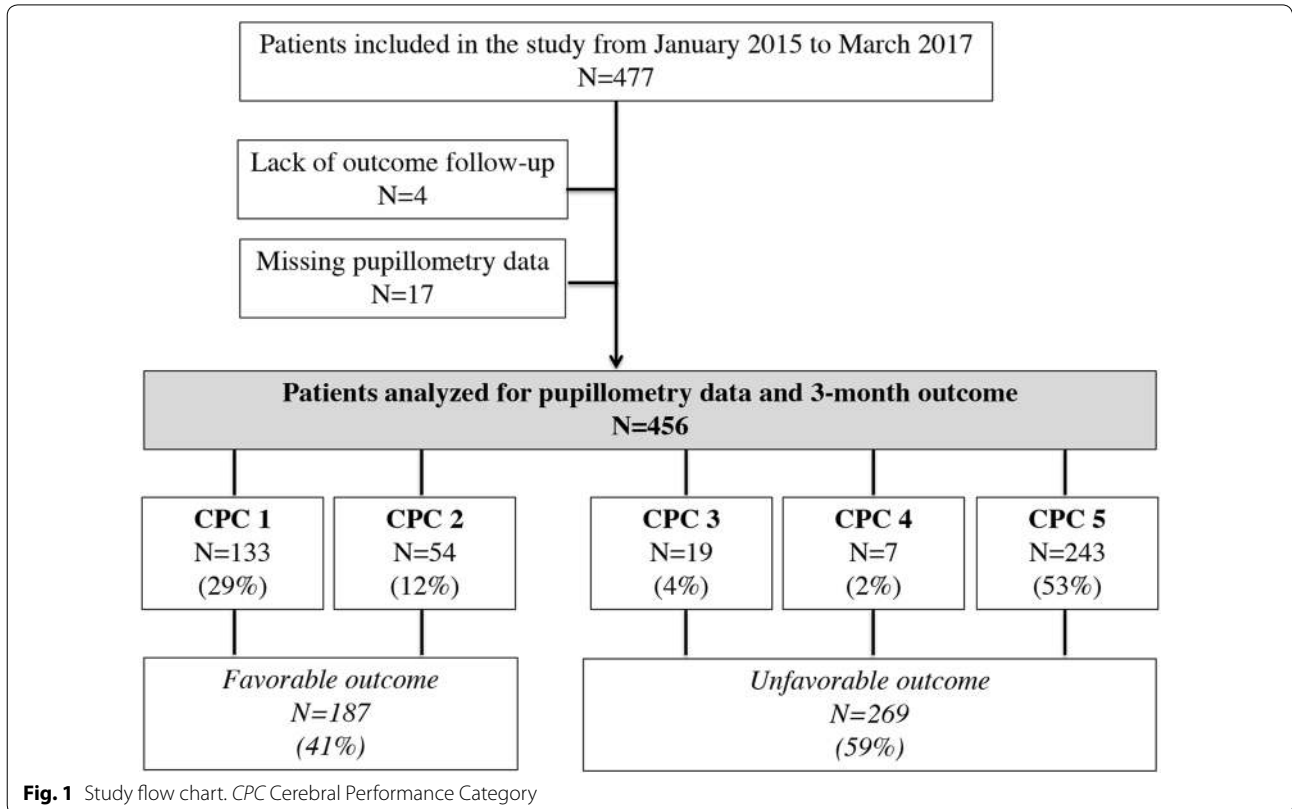
Fig. 1 Study flow chart. *CPC* Cerebral Performance Category

Table 1 Patient characteristics

Variable	Favorable outcome (*n* = 187)	Unfavorable outcome (*n* = 269)	*p* value
Age, years	60 (49–69)	65 (54–74)	0.0006
Male gender, *n* (%)	154 (80)	203 (72)	0.03
Non-shockable rhythm, *n* (%)	36 (19)	160 (57)	< 0.0001
Time from cardiac arrest to ROSC, min	19 (10–29)	28 (16–41)	< 0.0001
Admission norepinephrine dose, µg/min	5.7 (2.4–10.5)	5.9 (1.4–19)	0.71
Admission body temperature, °C	35.5 (34.6–36.1)	35 (34–35.8)	0.001
TTM at 33 °C, *n* (%)	85 (42%)	116 (58%)	0.73
TTM at 36 °C, *n* (%)	103 (40%)	152 (60%)	
TTM duration, hours	24 (24–24)	24 (24–24)	0.85
48-h propofol dose, mg/h	1.7 (1.1–2.4)	1.8 (1.3–3.7)	0.48
48-h midazolam dose, mg/kg/h	0.08 (0.04–0.15)	0.11 (0.06–0.37)	0.19
48-h fentanyl dose, µg/kg/h	0.49 (0.26–1.4)	0.55 (0.14–1.4)	0.94
48-h norepinephrine dose, µg/min	4.9 (1–15)	5.4 (1–15)	0.91

