Ultrasonographic Optic Nerve Sheath Diameter Measurements in Comatose Patients after Cardiac Arrest

Additional Predictive Value and Semi-Automatic Measurements

I.M. Visser

Ultrasonographic Optic Nerve Sheath Diameter Measurements in Comatose Patients after Cardiac Arrest

Additional Predictive Value and Semi-Automatic Measurements

Author

I.M. Visser

A thesis submitted for the degree of Master of Science
Technical Medicine - Medical Sensing and Stimulation Track
Faculty of Science and Technology, University of Twente, Enschede
Graduation internship at Intensive Care, Rijnstate Hospital, Arnhem

Examination committee

Chair	Prof.dr. J. Hofmeijer
Medical supervisor	Dr. M.J. Blans
Technical supervisor	Dr. C.O. Tan
Daily supervisor	M.M.L.H. Verhulst, MSc.
Process supervisor	Drs. P.A. van Katwijk
External member	T. Boers, MSc.

UNIVERSITY OF TWENTE.

Rijnstate

Abstract

Background Neurological outcome prediction after cardiac arrest is possible in only 28-47% of patients. Ultrasonographic measurements of the optic nerve sheath diameter (ONSD) may contribute to neurological outcome prediction. Our first objective was to evaluate the predictive value of ultrasonographic ONSD measurements, in addition to continuous electroencephalography (EEG) measurements, somatosensory evoked potentials (SSEP), and pupillary light reflexes (PLR) for neurological outcome in comatose patients after cardiac arrest. Our second objective was to develop and test a method to semi-automatically measure the ONSD from transorbital sonograms of comatose patients after cardiac arrest.

Methods We performed a prospective observational cohort study in adult comatose patients after cardiac arrest. ONSD was measured on days 1 to 3 using ultrasound. Continuous EEG, SSEP, and PLR were collected as part of standard care. Neurological outcome was classified using the Cerebral Performance Category (CPC) at 6 months (CPC 1-2 = good, CPC 3-5 = poor). For estimation of the additional predictive value of ONSD measurements, logistic regression models predicting neurological outcome were created based on EEG and SSEP, with and without ONSD. The additional predictive value of ONSD measurements was assessed by the increase in sensitivities for a poor (at 100% specificity) and good (at 90% specificity) neurological outcome. Semi-automatic ONSD measurements were performed using an active contour model. Agreement between manual and automatic ONSD measurements was assessed by visual inspection of segmentations, correlation, and Bland-Altman plots. Manual measurements were considered the gold standard.

Results We included 95 patients, of whom 41 (43.2%) died due to post-anoxic encephalopathy. ONSD measured on day 1 was larger in patients with a poor neurological outcome (6.40 [6.15 – 6.88] mm) than in those with a good neurological outcome (6.25 [5.68 – 6.63] mm) (p = 0.023). When adding ONSD measurements to predictions, sensitivity for a poor neurological outcome increased from 25% (95% confidence interval (CI): 0% – 50%) to 45% (95% CI: 25% – 65%) at 100% specificity. Sensitivity for a good neurological outcome raised from 8% (95% CI: 0 % – 23%) to 18% (95% CI: 5% – 36%) at 100% specificity. The PLR was not included as a predictor because of the low incidence of an absent PLR after 72 hours. The active contour model had a mean square error of 4.15 mm^2 and feasibility of 86.1%. A good estimation of the ONSD with an absolute error of ≤ 0.5 mm between manual and automatic ONSD measurements was obtained in 16.7% of sonograms. Bland-Altman plots showed a bias of -0.267 mm and 95% limits agreements between -4.230 and 3.696 mm. Manual and automatic ONSD measurements did not correlate.

Conclusion Ultrasonographic ONSD measurements on days 1 to 3 after cardiac arrest hold the potential to add predictive value in neurological outcome prediction in addition to EEG and SSEP recordings. Our semi-automatic method based on active contours is not suited for automatic ONSD measurements.

List of Abbreviations

AUC	Area under the curve
CI	Confidence interval
CPC	Cerebral performance category
CPR	Cardiopulmonary resuscitation
EEG	Electroencephalography
ICU	Intensive care unit
MSE	Mean square error
PLR	Pupillary light reflex
OHCA	Out-of-hospital cardiac arrest
ONSD	Optic nerve sheath diameter
OR	Odds ratio
pAUC	Partial area under the curve
ROC	Receiver operating characteristic
ROSC	Return of spontaneous circulation
SSEP	Somatosensory evoked potentials
WLST	Withdrawal of life-sustaining therapy

Table of Contents

1. Introduction

Every year, 67-170 per 100,000 Europeans have an out-of-hospital cardiac arrest, of whom ± 8% survive till hospital discharge [1]. Post-anoxic encephalopathy is the main cause of death and long-term disability among patients who are successfully resuscitated from cardiac arrest [2]. The majority of deaths follow the withdrawal of life-sustaining therapy (WLST) when a poor neurological outcome is predicted [3]. Accurate neurological outcome prediction is essential to prevent inappropriate WLST in patients with chances of neurological recovery and futile treatment in patients with no chance of neurologically meaningful survival, and to provide correct information to relatives [4]. Neurological outcome prediction is currently a multimodal approach of clinical examination (ocular reflexes, motor response, and presence of myoclonus), neurophysiology (somatosensory evoked potentials (SSEP) and electroencephalography (EEG)), biomarkers (neuron-specific enolase), and brain imaging (CT and MRI), as no single predictor is 100% accurate [5]. Based on these methods, a reliable prediction of a poor neurological outcome is possible in approximately 28-47% of patients [6–9]. Up to 64% of patients with an indeterminate outcome may have good recovery [7], which highlights the need for novel methods to aid in the prediction of neurological outcome in post-anoxic encephalopathy.

Ultrasonographic measurements of the optic nerve sheath diameter (ONSD) (**Figure 1**) are reported as a non-invasive and bedside tool to predict neurological outcome in comatose patients after cardiac arrest [10–17]. The optic nerve is surrounded by a sheath consisting of three meningeal layers, and its subarachnoid space is continuous with that of the central nervous system [18]. Therefore, elevations of intracranial pressure are transmitted along the optic nerve, dilating the optic nerve sheath [18]. Various studies demonstrated a good correlation between invasive intracranial pressure measurements and ultrasonographic measurements of the ONSD [10,18–25]. Patients with post-anoxic encephalopathy may develop intracranial hypertension as a result of cerebral oedema [26], which is associated with a poor neurological outcome [11–17]. Therefore, ultrasonographic measurements of the ONSD might contribute to neurologic outcome prediction in comatose patients after cardiac arrest.

