JAMA Neurology | Original Investigation

Serum Neurofilament Light Chain for Prognosis of Outcome After Cardiac Arrest

Marion Moseby-Knappe, MD; Niklas Mattsson, PhD; Niklas Nielsen, PhD; Henrik Zetterberg, PhD; Kaj Blennow, PhD; Josef Dankiewicz, PhD; Irina Dragancea, PhD; Hans Friberg, PhD; Gisela Lilja, PhD; Philip S. Insel, MS; Christian Rylander, PhD; Erik Westhall, PhD; Jesper Kjaergaard, PhD; Matt P. Wise, PhD; Christian Hassager, PhD; Michael A. Kuiper, PhD; Pascal Stammet, PhD; Michael C. Jaeger Wanscher, PhD; Jørn Wetterslev, PhD; David Erlinge, PhD; Janneke Horn, PhD; Tommaso Pellis, MD; Tobias Cronberg, PhD

IMPORTANCE Prognostication of neurologic outcome after cardiac arrest is an important but challenging aspect of patient therapy management in critical care units.

OBJECTIVE To determine whether serum neurofilament light chain (NFL) levels can be used for prognostication of neurologic outcome after cardiac arrest.

DESIGN, SETTING AND PARTICIPANTS Prospective clinical biobank study of data from the randomized Target Temperature Management After Cardiac Arrest trial, an international, multicenter study with 29 participating sites. Patients were included between November 11, 2010, and January 10, 2013. Serum NFL levels were analyzed between August 1 and August 23, 2017, after trial completion. A total of 782 unconscious patients with out-of-hospital cardiac arrest of presumed cardiac origin were eligible.

EXPOSURES Serum NFL concentrations analyzed at 24, 48, and 72 hours after cardiac arrest with an ultrasensitive immunoassay.

MAIN OUTCOMES AND MEASURES Poor neurologic outcome at 6-month follow-up, defined according to the Cerebral Performance Category Scale as cerebral performance category 3 (severe cerebral disability), 4 (coma), or 5 (brain death).

RESULTS Of 782 eligible patients, 65 patients (8.3%) were excluded because of issues with aliquoting, missing sampling, missing outcome, or transport problems of samples. Of the 717 patients included (91.7%), 580 were men (80.9%) and median (interquartile range [IQR]) age was 65 (56-73) years. A total of 360 patients (50.2%) had poor neurologic outcome at 6 months. Median (IQR) serum NFL level was significantly increased in the patients with poor outcome vs good outcome at 24 hours (1426 [299-3577] vs 37 [20-70] pg/mL), 48 hours (3240 [623-8271] vs 46 [26-101] pg/mL), and 72 hours (3344 [845-7838] vs 54 [30-122] pg/mL) ($P < .001$ at all time points), with high overall performance (area under the curve, 0.94-0.95) and high sensitivities at high specificities (eg, 69% sensitivity with 98% specificity at 24 hours). Serum NFL levels had significantly greater performance than the other biochemical serum markers (ie, tau, neuron-specific enolase, and S100). At comparable specificities, serum NFL levels had greater sensitivity for poor outcome compared with routine electroencephalogram, somatosensory-evoked potentials, head computed tomography, and both pupillary and corneal reflexes (ranging from 29.2% to 49.0% greater for serum NFL level).

CONCLUSIONS AND RELEVANCE Findings from this study suggest that the serum NFL level is a highly predictive marker of long-term poor neurologic outcome at 24 hours after cardiac arrest and may be a useful complement to currently available neurologic prognostication methods.

JAMA Neurol. doi:10.1001/jamaneurol.2018.3223
Published online October 29, 2018.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Marion Moseby-Knappe, MD, Department of Clinical Sciences Lund, Neurology, Lund University, Skåne University Hospital, Getingevägen 4 SE-222 41, Lund, Sweden (marion.moseby@med.lu.se).

Prognostication of neurologic outcome after cardiac arrest (CA) should be multimodal and typically include several of the following: clinical neurologic examination, electroencephalogram (EEG), somatosensory-evoked potentials (SSEP), neuroradiologic examination, and biochemical markers.[1,2] Objective and robust methods that combine very high specificity with high sensitivity would be valuable to provide safe and ethical care of patients. Biochemical markers have the potential to be objective measures of neuronal injury. The 2 most commonly studied blood-based biomarkers of brain injury after CA are S100 and neuron-specific enolase (NSE),[3-5] and NSE is the only blood biomarker currently recommended by guidelines for post CA care.[1,2]

A new potential biomarker is neurofilament light chain (NFL), in which elevated levels in cerebrospinal fluid or blood indicate axonal injury in several neurologic diseases.[6-11] One small pilot study found elevated levels of NFL in cerebrospinal fluid,[12] but performing lumbar punctures in patients following CA is impractical, while blood samples are easily collected. Two small pilot studies found that elevated blood levels of neurofilaments corresponded with poor neurologic outcome after CA, when samples were analyzed with standard immunoassays.[13,14]

In this study, we used a novel, ultrasensitive assay with an automated method for quantifying serum NFL levels on the single molecule array (Simoa) platform[15] to test prognostication of neurologic outcome after CA in a large cohort of patients. We compared serum NFL levels with other biomarkers and methods used in clinical practice and recommended in guidelines.

Key Points

Question What is the value of serum neurofilament light chain measurement for prognosis of outcome after cardiac arrest?

Findings In this analysis of biobank data from 717 patients at 29 sites, serum neurofilament light chain levels measured at 24 to 72 hours after cardiac arrest were a highly sensitive and specific marker for poor neurologic outcome 6 months later. The prognostic performance of serum neurofilament light is greater than for established serum biomarkers, head computed tomography, somatosensory-evoked potentials, electroencephalogram, and bedside clinical tests.

Meaning Measurement of serum neurofilament light levels may improve management of care for patients with cardiac arrest.