Data are presented as count (percentage) or median (interquartile range)

ROSC return of spontaneous circulation, *TTM* targeted temperature management

Fig. 2 Distributions of Neurological Pupil index (*NPi*) during the first 3 days after cardiac arrest according to patient 3-month outcome. Scatterplots showing individual NPi values on days 1, 2 and 3 according to 3-month Cerebral Performance Category (*CPC*) outcome group, dichotomized as favorable (CPC 1–2; *open circles*) and unfavorable (CPC 3–5; *filled squares*)

unfavorable outcome group (ESM_Table 1). Distributions across the two main outcome groups of individual NPi values at day 1 to 3 after CA are illustrated in Fig. 2. ESM_Figure 1 illustrates the distributions of qPLR at each time-point for both outcome groups. ESM_Table 2 summarizes the cut-off values associated to a specificity ranging from 95% to 100%, i.e. a false positive rate from 5% to 0%, along with those maximizing the Youden index

for both predictors; no qPLR values corresponding to 100% specificity at days 1 and 2 were found. At all time-points, the ability to predict outcome, assessed using the AUC (95% confidence interval), was higher for NPi than for qPLR: 0.77 (0.73–0.81) versus 0.76 (0.72–0.80) on day 1 ($p = 0.73$); 0.76 (0.72–0.81) versus 0.73 (0.68–0.78) on day 2 ($p = 0.81$); 0.78 (0.72–0.83) versus 0.71 (0.65–0.77) on day 3 ($p = 0.04$); Fig. 3.

A cut-off for NPi ≤ 2, occurring at any time between day 1 and day 3, provided maximum s ensitivity [32% (27–38)] to predict unfavorable outcome, while achieving 100% (98–100) specificity and 100% positive predictive value [0% (0–2) false-positive rate] (Table 2). Using the cut-off of abnormal NPi (< 3) [23] increased sensitivity [38% (32–44)] but at the expense of a lower specificity [96% (92–98); 6% FPR].

Prognostic value of standard manual pupillary light reflex (sPLR)