Figure 1: A transorbital sonogram showing the ultrasound probe, the eye globe and the optic nerve with the optic nerve sheath diameter (ONSD) [18].

The predictive value of the ONSD in addition to established predictors of neurological outcome in post-anoxic encephalopathy is still unclear. Therefore, the first objective of this thesis was to evaluate the predictive value of ultrasonographic ONSD measurements, in addition to continuous EEG measurements, SSEP recordings, and the pupillary light reflex (PLR) for neurological outcome in comatose patients after cardiac arrest.

Considering the small diameter of the optic nerve and its sheath, ranging between 3 to 8 mm, and that a small difference in ONSD can significantly change its interpretation, ONSD measurements need to be precise [27]. Studies on the intra- and interobserver reliability of ultrasonographic ONSD measurements are conflicting [27–34], with intraclass coefficients reported between 0.44 [27] and 0.84 [30]. This emphasizes the need for standardized methods. Automated methods to segment the optic nerve and its sheath from transorbital sonograms have already been reported [35–39], but these methods were tested in only small datasets of a few dozen sonograms. Therefore, the second objective of this thesis was to develop and test a method to semi-automatically measure the ONSD from transorbital sonograms of comatose patients after cardiac arrest, using a large dataset of hundreds of sonograms.

2. Methods

2.1 Study Design

We performed a prospective observational cohort study on the intensive care unit (ICU) of Rijnstate Hospital, The Netherlands. Subsequent comatose patients after cardiac arrest were included for ultrasonographic ONSD measurements on days 1 to 3 after cardiac arrest and followed for six months. Patients were included from December 2019 to October 2021 (Cohort 1) and from May 2022 to October 2022 (Cohort 2). This study was conducted according to the principles of the Declaration of Helsinki. The medical ethics committee Arnhem-Nijmegen approved the study protocol (2019-5586) and waived the need for informed consent before inclusion. In case of patient survival for up to 3 days, consent was obtained from the patient and/or relatives.

2.2 Study Population

Subsequent comatose (Glasgow Coma Scale ≤ 8 at admission) adult patients who were admitted to the ICU after an in-hospital or out-of-hospital cardiac arrest were included. Exclusion criteria were pregnancy, traumatic brain injury, previous relevant eye surgery, pre-existing dependency in daily living (Cerebral Performance Categories (CPC) 3 or 4), and any known progressive brain illness.

2.3 Standard of Care

Patients were monitored and treated according to local guidelines, which were in line with European guidelines for comatose patients after cardiac arrest [5], including targeted temperature management at 36°C for 24 hours, and normothermia for 72 hours. Patients generally received a combination of propofol, midazolam, and morphine for sedation and analgesia. Decisions on WLST were made according to European guidelines [5]. WLST was considered ≥ 72 hours after cardiac arrest during normothermia, and off-sedation. Treating physicians were blinded to the ultrasonographic ONSD measurements.

2.4 Outcome

Neurological outcome was expressed as the CPC at 6 months after cardiac arrest, which was assessed by a standardized telephone interview by the researchers. The CPC at 6 months is the most common scale to quantify the neurological outcome after cardiac arrest [2]. CPC 1 and 2 (no and moderate neurological disability, respectively) were dichotomised as good and CPC 3 − 5 (severe neurologic disability, persistent vegetative state, and (brain) death, respectively) as a poor neurological outcome. In the case of CPC 5, the cause of death was classified as neurological or non-neurological.

2.5 Data Acquisition

2.5.1 EEG, SSEP and PLR Measurements

Continuous EEG recordings were started in all patients as soon as possible after arrival at the ICU, always within 24 hours after cardiac arrest, and continued for at least three days or until the patient's decease or awakening, as part of standard care. Twenty-one electrodes were placed on the scalp according to the international 10-20 system. EEG recordings were performed using a Nihon Kohden system (VCM Medical, The Netherlands) from the start of our study to March 2021 and a BrainRT system (OSG, Belgium) from April 2021 onwards. Two reviewers independently classified EEG epochs at 6, 12, 24, 36, 48, and 72 hours after cardiac arrest, blinded to the timing of the epochs, the patient's clinical status, medication, and neurological outcome. In case of disagreement, a consensus was obtained by consulting a third reviewer. EEG patterns were classified as suppressed with or without superimposed synchronous activity, continuous, or other patterns (**Table A1 in Appendix A**) [8]. SSEP recordings were performed using a Nicolet EDX system (Natus Medical Inc., USA) as part of standard care at the treating physician's request, in case a patient remained comatose after 48 hours and had a motor score ≤ 3 when targeted temperature management and sedation were ceased. The presence of the PLR was tested daily by treating physicians as part of standard care.

2.5.2 Ultrasonographic ONSD Measurements

Trained investigators performed ultrasonographic ONSD measurements on days 1 (0 – 24 hours), 2 (24 – 48 hours) and 3 (48 – 72 hours) after cardiac arrest, or until the patient regained consciousness or died. Three consecutive measurements were performed per eye each day. Sonograms were obtained using the Affinity 70C Ultrasound system equipped with a linear probe with a frequency range of 3-12 MHz (Philips, the Netherlands). Ultrasound gel was applied to the probe and a probe cover was placed over it, preventing the gel from touching the eye. The probe was placed transversally on the superior-lateral part of the closed upper eyelid, angled caudally and medially with the patient's head 30° tilted. No pressure was asserted on the eye. The field depth was set to 4 cm. The optimal view of the optic nerve was sought. The ONSD was measured 3 mm posterior to the retina, as this is the site of maximal pressure changes along the optic nerve [40] (**Figure 2**). The markers were placed at the transition from the hyperechoic retrobulbar fat to the hypoechoic line, in the presence of hyperechoic striped bands, or at the transition from the hyperechoic retrobulbar fat to the hypoechoic region of the optic nerve, in absence of striped bands. These marker placements presumably both correspond to the outer edges of the dura mater [41]. Transorbital sonograms were exported from IntelliSpace Cardiovascular (version 5.1, Philips, The Netherlands) as anonymised DICOM files.
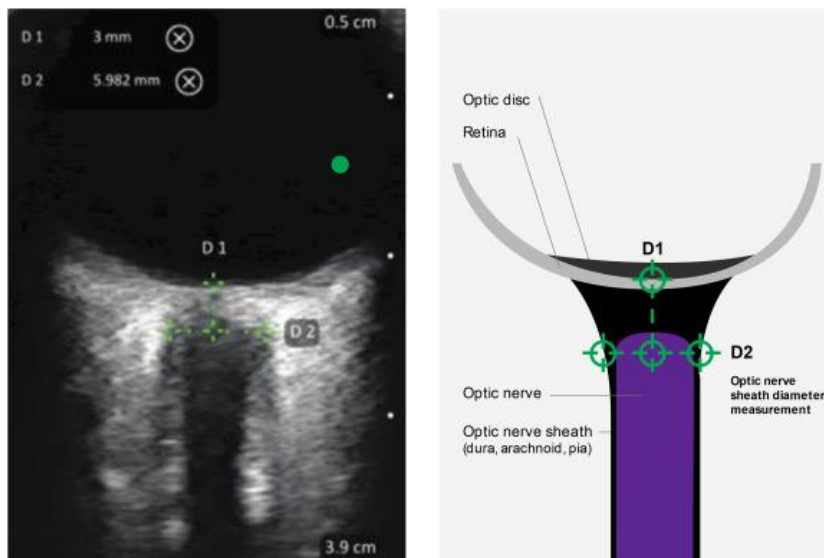