Methods

The Target Temperature Management After Cardiac Arrest (TTM) trial prospectively included unconscious patients after out-of-hospital CA and randomized them to targeted body temperature management of either 33°C or 36°C.[16] The trial design and main outcomes have been published.[17,18] Twenty-nine sites participated in the biobank part of the TTM trial, with 819 patients included between November 11, 2010, and January 10, 2013. The NFL data were available for 717 patients (eFigure 1 in the Supplement). Formal neurologic prognostication was performed, according to protocol,[18,19] for patients still comatose at a minimum of 108 hours after CA by a physician outside the critical care unit team who was blinded for group allocation as published.[20] This assessment included a clinical neurologic examination. Routine EEG was also part of the study protocol, whereas other investigations, such as computed tomography (CT) or SSEP, were performed on clinical indication typically before formal prognostication.

Written informed consent was waived or obtained from all patients or relatives in line with the Declaration of Helsinki[21] and according to each participating site's national legislation. The trial protocols were approved by ethical committees in each participating country. The Standards for Reporting of Diagnostic Accuracy Studies guidelines were followed.[22]

Exposure

The main exposure was serum NFL level measured in samples collected at 24, 48, and 72 hours after return of spontaneous circulation (ROSC). All samples were preanalytically processed at the site, aliquoted, and frozen to −80°C before shipment to the Integrated BioBank of Luxembourg for batch analysis after trial completion.

Serum NFL level was measured using the NFL assay (Simoa HD-1 analyzer with a Homebrew Kit, Quanterix). Details on the methods have been published.[10] The lower limit of quantification was 2.9 pg/mL, determined as 10 SDs above the mean of the blank signal. Samples were measured at a dilution of one-fourth or, for samples with a very high NFL level, at a dilution of 1/40. The intermediate precision values for 2 quality control samples were 11.8% for a 104-pg/mL quality control sample and 14.1% for a 31-pg/mL quality control sample.

Alternative biochemical markers included NSE and S100, both measured using a Cobas e601 instrument with electrochemiluminescent immunoassays (Roche Diagnostics),[4,5] and tau,[23] measured by the Human Total Tau kit (Quanterix), using the Simoa HD-1 analyzer.[15] Hemolysis was determined with the Cobas system. Board-certified laboratory technicians blinded to all clinical data analyzed all samples using single batches of reagents. All biochemical analyses were carried out after trial completion.

Other examinations included in this study and previously validated in this cohort were bedside clinical tests, including pupillary and corneal reflexes at the time point of prognostication (for each of these, pathologic finding was defined as bilateral absence of the reflex),[20] neurophysiologic tests, including median nerve SSEP[20] (pathologic finding defined as bilaterally absent cortical N20 responses), and routine EEG (pathologic finding defined as highly malignant EEG, with suppression with or without superimposed periodic discharges or burst suppression with or without superimposed discharges, as described previously[19]), and CT (pathologic finding defined as generalized edema[24]).

Outcomes

Neurologic outcome was determined by a face-to-face follow-up at 6 months after CA using the Cerebral Performance Category (CPC) scale.[25] The primary outcome of the study was

good outcome (good cerebral performance [CPC1] or moderate cerebral disability [CPC2]) vs poor outcome (severe cerebral disability [CPC3], coma [CPC4], or brain death [CPC5]). The modified Rankin Scale (mRS), ranging from mRS 0 (asymptomatic) to mRS 6 (death), was used as a secondary outcome.

Statistical Analysis

Bivariate associations were evaluated by Mann-Whitney test and Spearman correlation. Associations between NFL levels and neurologic outcome were tested by logistic regression adjusted for age, sex, and temperature arm. Diagnostic performance for poor outcome was tested with receiver-operating characteristic analysis by calculating the area under the receiver-operating curve (AUROC). Comparisons between paired receiver-operating curves (for comparisons of NFL levels with other biochemical markers) were done by a bootstrap procedure (n = 2000 iterations).

Because high specificity is the most critical metric for use in CA, in which false-positive predictions may lead to the death of the patient, we determined cutoffs at specificities of 100% to 95%. Cutoff points were also defined by the Youden index, which maximizes the sum of sensitivity and specificity.[26] Sensitivities and specificities at all cutoff points were determined by an out-of-sample cross-validation procedure. In each iteration (total n = 2000), we selected participants from 70% of the study sites (randomly chosen) as a training set and the remaining participants from the other sites as a test set. We determined cutoff points in the training set and evaluated them for sensitivity and specificity in the test set. None of the sites and patients used to train the models were therefore used to test the models. The reported sensitivities and specificities are the mean results among the 2000 test sets (with 95% CIs, based on 2.5th and 97.5th quantiles of all iterations).

We compared serum NFL levels with results of neurophysiologic, neuroimaging, and brain stem reflex tests. These comparisons were restricted to patients having data for 24-hour serum NFL level and the alternative method. For each alternative prognostication method, all of which had dichotomous data, we calculated the sensitivity and specificity for poor neurologic outcome. We then defined the serum NFL cutoff level that had a matching specificity and calculated the sensitivity for NFL at that cutoff level.

We compared regression models with 1 or several NFL measurements (24, 48, and 72 hours) as predictors of poor outcome. We compared models with and without NFL level added to clinical information (age, sex, time to ROSC, bystander cardiopulmonary resuscitation [CPR] [yes/no], and serum lactate level at admission) and models with or without neurologic examination, using bilaterally absent corneal reflexes, because this finding has demonstrated the best ability to predict poor outcome among clinical tests.[20] We used the Akaike information criterion (AIC) as a measure of overall model fit for model comparisons.[27] A lower AIC indicates a better model fit. A difference of 2 or more in AIC favors the model with the smallest AIC.

Longitudinal changes in NFL levels within CPC groups were tested in linear mixed-effects models with CPC level as the predictor, adjusted for age and sex (lme4 package, version 1.1-12).[28]

Table 1. Demographic Characteristics of 717 Patients

Characteristic	No. (%)
Age, median (IQR), y	65 (56-73)
Sex	
Men	580 (80.9)
Women	137 (19.1)
Time to ROSC, median (IQR), min	25 (17-39)
Bystander CPR[a]	
Yes	511 (71.3)
No	197 (27.5)
CPC at 6 mo	
1	313 (44.0)
2	44 (6.1)
3	28 (3.9)
4	8 (1.1)
5	324 (45.2)
mRS at 6 mo	
0	143 (19.9)
1	119 (16.6)
2	73 (10.2)
3	27 (3.8)
4	17 (2.4)
5	14 (2.0)
6	324 (45.2)

Abbreviations: CPC, Cerebral Performance Category scale; CPR, cardiopulmonary resuscitation; IQR, interquartile range; mRS, modified Rankin scale; ROSC, return of spontaneous circulation.