The specificity of bilaterally absent sPLR for unfavorable outcome was 90% on day 1 and 2, increasing to 94% (86–98) on day 3; sensitivity decreased from 35% (29–42) to 18% [11–26], respectively (Table 2). Five of 78 patients who had bilaterally absent sPLR on day 3 eventually had a favorable outcome (6% FPR). Of note, all five patients had smaller pupil size (1.9 ± 0.22 vs. 2.8 ± 1.05 mm in those
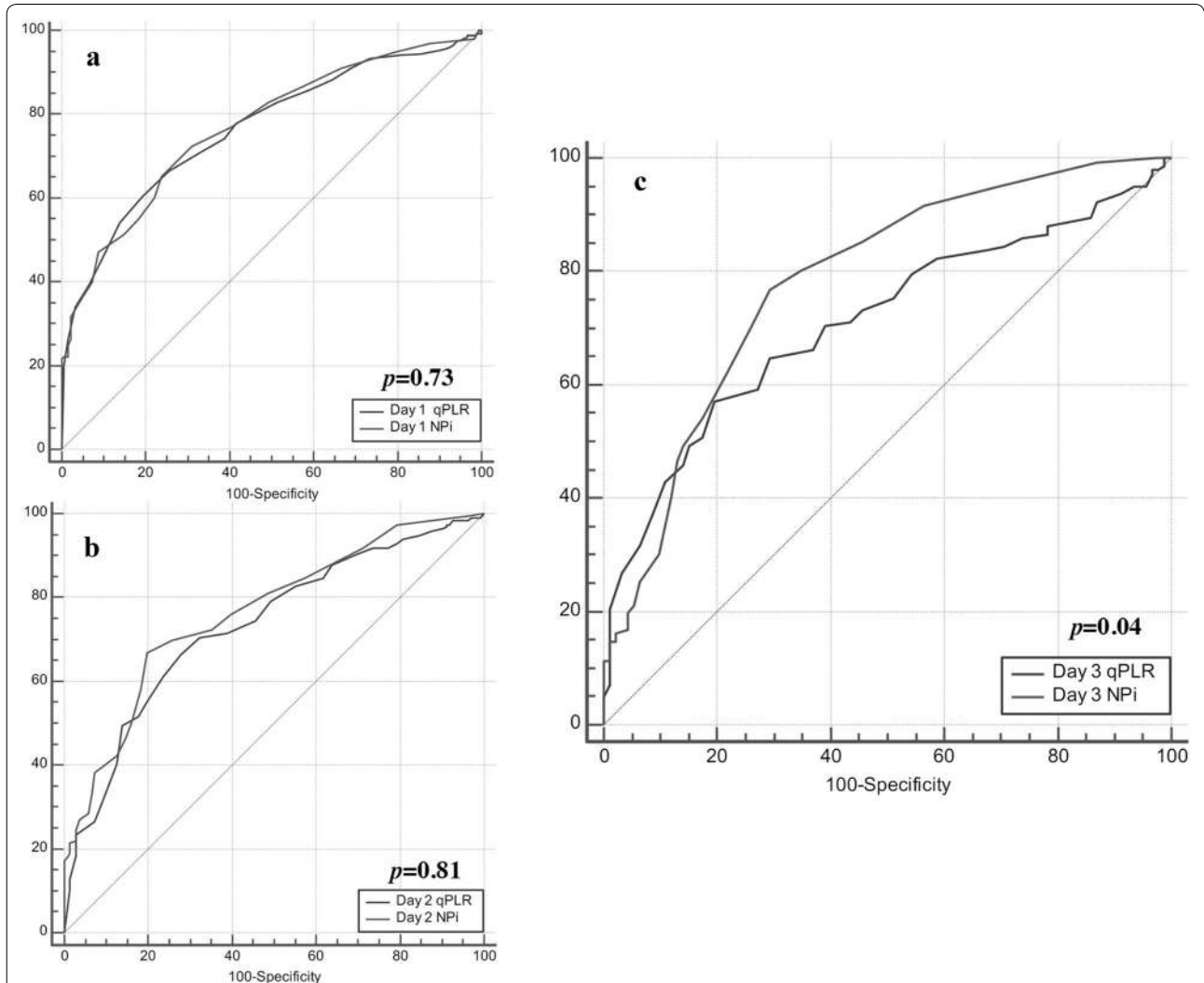
Fig. 3 Performance of Neurological Pupil index (*NPi*) versus quantitative pupillary light reflex (*qPLR*) to predict 3-month outcome (Cerebral Performance Category) at day 1 to 3 after cardiac arrest. Diagrams illustrate the area under the ROC curve (AUC) of qPLR (*blue line*) and NPi (*grey line* at days 1, 2 and 3; *p* for comparison of NPi AUC versus qPLR AUC. AUC of NPI to predict outcome at day 3 after cardiac arrest was significantly better than the AUC of qPLR [AUC 0.78 (0.72–0.83) vs. 0.71 (0.65–0.77) on day 3 ($p = 0.04$)]

Table 2 Specificity, sensitivity, positive predictive value, negative predictive value and false-positive rate for unfavorable outcome (CPC 3–5) of the different prognostic tests

Day after cardiac arrest	Sample size (*n*)	CPC 3–5 *n* (%)	Specificity % (95% CI)	Sensitivity % (95% CI)	Positive predictive value % (95% CI)	Negative predictive value % (95% CI)	False-positive rate % (95% CI)
Neurological pupil index (NPi) ≤ 2							
Day 1–3	456	269 (59)	100 (98–100)	32 (27–38)	100 (100–100)	51 (49–53)	0 (0–2) %
Day 1	450	264 (59)	100 (98–100)	22 (17–27)	100 (100–100)	47 (46–49)	0 (0–2) %
Day 2	361	213 (59)	100 (98–100)	19 (14–25)	100 (100–100)	46 (45–48)	0 (0–2) %
Day 3	271	166 (61)	100 (97–100)	17 (12–24)	100 (100–100)	43 (41–44)	0 (0–3) %
Bilaterally absent standard pupillary light reflex (sPLR)							
Day 1	392	225 (57)	90 (85–94)	35 (29–42)	83 (75–89)	51 (48–54)	10 (6–15) %
Day 2	278	163 (59)	90 (84–95)	29 (22–36)	81 (70–89)	47 (44–50)	10 (5–16) %
Day 3	206	128 (62)	94 (86–98)	18 (12–26)	82 (65–92)	41 (39–43)	6 (2–14) %
Bilaterally absent somatosensory evoked potentials (N20 wave)							
Day 2–3	188	133 (71)	100 (94–100)	48 (39–57)	100 (100–100)	44 (40–48)	0 (0–6) %
Combination of NPi ≤ 2 and bilaterally absent somatosensory evoked potentials							
Day 2–3	188	133 (71)	100 (94–100)	58 (49–66)	100 (100–100)	55 (50–59)	0 (0–6) %