Figure 2. Left panel: Transorbital sonogram showing the optic nerve sheath and surrounding tissues. D1 denotes the distance (3 mm) from the retina to the reference point of the ONSD measurement. D2 denotes the measured ONSD. Right panel: The corresponding image, labelling the anatomical components of the eye for direct comparison with the transorbital sonogram [42]*.*

2.6 Data Analysis

2.6.1 Objective 1: Additional Predictive Value of ONSD

The first objective of this thesis was to evaluate the predictive value of the ultrasonographic ONSD measurements in addition to established predictors for neurological outcome in comatose patients after cardiac arrest. The additional predictive value of ONSD measurements was assessed by creating two logistic regression models, one predicting neurological outcome solely based on established predictors, and one when also adding ONSD measurements. For our primary analysis, we excluded patients with a non-neurological cause of death. A secondary analysis including all patients is presented in **Appendix B**.

2.6.1.1 Outcome and Predictor Variables

The outcome variable was a good (CPC 1-2) or poor (CPC 3-5) neurological outcome. Predictor variables were the EEG, SSEP, PLR, and ONSD on days 1-3. Continuous EEG recordings were categorised as 1) continuous within 12 hours, 2) suppressed with or without superimposed synchronous activity after

24 hours, or 3) other [8] (**Table A1 in Appendix A**). SSEP were categorised as 1) bilaterally absent N20 responses or 2) present N20 responses or unavailable, as an absent N20 response is a highly-specific predictor of a poor neurological outcome, but no predictions about neurological outcome can be done in case of a present N20 response [5]. The PLR was categorised as 1) present or 2) bilaterally absent after 72 hours [5]. The six ONSD measurements obtained per day were averaged to one binocular ONSD.

2.6.1.2 Feature Selection
Predictor variables were selected by a filter feature selection method, which selected features by their statistical relevance to the outcome variable. Thereafter, multicollinearity between ONSD on days 1 to 3 and selected established predictors was measured by the variance inflation factors (VIF). A VIF > 5 indicated high collinearity [43].

2.6.1.3 Mixed-Effects Logistic Regression
To investigate the additional predictive value of ONSD measurements, we trained and validated a logistic regression model, based on the selected established predictors, and a mixed-effects logistic regression model, also including ONSD, in R (version 4.2.2). Patients were split into a training (70% of data) and validation set (30% of data) randomly. Mean binocular ONSD was normalised to z-scores. The mixed-effects logistic regression model with a random intercept was fitted using the BOBYQA optimiser. A random slope or a random intercept and slope model were not achievable due to many missing ONSD measurements. Fixed effects were the selected established predictors, mean binocular ONSD, and a categorical time variable (day 1, 2, or 3). Patient ID was used as a random effect. Predictive values of the models were evaluated using receiver operating characteristics (ROC): area under the curve (AUC), partial area under the curve (pAUC), sensitivity to predict a poor neurological outcome at 100% specificity, and sensitivity to predict a good neurological outcome at > 90% specificity. The pAUC for predicting a poor neurological outcome was calculated as the AUC where specificity was > 99% and the pAUC for predicting a good neurological outcome as the AUC where specificity was > 90%. A 100% specificity for poor neurological outcome prediction was required to prevent WLST in patients with chances of neurological recovery [2]. The additional predictive value of ONSD measurements was quantified as the increase in sensitivity for a poor and good neurological outcome.

2.6.2 Objective 2: Semi-Automatic ONSD Measurement
The second objective of this thesis was to develop and test a method to semi-automatically measure the ONSD from transorbital sonograms of comatose patients after cardiac arrest. This method consisted of two substeps: a semi-automatic segmentation model and an algorithm that automatically calculated the ONSD from the obtained segmentations (**Figure 3**), which were both developed in MATLAB (version R2021b, The MathWorks, Inc., Natick, MA, USA).
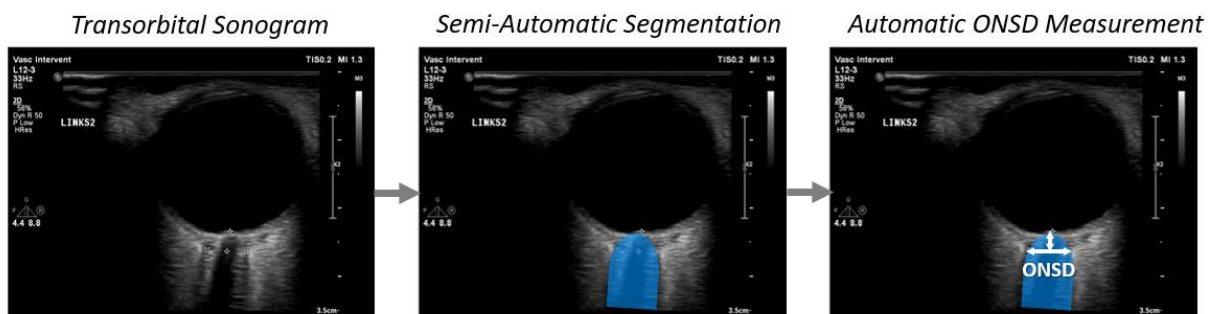

Figure 3: Visualisation of the two substeps for semi-automatic measurement of the optic nerve sheath diameter (ONSD) from transorbital sonograms. The optic nerve and its sheath were segmented from the transorbital sonograms (left panel) by a segmentation algorithm (middle panel) and the ONSD was calculated from the segmentations (right panel).