[a] Data missing on 9 patients.

Because there were only 8 patients with CPC4 status, we merged the CPC4 and CPC5 groups in this analysis to allow more reliable longitudinal estimates.

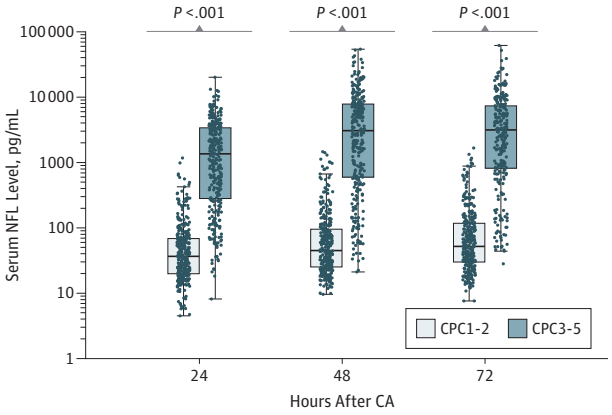
To reduce the skewness of the NFL measurements, we used $\log_{10}$-transformed data. However, for clarity, we present medians and interquartile ranges (IQRs) on the original scale when describing the results. Diagnostics of regression models included standard inspection of residuals, q-q plots, and correlations between residuals and predicted and observed data. Significance was set at 2-tailed $P < .05$. P values were adjusted for multiple comparisons by Hochberg correction.

Statistical analyses were done with SPSS, version 24.0 (IBM Corp) and R, version 3.3.2 (The R Foundation for Statistical Computing).

Results

A total of 717 patients with at least 1 NFL measurement were included (**Table 1**; eFigure 1 in the Supplement), of whom 360 (50.2%) had poor neurologic outcome at 6 months after CA. We found no systematic differences between patients with missing and observed NFL data (eTable 1A in the Supplement). The number of missing samples from patients who had at least 1 NFL level measurement and were alive at missing time points was low (eTable 1B in the Supplement). Of the included patients, 693 had serum NFL levels at 24 hours (96.7%

Figure 1. Serum Neurofilament Light Chain (NFL) Measures by Neurologic Outcome



Medians with interquartile ranges of serum NFL levels in patients with good outcome (Cerebral Performance Category Scale [CPC1-2]) vs poor outcome (CPC3-5) at 6 months for samples taken at 24, 48, and 72 hours. There were no significant differences in serum NFL levels between temperature arms for patients with good outcome.

of all patients), 659 at 48 hours (94.8% of 695 patients alive at 48 hours), and 609 patients at 72 hours (92.3% of 660 patients alive at 72 hours).

Higher serum NFL levels were associated with older age (ρ = 0.28, $P < .001$), longer time to ROSC (ρ = 0.39, $P < .001$), and absence of bystander CPR (median, 82 [IQR, 29-884] pg/mL with bystander CPR vs 356 [57-2406] pg/mL without bystander CPR, $P < .001$), but not with sex ($P = .13$) (eTable 1B in the Supplement). Age, sex, time to ROSC, bystander CPR, or CPC or mRS status at 6 months did not differ significantly between patient groups included in NFL analysis at 24, 48, or 72 hours (eTable 1B in the Supplement) or between the 2 temperature groups (eTable 1C in the Supplement). In patients with poor outcome, serum NFL levels at 72 hours were significantly higher in the 33°C arm (4205 [959-10193] pg/mL) compared with the 36°C arm (2693 [683-6660] pg/mL, $P = 0.02$), but there were no differences at 24 to 48 hours.

There was no significant association between hemolysis and NFL levels sampled at 24 hours or 48 hours, but NFL levels were lower at 72 hours in patients with hemolysis. We therefore performed a sensitivity analysis with all hemolytic samples removed, and all main results remained stable (eFigure 2 in the Supplement).

Serum NFL Levels and Neurologic Outcome

Median (IQR) serum NFL levels were higher in patients with poor outcome compared with patients with good outcome (24 hours: 1426 [299-3577] vs 37 [20-70] pg/mL; 48 hours: 3240 [623-8271] vs 46 [26-101] pg/mL; 72 hours: 3344 [845-7838] vs 54 [30-122] pg/mL) ($P < .001$ at all time points) (**Figure 1**). The associations between serum NFL level and neurologic outcome remained significant when adjusted for age, sex, and temperature arm (eTable 2 in the Supplement). Older age predicted poor neurologic outcome in these multivariable analyses.

We also evaluated each level of the CPC scale individually. The highest serum NFL levels were seen in the CPC4 and CPC5 groups. There were no statistically significant differences between CPC4 and CPC5, but otherwise, each CPC level had a higher NFL concentration than the previous CPC level. Similarly, more advanced levels of the alternative clinical scoring system (ie, mRS) were associated with higher concentrations of NFL (eFigure 3 in the Supplement).

Twenty-nine patients had poor neurologic outcome despite low or only moderately elevated NFL (≤100 pg/mL at 48 hours; the cutoff point 100 pg/mL corresponds to the upper IQR in patients with good outcome) (eTable 6 in the Supplement). Of these patients, 24 died (CPC5). The presumed cause of death was nonneurologic in 17 patients (10 [41.7%] cardiac, 4 [16.7%] multiorgan failure, and 3 [12.5%] other).

Prognostic Performance, Sensitivity, and Specificity of NFL

Serum NFL level had high performance for poor outcome at all time points (AUROC, 0.94) (**Figure 2**A-C). The AUROCs were greater than for the other biochemical markers (tau, NSE, and S100) at all time points ($P < .001$). We found no significant increases in performance when combining NFL level measurements at several time points (eTable 3 in the Supplement). Cutoff levels could be defined with high specificity and still produce relative high sensitivity. For example, the 98% specificity cutoff level for NFL concentration at 24 hours had sensitivity of 69% (95% CI, 57%-79%) (**Table 2** and Figure 2D-F). Youden index cutoff levels had sensitivities of 81% to 82% and specificities of 91% to 93% at the different time points.