CI confidence interval, *CPC* Cerebral Performance Category

with unfavorable outcome, $p = 0.009$) and lower qPLR (15 ± 2 vs. $24 \pm 8\%$, $p = 0.01$), but their NPi was normal (4.5 ± 0.1 vs. 4.4 ± 0.4, $p = 0.72$).

Compared to NPi ≤ 2, sPLR had a lower performance to predict an unfavorable outcome at each time point tested, both in terms of specificity ($p < 0.0001$ at day 1; $p < 0.001$ at day 2; $p = 0.06$ at day 3) and sensitivity ($p < 0.0001$ at day 1; $p = 0.0003$ at day 2; $p = 0.18$ at day 3; McNemar test).

Additional data: combination of quantitative NPi and SSEP to predict unfavorable outcome

In a subset of 188 patients for whom SSEP were available, we found that a bilaterally absent N20 wave had 48% (39–57) sensitivity and 100% (94–100) specificity for unfavorable neurological outcome at 3 months. The combination of bilaterally absent SSEP with an NPi ≤ 2 provided the highest sensitivity of all tests (58% (49–66), with comparable specificity 100% (94–100) (Table 2). Furthermore, among the 56/123 patients with normal SSEP and unfavorable outcome (i.e. false negatives), 15 (27%) had NPi ≤ 2 and were therefore correctly predicted only by quantitative pupillometry.

Discussion

This is the largest multicenter study on the prognostic value of pupillary reactivity in the acute phase of post-anoxic coma. Its main findings are that, in patients who are comatose following resuscitation from cardiac arrest, a quantitative measurement of the pupillary function, using automated infrared pupillometry, predicted an unfavorable neurological outcome at 3 months more accurately than standard pupillary examination using a manual device. The predictive value for an unfavorable outcome of NPi was superior to that of sPLR at each time-point tested (Table 2). Further analysis to compare prognostic performance between the two tests at separate time points provided significantly higher specificity and sensitivity to predict an unfavorable outcome for NPi versus sPLR; these differences were statistically significant at days 1 and 2 (both for specificity and sensitivity), while an almost significant trend in favor was found at day 3 for specificity of NPi versus sPLR. In particular, NPi ≤ 2 already had 100% specificity and positive predictive value at a very early time-point (day 1) following cardiac arrest. This is in contrast with the common finding that pupillary reactivity achieves its maximal specificity not earlier than day 3 [8, 9, 30]. A likely explanation of this finding is that—differently from the sPLR—NPi is a continuous scaled variable that allows discrimination between various degrees of severity of post-anoxic pupillary dysfunction. Our study further indicates that an accurate prediction of poor neurological outcome using PLR does not require a complete abolition of pupillary reactivity; in our cohort, in fact, a significant reduction of NPi (≤ 2) was a sufficient criterion. Regarding NPi, previous reports described NPi < 3 as the threshold for abnormality [23]: in our study, however, this cut-off did not achieve 100% specificity for a 3-month unfavorable outcome after CA.

Preliminary single-center studies, evaluating the value of quantitative pupillometry, focused exclusively on the

qPLR and found variable cut-offs of 100% specificity for poor prognosis, with thresholds varying from 10% [16] to 13% [17, 18]. These studies used other devices that did not compute automated NPi. Our study suggests that NPi may have better prognostic performance than qPLR, at least at day 3. In fact, we did not test different devices, and, therefore, we cannot speculate about potentially different prognostic performances between pupillometry devices. However, the advantage of using NPi in the setting of CA where sedation and analgesia may diminish pupil size, is that, contrary to qPLR, NPi is not affected by opioids and small pupil size [23].