2.6.2.1 Semi-Automatic Segmentation

Three segmentation methods were developed and tested to segment the optic nerve and its sheath from transorbital sonograms, being the region-growing algorithm, edge detection, and an active contour model. The active contour model [44], also known as 'snakes', was found most suitable to segment the optic nerve and its sheath, when visually inspecting segmentations (**Appendix C**). Therefore, the active contour model was chosen as the algorithm to proceed with. Active contour models are energy-minimising curves that deform to align with image features [44]. The contour deforms over a series of iterations from an initial contour as its energy is affected by internal, image, and external constraint forces (**Appendix D**). These forces are controlled by hyperparameters, that were optimised by performing a grid search.

Transorbital sonograms of cohort 1 were split into a dataset for the grid search (75%) and validation (25%). Transorbital sonograms were split randomly, but the six transorbital sonograms measured successively at one day in one patient were kept together, to prevent bias and maintain enough variability within the grid search set. Cohort 2 was intended as a test set, but due to poor grid search results, validation and testing were disregarded.

The transorbital sonograms were pre-processed to increase the contrast between the hyperechoic optic nerve sheath and the hyperechoic retrobulbar fat by performing standard deviation-based image stretching [45]. This stretching was done by linearly mapping pixel intensity values $\mathbf{x}$ of each sonogram between 0 and $\mathrm{mean}(\mathbf{x}) + n\,\mathrm{std}\,(\mathbf{x})$ to new intensity values between 0 and 1, and setting intensity values $> \mathrm{mean}(\mathbf{x}) + n\,\mathrm{std}\,(\mathbf{x})$ to 1, where an n of 3 gave the best enhanced image (**Figure E1 in Appendix E**). Intensity values near 0 were not mapped to new intensity values, as the black pixels were already dominant in the transorbital sonograms.

An active contour model [46] was adapted to segment the optic nerve and its sheath. The initial position of the active contour was set by manually selecting five to eight points inside the optic nerve and interpolating between these points using a lowpass interpolating filter [47] (**Figure 4**). The initial contour was positioned inside the optic nerve, as a visual inspection revealed that more accurate segmentations were obtained when the contour was grown outwards instead of shrinking towards the optic nerve. The behaviour of the active contour model was affected by fifteen hyperparameters **(Table 1).** Values of ten hyperparameters were tuned by educated guesses. A grid search was performed to tune the other five hyperparameters: the attraction to edges ($\mathrm{w_{edge}}$) and lines ($\mathrm{w_{lines}}$), the stiffness (α), the balloon force (δ), and the timestep (τ). The grid search was run on the AM High Performance Computing Cluster of the University of Twente. Jobs were sent to the Simple Linux Utility for Resource Management scheduler from Windows PowerShell. Data transfer from the local computer to the cluster was executed using FileZilla 3.60.2 using the Secure File Transfer Protocol. The 1024 hyperparameter combinations were evaluated based on the combined results of the active contour model and the algorithm that automatically calculated the ONSD from the segmentations.

2.6.2.2 Automatic Calculation of the ONSD

An algorithm to automatically calculate the ONSD from the segmentations obtained by the active contour model was developed next (**Figure 5**). The algorithm consisted of five steps. First, the centroid of the largest segmented area was calculated (Step 1). Thereafter, a reference point within the eye globe was marked halfway between the ultrasound probe and the centroid (Step 2). The retina was detected from this reference point by selecting the first pixel with an intensity value larger than 0.1 moving vertically (Step 3). The threshold value of > 0.1 was chosen by evaluating the intensity values of the hyperechoic retrobulbar fat in one hundred randomly selected transorbital sonograms of the grid search dataset. One pixel of the retrobulbar fat was selected in one hundred images. The pixel was chosen such that it was representative of the surrounding tissue. A pixel value of 1 (white) was most present (19%), and the other pixels were evenly spread in the range of 0.14 – 0.99 (**Figure E2 in Appendix E**). As the eye globe was black (intensity value 0) in absence of noise, a threshold of > 0.1 was chosen. If the retina was detected, a second reference point was selected 3 mm behind the retina

(Step 4). The number of pixels was related to anatomical distance by retrieving the pixel size from the metadata of the DICOM file. From this second reference point, the left and right boundaries of the segmentation area were detected by moving horizontally (Step 5), from which the ONSD was calculated.
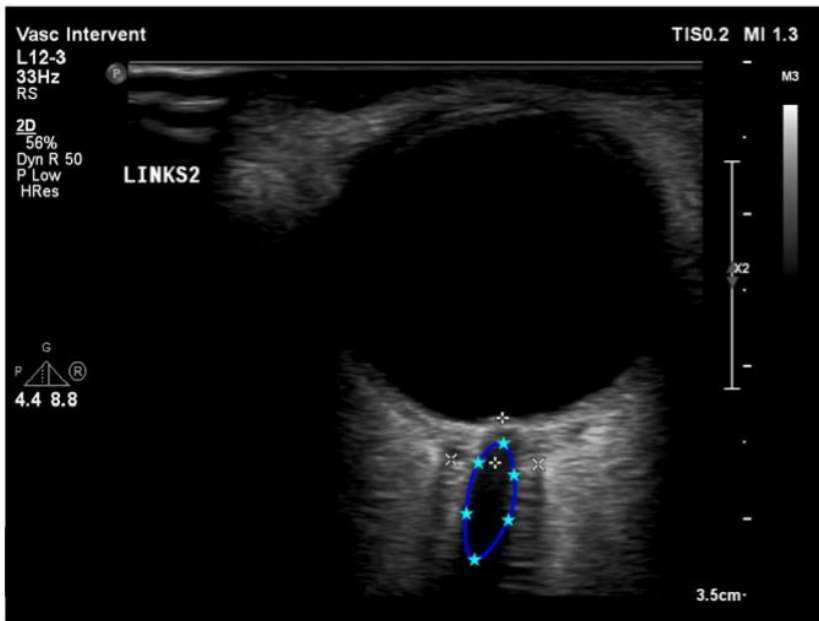