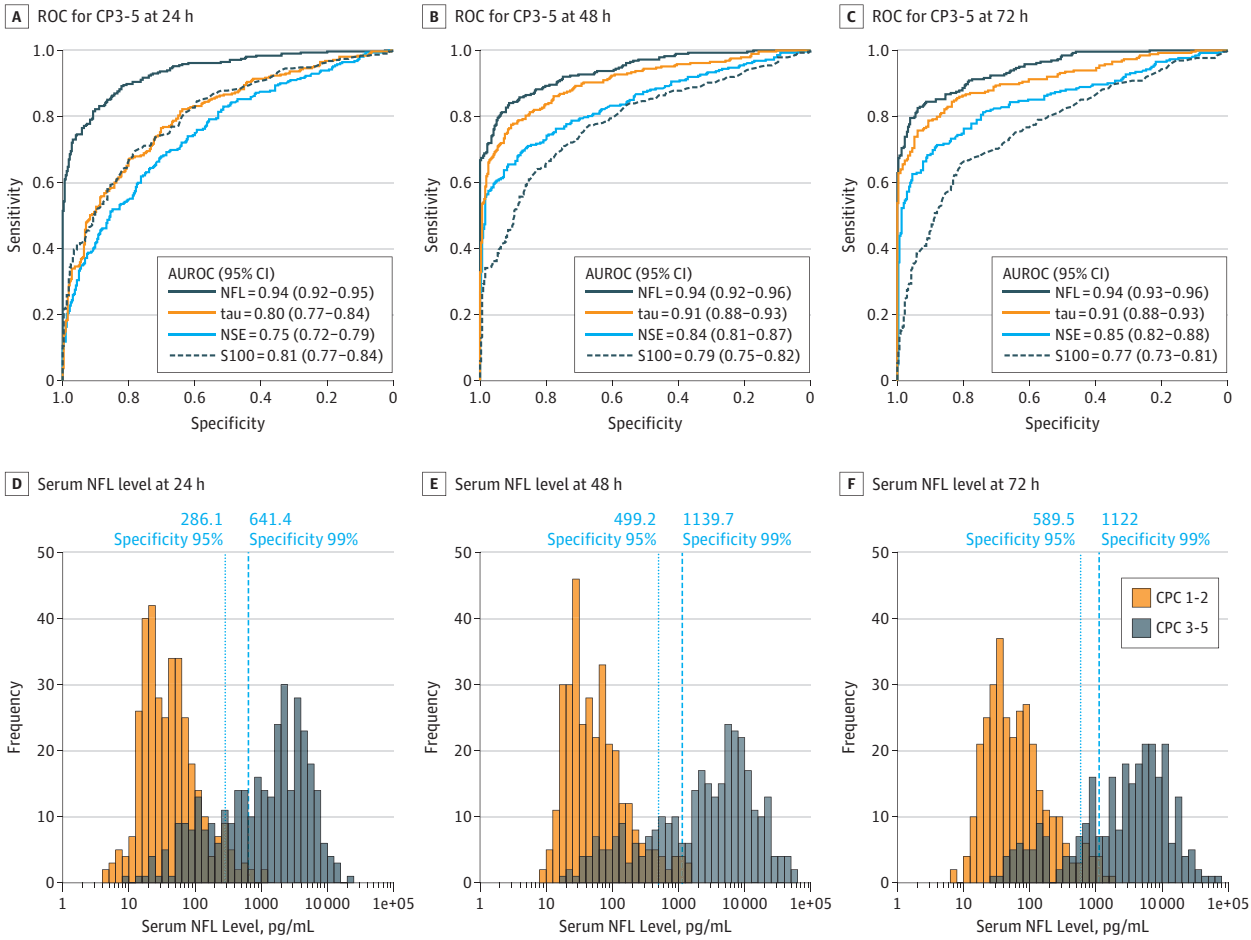
Comparing Serum NFL Levels With Other Prognostication Methods

We compared serum NFL levels with CT, SSEP, routine EEG, and clinical examination of brain stem reflexes (**Table 3**). At a matched specificity, the sensitivity for serum NFL level was always greater than for the competing method (ranging from 29.2% to 49.0% greater for serum NFL level). We used serum NFL level at 24 hours for all of these comparisons.

Adding Serum NFL Level to Clinical Information

Clinical information (age, sex, time to ROSC, bystander CPR [yes/no], and serum lactate level at admission) had moderate performance for poor outcome (AUROC, 0.76-0.78) (eTable 4 in the Supplement). The AUROC increased to 0.96 when the NFL level was added. The effect of adding the NFL level was significant at all time points ($P < .001$). The coefficient for the NFL level changed only marginally when the NFL level was used alone or together with clinical information. The AICs favored models that included both NFL level and clinical information (eTable 4 in the Supplement). We also evaluated the combination of clinical information, bedside neurologic examination (represented by corneal reflexes), and NFL level (eTable 5 in the Supplement). The combination of clinical information and bedside testing had an AUROC value of 0.87, which was improved to 0.98 when NFL level was added. Both NFL level and neurologic examination were significant predictors of poor outcome in the combined model. The AICs fa-

Figure 2. Accuracy of Biomarkers to Predict Poor Neurologic Outcome



A-C, Receiver-operating characteristic (ROC) analyses for prediction of Cerebral Performance Category (CPC)1-2 vs CPC3-5 status 6 months after cardiac arrest for serum samples obtained at 24, 48, and 72 hours. Serum neurofilament light chain (NFL) level always had significantly greater area under the ROC (AUROC) than the other biomarkers ($P < .001$). D-F, CPC1-2 vs CPC3-5 cutoff specificity. These tests were done on participants for whom data were available for all markers (24 hours, n = 641; 48 hours, n = 608; 72 hours, n = 573). NSE indicates neuron-specific enolase.

vored the model that included both NFL level and bedside neurologic examination (eTable 5 in the Supplement).