Differently from previous prognostication studies, we could not confirm 100% specificity for a bilaterally absent sPLR on day 3. In fact, sPLR was scored as absent on day 3 in 5/79 patients who eventually had a favorable outcome (FPR 6%). Although unexpected, this finding is not completely unprecedented; previous studies by Bouwes et al. [31] and Dragancea et al. [32] found a FPR > 0% for bilaterally absent sPLR (FPR 1% and 2.1%, respectively). It should be noted that, similar to our study, both the above studies had a large sample size, were multicentered, and involved non-neurologists for the assessment of the standard PLR. One possible explanation of our findings may be that in a few patients standard PLR recording may have been suboptimal. If this was the case, our study indeed supports the notion that standardized objective pupillary assessment using automated pupillometry—rather than manual subjective assessment—may limit the risk of an incorrect PLR test [13–15], and, consequently, reduce falsely pessimistic prognostic prediction. Another plausible explanation for the observed five false positive manual PLR results was that in all these patients pupils were miotic and qPLR was on average 15%, corresponding to a variation in pupil size of only about 0.25 mm, which most likely was too small to be detected by manual assessment. This again reinforces the utility of quantitative pupillometry and further underlines the potential advantage of NPi, which is unaffected by pupil size [23], since NPi was completely normal in all of these patients.

In line with current ERC–ESICM guidelines on prognostication after CA [5, 6], the results of our study confirm that the assessment of PLR allows the prediction of an unfavorable outcome with high accuracy and precision and further support the use of quantitative NPi to minimize the risk of a false positive prediction. In addition, our study suggests that the outcome may be predicted earlier than what is currently recommended using sPLR. As for other quantitative prognosticators, such as serum NSE [33], NPi results could, therefore, be similarly collected during the first 3 days after CA and considered at the recommended time point of ≥ 72 h from CA [6]. Finally, we confirmed that the recommended combination of PLR and SSEP increases the sensitivity to predict an unfavorable outcome 3 days after CA, but the novelty of our findings was that 100% specificity could be reached only when quantitative NPi was used.

Study limitations

First, this study was observational and, as such, no homogeneous WLST protocol was imposed to the treating teams of the participating centers. However, all centers consistently applied their local WLST protocol, which was based on ERC–ESICM prognostication guidelines [6, 34] or equivalent algorithms [25, 35]. These recommendations should be the benchmark of every current prognostication study. Second, although clinicians involved in patient care were blinded to the results of quantitative pupillometry, this did not occur for sPLR, which was part of prognostication protocol. WLST decisions, however, were never based on a single predictor. NPi and sPLR are not equivalent, as demonstrated by their different specificities; however, as both tests are based on pupillary light reactivity, although to a lesser extent than for sPLR, some degree of self-fulfilling bias also cannot be completely excluded for NPi. Third, prognostic tests (especially at days 1 and 2) were performed in patients who were under sedation and analgesia, which may have reduced the accuracy of manual PLR assessments; as discussed above, miosis was associated with false positive predictions for s-PLR. In this setting, since it is not altered by sedatives/analgesics, NPi may confer a significant advantage over sPLR and provide accurate prognostic information, particularly in those patients with sedation or delayed awakening, which may represent up to 30% of cases [35]. Fourth, there was no sequential order for NPi and sPLR, and we cannot entirely rule out that in some cases the sPLR could have been recorded from an operator who was aware of the automated pupillometry results.

Conclusions

Our main findings are that a NPi ≤ 2, performed at any time between day 1 and day 3 following hospital admission, was 100% specific to predict an unfavorable 3-month neurological outcome and provided greater prognostic performance than standard manual pupillary light reactivity. Our data further suggest that using a prognostic approach that combines the NPi with somatosensory evoked potentials improved the sensitivity to predict an unfavorable outcome in patients with cardiac arrest, while providing equal 100% specificity. Future studies are needed to examine the value of a multimodal, prognostic algorithm that integrates a quantitative NPi

Author details

[1] Department of Intensive Care Medicine, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne, CH-1011 Lausanne, Switzerland. [2] Department of Anesthesiology and Intensive Care, Catholic University School of Medicine, Rome, Italy. [3] School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy. [4] Neurointensive Care, San Gerardo Hospital, Monza, Italy. [5] Department of Intensive Care, Academic Medical Center, Amsterdam, The Netherlands. [6] Department of Clinical Sciences, Anesthesiology and Intensive Care Medicine, Skåne University Hospital, Lund University, Lund, Sweden. [7] Réanimation Médicale-Hôpital Cochin, Paris, France. [8] Université Paris Descartes, Paris, France. [9] Department of Anesthesia and Critical Care, Grenoble Alpes University Hospital, Grenoble, France. [10] Department of Internal Medicine, Nephrology and Intensive Care, Charité-University, Berlin, Germany. [11] Medical and Health Department, National Fire and Rescue Corps, Luxembourg, Luxembourg. [12] Department of Intensive Care, Erasme Hospital, Université Libre de Bruxelles, Brussels, Belgium.