Figure 4: Example of a transorbital sonogram where the initial contour of the active contour model (dark blue contour) was positioned inside the optic nerve by first selecting six points (light blue stars) and thereafter interpolating between these points.
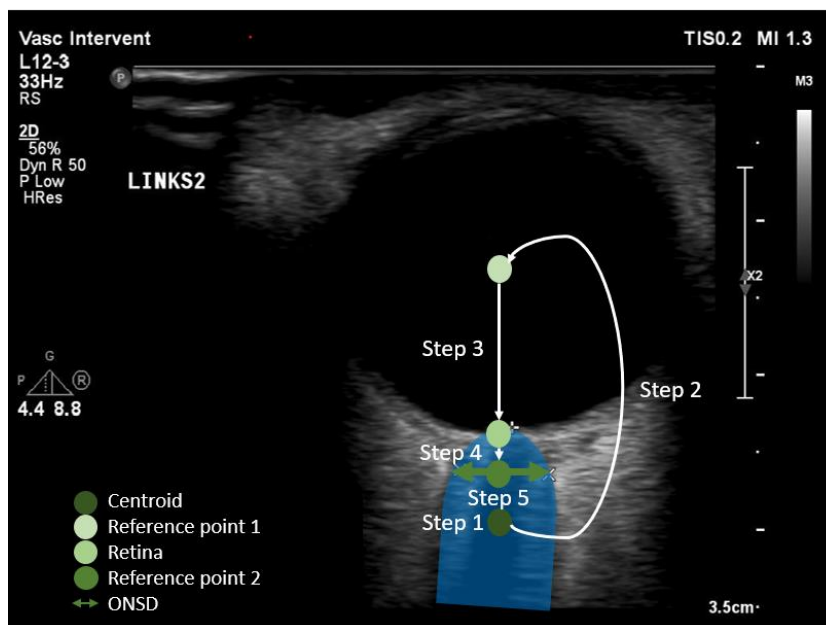

Figure 5: Visualisation of the five consecutive steps to automatically measure the optic nerve sheath diameter (ONSD) from the segmented area of the transorbital sonograms. Step 1: the centroid of the largest segmented area was calculated. Step 2: a reference point within the eye globe was selected halfway between the ultrasound probe and the centroid. Step 3: the retina was searched from this reference point by moving vertically. Step 4: a second reference point was selected 3 mm behind the retina. Step 5: the left and right boundaries of the segmented area were detected by moving horizontally, from which the ONSD was calculated.

Table 1: Hyperparameters of the active contour model.

Grid search		
Parameter	**Description**	**Grid**
w_{edges}	Attraction to edges	[1 2 3 4] *Too large values resulted in a collapse within a local minimum.*
w_{lines}	Attraction to lines	[-0.08 -0.06 -0.04 -0.02] *Negative values attracted the model towards dark lines.*
α	Stiffness	[0.02 0.04 0.06 0.08] *Penalized changes in distances between contour points.*
δ	Balloon force	[0.05 0.10 0.15 0.20] *Inflated the active contour.*
γ	Time step	[0.05 0.2 0.5 1] *Weighed the influence of the current contour position and the image forces on the contour movement. Large γ increased the weight of the current contour position.*
Constant values		
Parameter	**Description**	**Value**
n	Number of contour points	100 *Enough detail was assumed to be obtained with 100 contour points.*
i	Number of iterations	500 *Segmentations did not change largely after 500 iterations.*
β	Smoothness	0.2 *Penalized oscillations in the contour. Segmentations did not change largely when changing β.*
σ_1	Sigma to calculate energy functional for attracting lines using a Gaussian filter	20 *Determined the kernel size of the Gaussian filter and thereby affected the amount of detail [48]. σ was kept unaltered from the original MATLAB script.*
σ_2	Sigma to calculate the gradient of snake energy using a Gaussian filter	20 *Unaltered from the original MATLAB script.*
σ_3	Sigma to calculate the Laplacian in gradient vector flow.	1 *Unaltered from the original MATLAB script, as a good vector field was assumed to be obtained with $\sigma = 1$.*
w_{term}	Attraction to terminations of lines	0 *No line terminations or corners had to be detected.*
κ	Weight of external image force	2 *Trained indirectly by varying the other parameters.*
μ	Regularisation term	*0.2* *A well-smoothed vector field was assumed to be obtained with a μ of 0.2.*
$GI_{\mathrm{iteration}}$	Number of gradient vector flow iterations	200 *A good vector field was assumed to be obtained with 200 iterations.*

Attraction to edges (w_{edges}) and lines (w_{lines}), the stiffness of the contour (α), the balloon force (δ), and the timestep (τ) were tuned by performing a grid search. The other ten hyperparameters were set by educated guesses.

2.6.2.3 Hyperparameter Tuning

The performances of the 1024 hyperparameter combinations were quantified in terms of the mean square error (MSE) between the manually and semi-automatically measured ONSD, and the feasibility, defined as the percentage of transorbital sonograms from which the algorithm was able to calculate an ONSD. The hyperparameter combination with the smallest MSE was chosen as the best one and used for further analysis. We chose to not evaluate the performance of the hyperparameter combinations in terms of segment maps, as the boundaries of the posterior part were often unsharp and only accurate segmentations of the anterior part of the optic nerve were required to calculate the ONSD.

2.6.2.4 Agreement between Manual and Automatic Measurements

The hyperparameter combination with the smallest MSE was thus chosen for further agreement analysis, where the manual ONSD measurements were considered the gold standard. First, a visual inspection of segmentations was done to assess the general performance of the semi-automatic method. Thereafter, the agreement between manual and automatic ONSD measurements was investigated with correlation and Bland-Altman plots. We set a priori acceptable limits of agreement of ± 0.5 mm.

2.6.2.5 Contrast and Histograms

We compared the contrast and histograms between three categories of transorbital sonograms: those in which the ONSD was estimated with an absolute error of ≤ 0.5 mm, underestimated by > 0.5 mm, or overestimated by > 0.5 mm. As only the image regions showing the optic nerve and its surroundings were of interest, we automatically positioned boxes around the optic nerve (**Figure 6**). The box positions were derived from the initial contours. The upper boundary of the box was positioned at the most anterior point of the initial contour, which prevented including the black eye globe within the region. The lower boundary of the box was set as the posterior edge of the ultrasound image, which was the same in sonograms of equal size. The left and right boundaries of the box were set such that the box had a width of 25% of the sonogram and was centred around the centroid of the initial contour. The contrast within the box was calculated by subtracting the average intensity value of the 10% darkest pixels from the average intensity value of the 10% brightest pixels. Moreover, a histogram was calculated for each box. The average histograms were calculated by stacking the individual histograms of each category and dividing them by the number of histograms within each category.