Longitudinal Serum NFL Level and Neurologic Outcome

Serum NFL level was correlated across the different time points (ρ = 0.94-0.98, $P < .001$). There were no significant changes in NFL level over time in the different CPC groups, except for an increase in NFL level between 24 and 48 hours in the CPC4-CPC5 group (effect of interaction between CPC4-CPC5 and the 48-hour time point: β, 4636; SE, 294; $P < .001$); concentrations at different time points in different CPC groups are seen in eFigure 3 in the Supplement.

Discussion

Serum NFL level measured by an ultrasensitive Simoa assay was a reliable and highly sensitive predictor of poor neurologic outcome at 6 months after CA with a small risk of misclassifying patients with good prognosis. Compared with routine tests for prognostication, serum NFL level had greater sensitivity for poor outcome. The NFL level also differentiated between various degrees of neurologic function according to the CPC and mRS scales, offering unique opportunities to a more nuanced assessment. Taken together, our results indicate that the serum NFL level is a good novel marker for prognosticating long-term neurologic outcome after CA.

The diagnostic performance of the serum NFL level was stable from 24 to 72 hours (AUROC, 0.94) and was slightly further increased when combined with clinical examination (AUROC, 0.96) or bedside neurologic examination (AUROC, 0.98). The performance did not increase significantly when combining serum NFL levels at different time points, which suggests that the trend might not be important for predicting outcome.

When comparing NFL levels with the serum biomarkers NSE, S100, and tau, NFL level was the only marker that predicted poor outcome with high performance at 24 hours post CA, and NFL level remained superior to the other biomarkers at 48 and 72 hours. Currently, NSE is the only blood-based bio-

Table 2. Serum Neurofilament Light Chain (NFL) Specificity, Cutoff Levels, and Sensitivity

Time, h	Specificity (95% CI)[a]	Cutoff Level, pg/mL[b]	Sensitivity (95% CI)[a]	Patients, No. (%)				
				True-Positive	False-Negative	False-Positive	True-Negative	Total
24	1.00 (0.98-1.00)	12 317	0.53 (0.41-0.64)	178 (25.7)	164 (23.7)	0	351 (50.6)	693
	0.99 (0.96-1.00)	641	0.64 (0.53-0.74)	216 (31.2)	126 (18.2)	4 (0.6)	347 (50.0)	693
	0.98 (0.95-1.00)	478	0.69 (0.57-0.79)	237 (34.2)	105 (15.2)	7 (1.0)	344 (49.6)	693
	0.97 (0.94-1.00)	365	0.73 (0.62-0.81)	250 (36.1)	92 (13.3)	11 (1.6)	340 (49.1)	693
	0.96 (0.92-1.00)	300	0.74 (0.63-0.84)	256 (36.9)	86 (12.4)	14 (2.0)	337 (48.6)	693
	0.95 (0.78-0.96)	154	0.75 (0.65-0.85)	257 (37.1)	85 (12.3)	18 (2.6)	333 (48.1)	693
	0.88 (0.90-0.98)	286	0.82 (0.67-0.93)	281 (40.5)	61 (8.8)	34 (4.9)	317 (45.7)	693
48	1.00 (0.98-1.00)	1539	0.65 (0.55-0.74)	208 (31.6)	114 (17.3)	0	337 (51.1)	659
	0.99 (0.95-1.00)	1140	0.67 (0.57-0.75)	214 (32.5)	108 (16.4)	4 (0.6)	333 (50.5)	659
	0.98 (0.93-1.00)	943	0.69 (0.59-0.79)	225 (34.1)	97 (14.7)	7 (1.1)	330 (50.1)	659
	0.97 (0.92-1.00)	808	0.72 (0.61-0.82)	229 (34.7)	93 (14.1)	11 (1.7)	326 (49.5)	659
	0.96 (0.90-1.00)	614	0.75 (0.63-0.85)	242 (36.7)	80 (12.1)	14 (2.1)	323 (49.0)	659
	0.95 (0.89-0.99)	499	0.77 (0.67-0.86)	250 (37.9)	72 (10.9)	17 (2.6)	320 (48.6)	659
	0.91 (0.85-0.97)	269	0.81 (0.70-0.89)	267 (40.5)	55 (8.3)	30 (4.6)	307 (46.6)	659
72	1.00 (0.97-1.00)	1756	0.64 (0.53-0.74)	178 (29.2)	108 (17.7)	0	323 (53.0)	609
	0.99 (0.95-1.00)	1122	0.70 (0.59-0.79)	195 (32.0)	91 (14.9)	4 (0.7)	319 (52.4)	609
	0.98 (0.94-1.00)	963	0.73 (0.62-0.83)	207 (34.0)	79 (13.0)	7 (1.1)	316 (51.9)	609
	0.97 (0.92-1.00)	768	0.76 (0.65-0.85)	218 (35.8)	68 (11.2)	10 (1.6)	313 (51.4)	609
	0.96 (0.91-1.00)	668	0.79 (0.69-0.88)	226 (37.1)	60 (9.9)	13 (2.1)	310 (50.9)	609
	0.95 (0.90-1.00)	590	0.80 (0.71-0.89)	227 (37.3)	59 (9.7)	17 (2.8)	306 (50.2)	609
	0.93 (0.88-0.98)	493	0.81 (0.72-0.90)	234 (38.4)	52 (8.5)	19 (3.1)	304 (50.0)	609

Abbreviation: CPC, Cerebral Performance Category.

[a] Sensitivity and specificity for serum NFL levels (measured at 24, 48, or 72 h) to separate poor outcomes (CPC3-5) from good outcomes (CPC1-2) at 6-mo follow-up. Sensitivities, specificities, and 95% CIs were generated by an out-of-sample cross-validation procedure, which preserved independence of study sites between training and test sets (n = 2000 iterations).

[b] Cutoff levels were identified by the Youden index, which maximizes the combination of sensitivity and specificity and at specificities of 100% to 95%.

chemical marker recommended by European and American guidelines for prognostication,[1,2] but our results indicate that NSE is inferior to NFL. One of the reasons for this difference might be that NFL level is less sensitive to hemolysis than NSE.