Compliance with ethical standards

Conflicts of interest
Mauro Oddo has received speaker honoraria from Neuroptics®. All the remaining authors declare no conflicts of interest related to this study.

References

1. Nolan JP, Berg RA, Bernard S, Bobrow BJ, Callaway CW, Cronberg T, Koster RW, Kudenchuk PJ, Nichol G, Perkins GD et al (2017) Intensive care medicine research agenda on cardiac arrest. Intensive Care Med 43:1282–1293
2. Callaway CW, Donnino MW, Fink EL, Geocadin RG, Golan E, Kern KB, Leary M, Meurer WJ, Peberdy MA, Thompson TM et al (2015) Part 8: post-cardiac arrest care: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 132:S465–S482
3. Dragancea I, Rundgren M, Englund E, Friberg H, Cronberg T (2013) The influence of induced hypothermia and delayed prognostication on the mode of death after cardiac arrest. Resuscitation 84:337–342
4. Lemiale V, Dumas F, Mongardon N, Giovanetti O, Charpentier J, Chiche JD, Carli P, Mira JP, Nolan J, Cariou A (2013) Intensive care unit mortality after cardiac arrest: the relative contribution of shock and brain injury in a large cohort. Intensive Care Med 39:1972–1980
5. Nolan JP, Soar J, Cariou A, Cronberg T, Moulaert VR, Deakin CD, Bottiger BW, Friberg H, Sunde K, Sandroni C et al (2015) European resuscitation council and european society of intensive care medicine 2015 guidelines for post-resuscitation care. Intensive Care Med 41:2039–2056
6. Sandroni C, Cariou A, Cavallaro F, Cronberg T, Friberg H, Hoedemaekers C, Horn J, Nolan JP, Rossetti AO, Soar J (2014) 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 40:1816–1831
7. Kamps MJ, Horn J, Oddo M, Fugate JE, Storm C, Cronberg T, Wijman CA, Wu O, Binnekade JM, Hoedemaekers CW (2013) Prognostication of neurologic outcome in cardiac arrest patients after mild therapeutic hypothermia: a meta-analysis of the current literature. Intensive Care Med 39:1671–1682
8. Sandroni C, Cavallaro F, Callaway CW, D'Arrigo S, Sanna T, Kuiper MA, Biancone M, Della Marca G, Farcomeni A, Nolan JP (2013) Predictors of poor neurological outcome in adult comatose survivors of cardiac arrest: a systematic review and meta-analysis. Part 2: patients treated with therapeutic hypothermia. Resuscitation 84:1324–1338
9. Sandroni C, Cavallaro F, Callaway CW, Sanna T, D'Arrigo S, Kuiper M, Della Marca G, Nolan JP (2013) Predictors of poor neurological outcome in adult comatose survivors of cardiac arrest: a systematic review and meta-analysis. Part 1: patients not treated with therapeutic hypothermia. Resuscitation 84:1310–1323
10. Sandroni C, Geocadin RG (2015) Neurological prognostication after cardiac arrest. Curr Opin Crit Care 21:209–214
11. Larson MD, Behrends M (2015) Portable infrared pupillometry: a review. Anesth Analg 120:1242–1253
12. Larson MD, Singh V (2016) Portable infrared pupillometry in critical care. Crit Care 20:161
13. Clark A, Clarke TN, Gregson B, Hooker PN, Chambers IR (2006) Variability in pupil size estimation. Emerg Med J 23:440–441
14. Couret D, Boumaza D, Grisotto C, Triglia T, Pellegrini L, Ocquidant P, Bruder NJ, Velly LJ (2016) Reliability of standard pupillometry practice in neurocritical care: an observational, double-blinded study. Crit Care 20:99
15. Olson DM, Stutzman S, Saju C, Wilson M, Zhao W, Aiyagari V (2016) Interrater reliability of pupillary assessments. Neurocrit Care 24:251–257
16. Heimburger D, Durand M, Gaide-Chevronnay L, Dessertaine G, Moury PH, Bouzat P, Albaladejo P, Payen JF (2016) Quantitative pupillometry and transcranial Doppler measurements in patients treated with hypothermia after cardiac arrest. Resuscitation 103:88–93
17. Solari D, Rossetti AO, Carteron L, Miroz JP, Novy J, Eckert P, Oddo M (2017) Early prediction of coma recovery after cardiac arrest with blinded pupillometry. Ann Neurol 81:804–810
18. Suys T, Bouzat P, Marques-Vidal P, Sala N, Payen JF, Rossetti AO, Oddo M (2014) Automated quantitative pupillometry for the prognostication of coma after cardiac arrest. Neurocrit Care 21:300–308
19. Tamura T, Namiki J, Sugawara Y, Sekine K, Yo K, Kanaya T, Yokobori S, Roberts R, Abe T, Yokota H et al (2018) Quantitative assessment of pupillary light reflex for early prediction of outcomes after out-of-hospital cardiac arrest: a multicentre prospective observational study. Resuscitation 131:108–113
20. Chen JW, Gombart ZJ, Rogers S, Gardiner SK, Cecil S, Bullock RM (2011) Pupillary reactivity as an early indicator of increased intracranial pressure: the introduction of the neurological pupil index. Surg Neurol Int 2:82
21. Chen JW, Vakil-Gilani K, Williamson KL, Cecil S (2014) Infrared pupillometry, the neurological pupil index and unilateral pupillary dilation after traumatic brain injury: implications for treatment paradigms. Springerplus 3:548