2.7 Statistical Analyses

Statistical analysis was performed using SPSS (version 28, IBM SPSS Statistics, IBM, Armonk, New York, USA), R (version 4.2.2), and MATLAB (version R2021b, The MathWorks, Inc., Natick, MA, USA). Normality was assessed using histograms and Q-Q plots. Data were presented as medians with interquartile ranges, as our data were not normally distributed. Differences between continuous variables were assessed using Mann-Whitney U tests. In the case of three groups, differences between continuous variables were measured with a Friedman test. Differences in categorical variables were tested by Chi-squared or Fishers's exact test in case expected frequencies were below five. Correlation coefficients were calculated using Spearman's rank correlation coefficient. Violin plots were drawn as probability density plots based on a normal kernel function with an optimised bandwidth [49]. Boxplots were drawn according to Tukey's method. Differences between histograms were assessed with Chi-squared tests. Values below $p < 0.05$ were considered statistically significant.
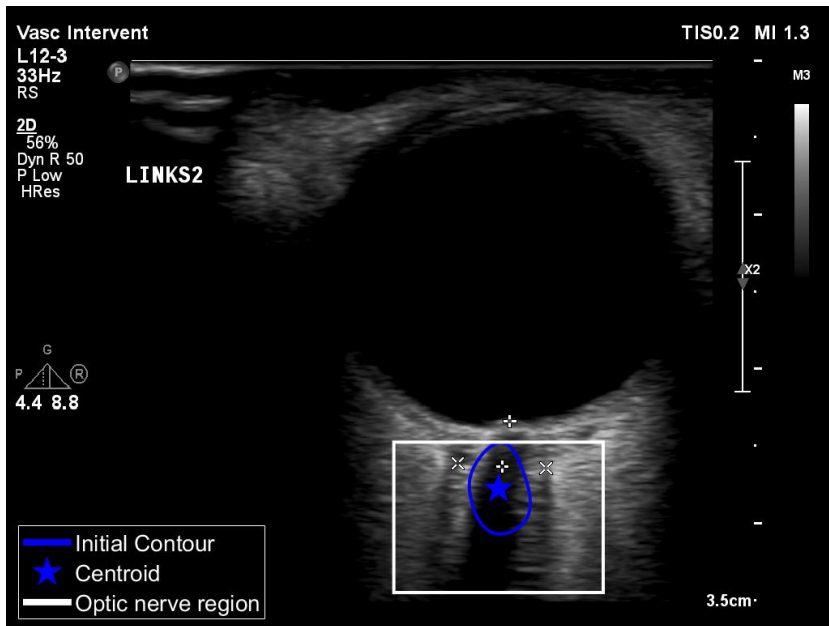
Figure 6: Transorbital sonogram in which the initial contour (blue line), the centroid of the initial contour (blue pentagram), and a region surrounding the optic nerve (white box) are highlighted. The upper boundary was positioned at the most anterior point of the initial contour. The lower boundary was set as the posterior edge of the ultrasound image. The left and right boundaries were set such that the region had a width of 25% of the sonogram and was centred around the centroid of the initial contour.

3. Results

3.1 Study Population

In total, 171 patients who were admitted to the ICU after a cardiac arrest were screened, of whom 95 (56%) were included in this study (**Figure 7**). Baseline characteristics are presented in **Table 2**. A poor outcome was observed in 53 (56%) patients, of whom 41 (77%) died due to post-anoxic encephalopathy. Patients with a good outcome were younger, more frequently had a shockable first cardiac rhythm, and had a shorter time from collapse to the return of spontaneous circulation (ROSC). A continuous EEG pattern within 12 hours (suggestive of a good neurological outcome) was observed in eleven (26%) patients with a good outcome and three (6%) patients with a poor outcome [8]. Two of the patients with poor outcomes and continuous EEG patterns died because of non-neurological causes. Of the patients with a poor outcome, five (9%) had a suppressed EEG pattern with or without superimposed synchronous activity after 24 hours, seven (13%) a bilaterally absent SSEP N20 response after 48 hours, and one (2%) an absent PLR after 72 hours (all suggestive of a poor neurological outcome) [5,8].
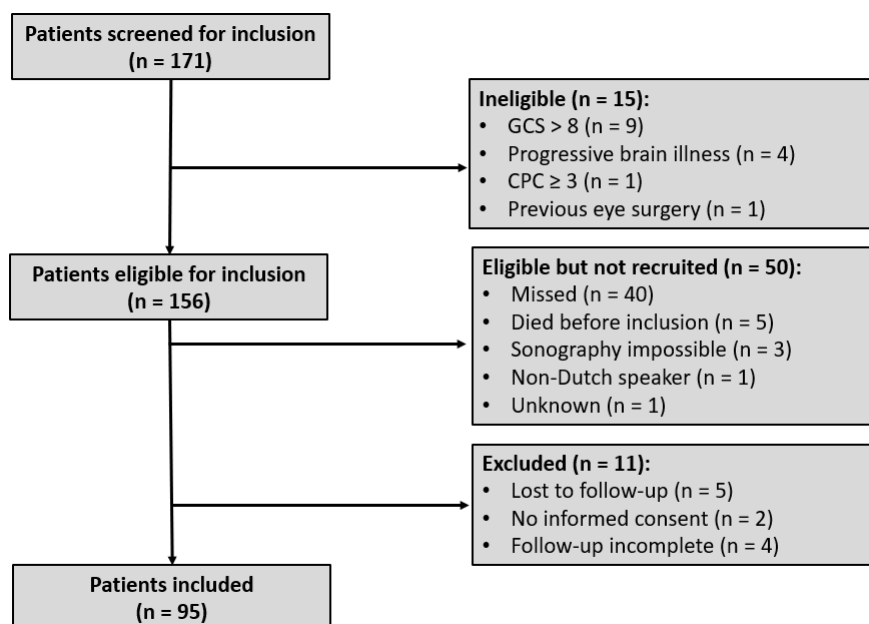

Figure 7: STROBE (Strengthening the Reporting of Observational studies in Epidemiology) flowchart of patient inclusion. GSC: Glasgow Coma Scale, CPC: Cerebral Performance Category.

3.2 Available Data

Transorbital sonography was performed in 72 (75.8%) patients on day 1, 57 (60.0%) patients on day 2, and 39 (41.1%) patients on day 3 (**Table 3**). The most important reasons for missing measurements were regaining consciousness in patients with a good outcome, and decease in patients with a poor outcome. In cohort 1, 858 transorbital sonograms were obtained, whereof 811 (94.5%) were available in IntelliSpace Cardiovascular. Of these sonograms, 612 (75%) were included in the grid search.

Table 2: Baseline characteristics of included patients with a good and poor outcome six months after cardiac arrest.