The NFL level predicted poor neurologic outcome with higher sensitivity at identical specificities compared with other routine methods of prognostication, including head CT, EEG, SSEP, and absence of pupillary or corneal reflexes. Head CT and SSEP were not part of the TTM trial protocol, but they were performed at the responsible physicians' discretion.[20,24] It is therefore possible that patients with poor prognosis were selected for these tests, which may have inflated the sensitivity of CT and SSEP to detect poor outcome. Despite this possible bias, which would lead to an apparent disadvantage for NFL, NFL testing had superior performance. Furthermore, in those previous prognostication studies, the possibility of a self-fulfilling prophecy of poor outcome cannot be excluded because results of the investigations were available to the treating physicians.[19,20,24] In contrast, measurements of NFL were not done selectively and were performed after trial completion, which excludes both the risk of selection bias and self-fulfilling prophecy.

Little is known about the development of serum NFL levels after CA, especially before 24 hours and after 72 hours. In patients with poor outcome, the median NFL levels nearly doubled from 24 to 48 hours after CA in our cohort, and between 48 and 72 hours, the NFL level seems to have reached steady state. Further research is required to examine the kinetics of NFL beyond the first days after CA. One study in boxers suggests that it may take several weeks or months after a traumatic brain injury before NFL levels are normalized,[29] which should be taken into consideration when evaluating patients with such previous injuries.

Comparisons of NFL levels measured in different studies must be done with caution because studies may differ in preanalytical factors and there is no certified reference standard for the method; therefore, levels may differ between studies, especially if they did not include common samples for calibration. However, in a recent study of parkinsonian diseases,[7] the control group had median NFL levels of 9 pg/mL and, in a study on Alzheimer disease,[8] the control group had a mean NFL level of 35 pg/mL, which is close to the least-impaired patients in the present study (CPC1 or mRS0) (eFigure 3 in the Supplement). Serum or plasma NFL levels in parkinsonian disorders, Alzheimer disease, and cerebral infarctions are relatively low and are therefore unlikely to influence prognostication in patients with CA.[7-9] The only conditions where markedly elevated NFL serum levels have been described previously are amyotrophic lateral sclerosis,[11] extensive traumatic brain injury,[6] and HIV-associated dementia,[10] which may influence the prognostication performance in CA.

Strengths and Limitations

The main strengths of this study include the multicenter design, large sample size, prospective and blinded design, de-

Table 3. Comparing Serum Neurofilament Light Chain (NFL) Levels With Other Prognostication Methods[a]

Method[b]	No. of Patients	Poor Outcome, No. of Patients (%)	Time From CA, Median (IQR), h	Specificity, % (No. of Patients)	Sensitivity, % (No. of Patients)	Serum NFL Level at 24 h (No. of Patients), %[c]		Δ Sensitivity, %
						Cutoff Level, pg/mL	Sensitivity	
CT, generalized edema	261 (Positive, 65; negative, 196)	169 (64.8)	23 (2-91)	97.8 (negative, 90 of 92)	37.3 (positive, 63 of 169)	674	68.6 (positive, 116 of 169)	+31.3
SSEP, bilaterally absent N20-response	170 (Positive, 65; negative, 105)	137 (80.6)	93 (65-117)	97.0 (negative, 32 of 33)	46.7 (positive, 64 of 137)	618	75.9 (positive, 104 of 137)	+29.2
Routine EEG, highly malignant pattern	81 (Positive, 30; negative, 51)	63 (77.8)	67 (51-97)	100 (negative, 18 of 18)	47.6 (positive, 30 of 63)	295	81.0 (positive, 51 of 63)	+33.4
Pupillary reflex, bilaterally absent	245 (Positive, 52; negative,193)	210 (85.7)	119 (96-140)	97.1 (negative, 34 of 35)	24.3 (positive, 51 of 210)	634	73.3 (positive, 154 of 210)	+49.0
Corneal reflex, bilaterally absent	245 (Positive, 79; negative, 166)	210 (85.7)	119 (96-140)	97.1 (negative, 34 of 35)	37.1 (positive, 78 of 210)	634	73.3 (positive, 154 of 210)	+36.2

Abbreviations: CA, cardiac arrest; CT, computed tomography; EEG, electroencephalogram; IQR, interquartile range; SSEP, somatosensory-evoked potentials.

[a] For each comparison, the analysis was restricted to people who had both data for the 24-h serum NFL level and the alternative method (the exact sample therefore differs between each row). Because the number of patient samples for each test differed, the NFL cutoff level for a given specificity varies for each comparison.

[b] Each alternative method had a dichotomous outcome. *Positive* indicates a result associated with worse outcome (ie, generalized edema on CT, bilaterally absent SSEP N20-response, highly malignant EEG pattern, or bilaterally absent corneal reflexes), and *negative* indicates results associated with a better outcome.

[c] Serum NFL level at 24 h vs alternative methods for prognostication of poor outcome (Cerebral Performance Category 3-5) at 6 mo. Serum NFL cutoff levels were determined for matching specificities with each alternative method. For example, 261 patients had both CT data and 24-h serum NFL data, and 65 of these patients had edema on CT, a finding with 97.8% specificity (90 negative CT scans in 92 patients with good outcome) and 37.3% sensitivity (63 positive CT scans in 169 patients with poor outcome). In the same patients, the cutoff level for serum NFL with the same specificity (97.8%) was 674 pg/mL, and that cutoff level had 68.6% sensitivity, representing a difference in sensitivity of 31.3% in favor of serum NFL levels compared with CT.

tailed face-to-face long-term follow-up, and inclusion of several different validated methods of prognostication for comparison with NFL level. Potentially, a measurement of NFL level 24 hours after CA may become an important tool to detect severe hypoxic ischemic brain injury with high performance. The NFL level may therefore be an important complement to existing prognostic methods, which potentially allows for earlier neurologic prognostication in select patients compared with current guidelines.

The study also has limitations. The Simoa instrument has limited availability in routine laboratories. External validation of the specific cutoff levels presented here and establishment of robust laboratory reference ranges are necessary before serum NFL levels can be used for decision making in the care of patients who have undergone CA.