22. Rollins MD, Feiner JR, Lee JM, Shah S, Larson M (2014) Pupillary effects of high-dose opioid quantified with infrared pupillometry. Anesthesiology 121:1037–1044

23. Shoyombo I, Aiyagari V, Stutzman SE, Atem F, Hill M, Figueroa SA, Miller C, Howard A, Olson DM (2018) Understanding the relationship between the neurologic pupil index and constriction velocity values. Sci Rep 8:6992

24. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC (2015) STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 351:h5527

25. Rossetti AO, Rabinstein AA, Oddo M (2016) Neurological prognostication of outcome in patients in coma after cardiac arrest. Lancet Neurol 15:597–609

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

27. Nielsen N, Wetterslev J, Cronberg T, Erlinge D, Gasche Y, Hassager C, Horn J, Hovdenes J, Kjaergaard J, Kuiper M et al (2013) Targeted temperature management at 33° C versus 36° C after cardiac arrest. N Engl J Med 369:2197–2206

28. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845

29. Jones SR, Carley S, Harrison M (2003) An introduction to power and sample size estimation. Emerg Med J 20:453–458

30. Zandbergen EG, Hijdra A, Koelman JH, Hart AA, Vos PE, Verbeek MM, de Haan RJ (2006) Prediction of poor outcome within the first 3 days of postanoxic coma. Neurology 66:62–68

31. Bouwes A, Binnekade JM, Kuiper MA, Bosch FH, Zandstra DF, Toornvliet AC, Biemond HS, Kors BM, Koelman JH, Verbeek MM et al (2012) Prognosis of coma after therapeutic hypothermia: a prospective cohort study. Ann Neurol 71:206–212

32. Dragancea I, Horn J, Kuiper M, Friberg H, Ullen S, Wetterslev J, Cranshaw J, Hassager C, Nielsen N, Cronberg T et al (2015) Neurological prognostication after cardiac arrest and targeted temperature management 33° C versus 36° C: results from a randomised controlled clinical trial. Resuscitation 93:164–170

33. Stammet P, Collignon O, Hassager C, Wise MP, Hovdenes J, Aneman A, Horn J, Devaux Y, Erlinge D, Kjaergaard J et al (2015) Neuron-specific enolase as a predictor of death or poor neurological outcome after out-of-hospital cardiac arrest and targeted temperature management at 33° C and 36° C. J Am Coll Cardiol 65:2104–2114

34. Nolan JP, Cariou A (2015) Post-resuscitation care: ERC–ESICM guidelines 2015. Intensive Care Med 41:2204–2206

35. Paul M, Bougouin W, Geri G, Dumas F, Champigneulle B, Legriel S, Charpentier J, Mira JP, Sandroni C, Cariou A (2016) Delayed awakening after cardiac arrest: prevalence and risk factors in the Parisian registry. Intensive Care Med 42:1128–1136