Characteristic	Good outcome (n = 42)	Poor outcome (n = 53)	p-value
Age (years)	60.0 [52.5 – 67.3]	70.0 [63.5 – 75.5]	p < 0.001*
Female (n)	11 (26%)	16 (30%)	p = 0.668
Bystander CPR (n)	23 (55%)	25 (47%)	p = 0.537
OHCA (n)	42 (100%)	50 (94%)	p = 0.252
Shockable first rhythm (n)	42 (100%)	36 (68%)	p < 0.001*
Cardiac cause (n)	40 (95%)	46 (87%)	p = 0.428
Time from collapse to ROSC (min)	15 [10 – 20]	19 [15 – 32]	p = 0.002*
EEG (n)			p = 0.003*
Continuous < 12 hours	11 (26%)	3 (6%)	
Suppressed > 24 hours	0 (0%)	5 (9%)	
Inconclusive	31 (74%)	45 (85%)	
SSEP (n)			p < 0.001*
Present N20	2 (5%)	20 (38%)	
Absent N20	0 (0%)	7 (13%)	
Not tested	40 (95%)	26 (50%)	
Absent PLR > 72 hours (n)	0 (0%)	1 (2%)	p > 0.999
CPC at 6 months (n)			-
1 – No disability	32 (76%)	0 (0%)	
2 – Moderate disability	10 (24%)	0 (0%)	
3 – Severe disability	0 (0%)	0 (0%)	
4 – Vegetative state	0 (0%)	0 (0%)	
5 – (Brain) death	0 (0%)	53 (100%)	
• Post-anoxic encephalopathy	0 (0%)	41 (77%)	
• Other	0 (0%)	12 (23%)	

*Dichotomous variables are listed as n (%). Continuous variables are listed as median [interquartile range]. Significant differences are indicated by *. CPR: cardiopulmonary resuscitation, OHCA: out-of-hospital cardiac arrest, ROSC: the return of spontaneous circulation, EEG: electroencephalogram, SSEP: somatosensory evoked potentials, PLR: pupillary light reflex, CPC: cerebral performance category.*

Table 3: Available ultrasonographic measurements of the optic nerve sheath diameter (ONSD) on days 1-3 after cardiac arrest in patients with a good and poor outcome.

Outcome	Day 1	Day 2	Day 3
Good (n = 42)	33	24	15
Poor (n = 53)	39	33	24
• Post-anoxic encephalopathy (n = 41)	29	26	21
• Other (n = 12)	10	7	3
Total (n = 95)	**72 (75.8%)**	**57 (60.0%)**	**39 (41.1%)**

3.3 Objective 1: Additional Predictive Value of ONSD

3.3.1 ONSD in Good and Poor Neurological Outcome

Mean binocular ONSD measured on day 1 was larger in patients with a poor neurological outcome (6.40 [6.15 – 6.88] mm) than in those with a good neurological outcome (6.25 [5.68 – 6.63] mm) (p = 0.023) (**Figure 8 and Table A2 in Appendix A**). Mean binocular ONSD measured on days 2 (6.32 [5.84 – 5.56] mm versus 6.51 [6.06 – 6.88] mm, p = 0.130) and 3 (6.45 [5.76 – 6.89] mm versus 6.67 [6.37 – 6.92] mm, p = 0.490) did not differ between patients with good and poor neurological outcomes, respectively.
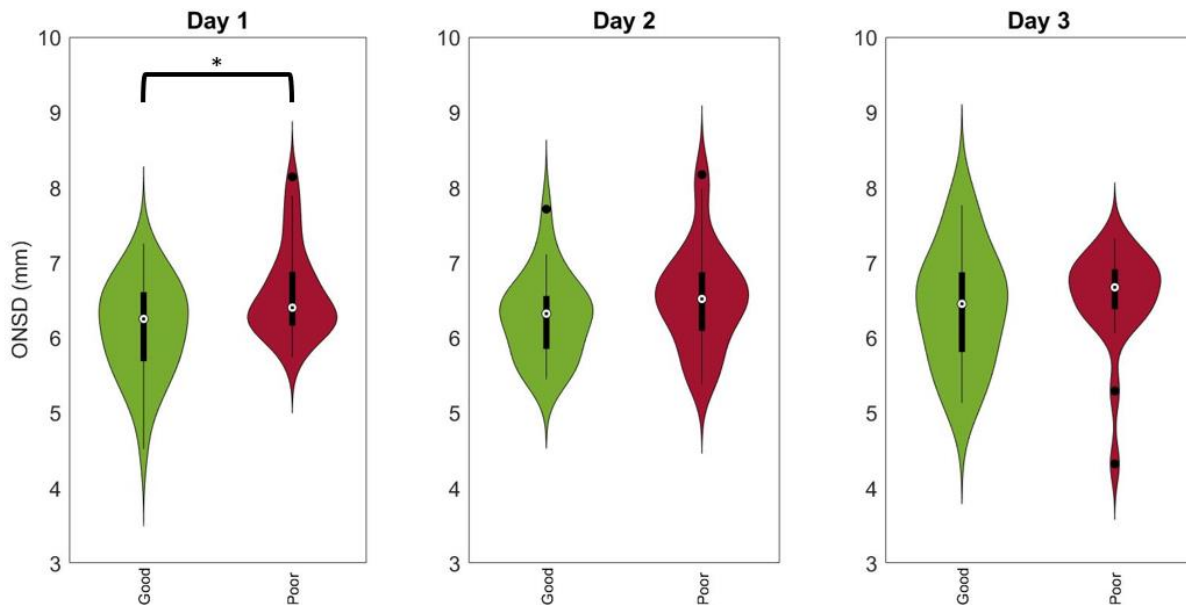
Figure 8: Violin plots of the mean binocular optic nerve sheath diameter (ONSD) measured at day 1 (n = 62), 2 (n = 50), and 3 (n = 36) in patients with a good (in green) and poor (in red) neurological outcome. ONSD on day 1 was larger in patients with a poor neurological outcome (6.40 [6.15 – 6.88] mm) than in those with a good neurological outcome (6.25 [5.68 – 6.63] mm) (p = 0.023). ONSD on days 2 and 3 did not differ between patients with good and poor neurological outcomes. Significant differences are indicated by *.

3.3.2 Feature Selection

Of the established predictors of neurological outcome, EEG pattern and SSEP results were associated with neurological outcome (both p < 0.001), and PLR was not (p > 0.999). Only one (2%) patient with a poor neurological outcome had an absent PLR after 72 hours. Therefore, we chose to only include the EEG and SSEP as predictor variables. The VIFs between ONSD on days 1 to 3 and EEG and SSEP were 1.14, 1.15, and 1.14 respectively, indicating low collinearity.