Conclusions

Serum NFL level is an early and highly specific indicator of neurologic outcome in CA. It performs better than other biochemical, clinical, neuroimaging, and electrophysiologic methods.

ARTICLE INFORMATION

Accepted for Publication: August 6, 2018.

Published Online: October 29, 2018. doi:10.1001/jamaneurol.2018.3223

Author Affiliations: Department of Clinical Sciences Lund, Neurology, Lund University, Skåne University Hospital, Lund, Sweden (Moseby-Knappe, Mattsson, Dragancea, Lilja, Cronberg); Clinical Memory Research Unit, Faculty of Medicine, Lund University, Lund, Sweden (Mattsson, Insel); Department of Clinical Sciences Lund, Anesthesia and Intensive Care, Lund University, Helsingborg Hospital, Lund, Sweden (Nielsen); Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden (Zetterberg, Blennow); Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden (Zetterberg, Blennow); Department of Molecular Neuroscience, University College of London Institute of Neurology, London, United Kingdom (Zetterberg); United Kingdom Dementia Research Institute, London, United Kingdom (Zetterberg); Department of Clinical Sciences Lund, Cardiology, Lund University, Skåne University Hospital, Lund, Sweden (Dankiewicz, Erlinge); Department of Clinical Sciences Lund, Anesthesia and Intensive Care, Lund University, Skåne University Hospital, Lund, Sweden (Friberg); Department of Anesthesiology and Intensive Care Medicine, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden (Rylander); Department of Clinical Sciences Lund, Clinical Neurophysiology, Skåne University Hospital, Lund, Sweden (Westhall); Departments of Cardiology, Rigshospitalet and Clinical Medicine, University of Copenhagen, Copenhagen, Denmark (Kjaergaard, Hassager); Adult Critical Care, University Hospital of Wales, Cardiff, United Kingdom (Wise); Department of Intensive Care, Medical Center Leeuwarden, Leeuwarden, the Netherlands (Kuiper); National Rescue Services, Luxembourg City, Luxembourg (Stammet); Department of Cardiothoracic Anaesthesia, The Heart Center, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark (Wanscher); Copenhagen Trial Unit, Centre for Clinical Intervention Research Department, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark (Wetterslev); Department of Intensive Care, Academic Medical Center, Amsterdam, the Netherlands (Horn); Anesthesia and Intensive Care, Card. G. Panico Hospital Agency, Tricase, Italy (Pellis).

Author Contributions: Drs Moseby-Knappe and Mattsson contributed equally to the study, had full access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Moseby-Knappe, Mattsson, Nielsen, Zetterberg, Blennow, Friberg, Lilja, Rylander, Kjaergaard, Hassager, Kuiper, Wanscher, Wetterslev, Erlinge, Horn, Cronberg.
Acquisition, analysis, or interpretation of data: Moseby-Knappe, Mattsson, Nielsen, Zetterberg, Blennow, Dankiewicz, Dragancea, Lilja, Insel, Rylander, Westhall, Kjaergaard, Wise, Hassager, Kuiper, Stammet, Wanscher, Wetterslev, Erlinge, Horn, Pellis.
Drafting of the manuscript: Moseby-Knappe, Mattsson, Cronberg.

Critical revision of the manuscript for important intellectual content: Nielsen, Zetterberg, Blennow, Dankiewicz, Dragancea, Friberg, Lilja, Insel, Rylander, Westhall, Kjaergaard, Wise, Hassager, Kuiper, Stammet, Wanscher, Wetterslev, Erlinge, Horn, Pellis.
Statistical analysis: Moseby-Knappe, Mattsson, Insel.
Obtained funding: Moseby-Knappe, Nielsen, Zetterberg, Blennow, Friberg, Wanscher, Wetterslev, Cronberg.
Administrative, technical, or material support: Moseby-Knappe, Nielsen, Zetterberg, Dankiewicz, Lilja, Westhall, Wise, Hassager, Kuiper, Stammet, Wanscher, Wetterslev, Cronberg.
Supervision: Mattsson, Nielsen, Dragancea, Friberg, Horn, Pellis, Cronberg.

Conflict of Interest Disclosures: Drs Zetterberg and Blennow are cofounders of Brain Biomarker Solutions in Gothenburg AB, a venture-based platform company at the University of Gothenburg. Dr Blennow has served as a paid consultant or served on advisory boards for Alzheon, BioArctic, Biogen, Eli Lilly, Fujirebio Europe, IBL International, Merck, Novartis, Pfizer, and Roche Diagnostics. No other conflicts were reported.

Funding/Support: Funding for the study was provided by the Swedish Research Council (grants 134281, 296161, 286321); Swedish Heart Lung Foundation (grant 20090275); Arbetsmarknadens Försäkringsaktiebolag Insurance Foundation (grant 100001); the Skåne University Hospital Foundations, the Gyllenstierna-Krapperup Foundation, and governmental funding of clinical research within the Swedish National Health System (grants M2010/1837, M2010/1641, 353301); the County Council of Skåne; the Swedish Society of Medicine; the Koch Foundation; TrygFonden (Denmark); European Clinical Research Infrastructures Network; Thelma Zoega Foundation; Stig and Ragna Gorthon Foundation; Thure Carlsson Foundation; Hans-Gabriel and Alice Trolle-Wachtmeister Foundation for Medical Research; Lions Researchfund Skåne; South Swedish Hospital Region Research Funds (Södra Sjukvårdsregionen Regionalt Forsknigsstöd); the Swedish Brain Foundation; and the Torsten Söderberg foundation at the Royal Swedish Academy of Sciences.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: We thank the Target Temperature Management After Cardiac Arrest trial investigators and the laboratory technicians at the Clinical Neurochemistry Laboratory in Mölndal, Sweden for their contribution to this publication.

Additional Information: Samples used in this study were stored and processed at the Integrated BioBank of Luxembourg in compliance with ISO 9001:2008, NF S96-900:2011, and ISO 17025: 2005 standards and International Society for Biological and Environmental Repositories Best Practices.

REFERENCES

1. Callaway CW, Donnino MW, Fink EL, et al. Part 8: post-cardiac arrest care: *2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care*. *Circulation*. 2015;132(18)(suppl 2):S465-S482. doi:10.1161/CIR.0000000000000262

2. Nolan JP, Cariou A. Post-resuscitation care: ERC–ESICM guidelines 2015. *Intensive Care Med*. 2015;41(12):2204-2206. doi:10.1007/s00134-015-4094-5

3. Rossetti AO, Rabinstein AA, Oddo M. Neurological prognostication of outcome in patients in coma after cardiac arrest. *Lancet Neurol*. 2016;15(6):597-609. doi:10.1016/S1474-4422(16)00015-6

4. Stammet P, Collignon O, Hassager C, et al; TTM-Trial Investigators. 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*. 2015;65(19):2104-2114. doi:10.1016/j.jacc.2015.03.538

5. Stammet P, Dankiewicz J, Nielsen N, et al; Target Temperature Management After Out-of-Hospital Cardiac Arrest (TTM) trial investigators. Protein S100 as outcome predictor after out-of-hospital cardiac arrest and targeted temperature management at 33 °C and 36 °C. *Crit Care*. 2017;21(1):153. doi:10.1186/s13054-017-1729-7

6. Shahim P, Gren M, Liman V, et al. Serum neurofilament light protein predicts clinical outcome in traumatic brain injury. *Sci Rep*. 2016;6:36791. doi:10.1038/srep36791

7. Hansson O, Janelidze S, Hall S, et al; Swedish BioFINDER study. Blood-based NfL: a biomarker for differential diagnosis of parkinsonian disorder. *Neurology*. 2017;88(10):930-937. doi:10.1212/WNL.0000000000003680

8. Mattsson N, Andreasson U, Zetterberg H, Blennow K; Alzheimer's Disease Neuroimaging Initiative. Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. *JAMA Neurol*. 2017;74(5):557-566. doi:10.1001/jamaneurol.2016.6117

9. De Marchis GM, Katan M, Barro C, et al. Serum neurofilament light chain in patients with acute cerebrovascular events. *Eur J Neurol*. 2018;25(3):562-568. doi:10.1111/ene.13554

10. Gisslén M, Price RW, Andreasson U, et al. Plasma concentration of the neurofilament light protein (NFL) is a biomarker of CNS injury in HIV infection: a cross-sectional study. *EBioMedicine*. 2015;3:135-140. doi:10.1016/j.ebiom.2015.11.036

11. Feneberg E, Oeckl P, Steinacker P, et al. Multicenter evaluation of neurofilaments in early symptom onset amyotrophic lateral sclerosis. *Neurology*. 2018;90(1):e22-e30. doi:10.1212/WNL.0000000000004761

12. Rosén C, Rosén H, Andreasson U, et al. Cerebrospinal fluid biomarkers in cardiac arrest survivors. *Resuscitation*. 2014;85(2):227-232. doi:10.1016/j.resuscitation.2013.10.032

13. Rana OR, Schröder JW, Baukloh JK, et al. Neurofilament light chain as an early and sensitive predictor of long-term neurological outcome in patients after cardiac arrest. *Int J Cardiol*. 2013;168(2):1322-1327. doi:10.1016/j.ijcard.2012.12.016

14. Rundgren M, Friberg H, Cronberg T, Romner B, Petzold A. Serial soluble neurofilament heavy chain in plasma as a marker of brain injury after cardiac arrest. *Crit Care*. 2012;16(2):R45. doi:10.1186/cc11244

15. Wilson DH, Rissin DM, Kan CW, et al. The Simoa HD-1 Analyzer: a novel fully automated digital immunoassay analyzer with single-molecule sensitivity and multiplexing. *J Lab Autom*. 2016;21(4):533-547. doi:10.1177/2211068215589580

16. Target Temperature Management After Cardiac Arrest. https://clinicaltrials.gov/ct2/show/NCT01020916. Accessed September 22, 2018.

17. Nielsen N, Wetterslev J, Cronberg T, et al; TTM Trial Investigators. Targeted temperature management at 33°C versus 36°C after cardiac arrest. *N Engl J Med*. 2013;369(23):2197-2206. doi:10.1056/NEJMoa1310519

18. Nielsen N, Wetterslev J, al-Subaie N, et al. Target temperature management after out-of-hospital cardiac arrest: a randomized, parallel-group, assessor-blinded clinical trial—rationale and design. *Am Heart J*. 2012;163(4):541-548. doi:10.1016/j.ahj.2012.01.013

19. Westhall E, Rossetti AO, van Rootselaar AF, et al; TTM Trial investigators. Standardized EEG interpretation accurately predicts prognosis after cardiac arrest. *Neurology*. 2016;86(16):1482-1490. doi:10.1212/WNL.0000000000002462

20. Dragancea I, Horn J, Kuiper M, et al; TTM Trial Investigators. Neurological prognostication after cardiac arrest and targeted temperature management 33°C versus 36°C: results from a randomised controlled clinical trial. *Resuscitation*. 2015;93:164-170. doi:10.1016/j.resuscitation.2015.04.013

21. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194. doi:10.1001/jama.2013.281053

22. Bossuyt PM, Reitsma JB, Bruns DE, et al; Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326(7379):41-44. doi:10.1136/bmj.326.7379.41

23. Mattsson N, Zetterberg H, Nielsen N, et al. Serum tau and neurological outcome in cardiac arrest. *Ann Neurol*. 2017;82(5):665-675. doi:10.1002/ana.25067

24. Moseby-Knappe M, Pellis T, Dragancea I, et al; TTM Trial investigators. Head computed tomography for prognostication of poor outcome in comatose patients after cardiac arrest and targeted temperature management. *Resuscitation*. 2017;119:89-94. doi:10.1016/j.resuscitation.2017.06.027

25. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet*. 1975;1(7905):480-484.

26. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35.

27. Burnham KP. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. http://www.springer.com/us/book/9780387953649. Accessed April 12, 2017.

28. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1-48.

29. Shahim P, Zetterberg H, Tegner Y, Blennow K. Serum neurofilament light as a biomarker for mild traumatic brain injury in contact sports. *Neurology*. 2017;88(19):1788-1794. doi:10.1212/WNL.0000000000003912