3.3.3 Predictive Values

When predicting neurological outcomes with a logistic regression model based on EEG and SSEP, a continuous EEG pattern was significantly associated with lower odds of a poor neurological outcome (odds ratio (OR) = 0.100, 95% confidence interval (CI): 5.19e-03 - 0.599, p = 0.029) (**Table 4**). Similarly, a continuous EEG pattern was significantly associated with lower odds of a poor neurological outcome (OR = 0.159, 95% CI: 0.0305 - 0.833, p = 0.029) when predicting neurological outcomes with a mixed-effects logistic regression model based on ONSD, EEG and SSEP (**Table 5**). A suppressed EEG pattern (p = 0.995), an absent SSEP (p = 0.994), and ONSD measurements (p = 0.356) were not significantly associated with neurological outcome.

Table 4: Summary of a logistic regression model predicting neurological outcome using EEG and SSEP results.

Predictor	Estimated coefficient	Odds ratio	95% CI for odds ratio		p-value
			Lower	Upper	
Intercept	5.12e-16	1.00	0.527	1.90	p > 0.999
Suppressed EEG	18.6	1.16e+08	1.16e-154	∞	p = 0.995
Continuous EEG	-2.30	0.100	5.19e-03	0.599	p = 0.036*
Absent SSEP	18.6	1.16e+08	7.688342e-124	∞	p = 0.995

EEG: electroencephalography, SSEP: somatosensory evoked potentials, CI: Confidence interval. Significance is indicated by *.

Table 5: Summary of a mixed-effects logistic regression model predicting neurological outcome using EEG pattern, SSEP results and ONSD measurements.

Predictor	Estimated coefficient	Odds ratio	95% CI for odds ratio		p-value
			Lower	Upper	
Intercept	-0.0834	0.920	0.405	2.09	p = 0.842
Suppressed EEG	17.3	3.39e+07	0.000	∞	p = 0.995
Continuous EEG	-1.84	0.159	0.0305	0.833	p = 0.029*
Absent SSEP	17.6	4.51e+07	0.000	∞	p = 0.994
ONSD (mm)	0.273	1.31	0.736	2.34	p = 0.356
Day 2	0.0887	1.09	0.355	3.37	p = 0.877
Day 3	0.465	1.59	0.475	5.34	p = 0.451

*EEG: electroencephalography, SSEP: somatosensory evoked potentials, ONSD: optic nerve sheath diameter, CI: Confidence interval. Significance is indicated by *.*

Performance measures of the mixed-effects logistic regression and logistic regression are listed in **Table 5**. ROC curves are given in **Figures 9** and **10** for the mixed-effects logistic regression model and the logistic regression model, respectively. When adding ONSD measurements to neurological outcome predictions, sensitivity for a poor neurological outcome increased from 25% (95% CI: 0% – 50%) to 45% (95% CI: 25% – 65%). Sensitivity for a good neurological outcome raised from 7.7% (95% CI: 0% – 23%) to 18% (95% CI: 4.5% – 36%). AUC increased from 0.65 (95% CI: 0.52 – 0.79) to 0.76 (95% CI: 0.62 – 0.92). The pAUC for predicting a poor neurological outcome with a specificity > 99% increased from 0.0025 (95% CI: 0.000059 – 0.005) to 0.0045 (95% CI: 0.0025 – 0.0075), proportional to the sensitivity for poor neurological outcome. The pAUC curve for predicting a good neurological outcome with a specificity > 90% increased from 0.014 (95% CI: 0.0055 – 0.029) to 0.018 (95% CI: 0.0045 – 0.048).

Table 5: Performance measures of the logistic regression model, predicting neurological outcome based on EEG pattern and SSEP results, and the mixed-effects logistic regression model, predicting neurological outcome based on EEG pattern, SSEP results, and ONSD on days 1-3. When adding ONSD measurements to neurological outcome predictions, (p)AUC and sensitivity for both a good and poor neurological outcome increased.

	EEG + SSEP	EEG + SSEP + ONSD
AUC	0.65 (0.52 – 0.79)	0.76 (0.62 – 0.92)
Poor outcome		
Sensitivity (%)	25 (0 – 50)	45 (25 – 65)
Specificity (%)	100 (100 – 100)	100 (100 – 100)
pAUC	0.0025 (0.000059 – 0.005)	0.0045 (0.0025 – 0.0075)
Good outcome		
Sensitivity (%)	7.7 (0.0 – 23)	18 (4.5 – 36)
Specificity (%)	100 (100 – 100)	100 (100 – 100)
pAUC	0.014 (0.0055 – 0.029)	0.018 (0.0045 – 0.048)

Performance measures are presented with 95% confidence intervals. EEG: electroencephalography, SSEP: somatosensory evoked potentials, ONSD: optic nerve sheath diameter, AUC: area under the curve, pAUC: partial area under the curve.

EEG + SSEP

EEG + SSEP + ONSD

Poor: sensitivity 25% (0% – 50%)
Good: sensitivity 7.7% (0% – 23%)
AUC: 0.65 (0.52 – 0.79)

Poor: sensitivity 45% (25% – 65%)
Good: sensitivity 18% (4.5% – 36%)
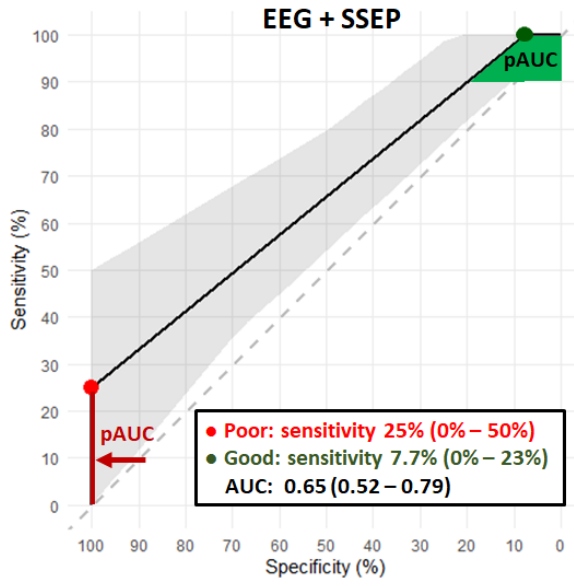AUC: 0.76 (0.62 – 0.92)

Figure 9: ROC curve of the logistic regression model predicting neurological outcomes by the electroencephalography (EEG) pattern and somatosensory evoked potentials (SSEP) results. The red point marks the probability threshold for predicting a poor neurological outcome. The green point marks the probability threshold for predicting a good neurological outcome. The red area is the pAUC for predicting poor outcome with a specificity > 99%, and the green area the pAUC for predicting a good neurological outcome with a specificity > 90%.
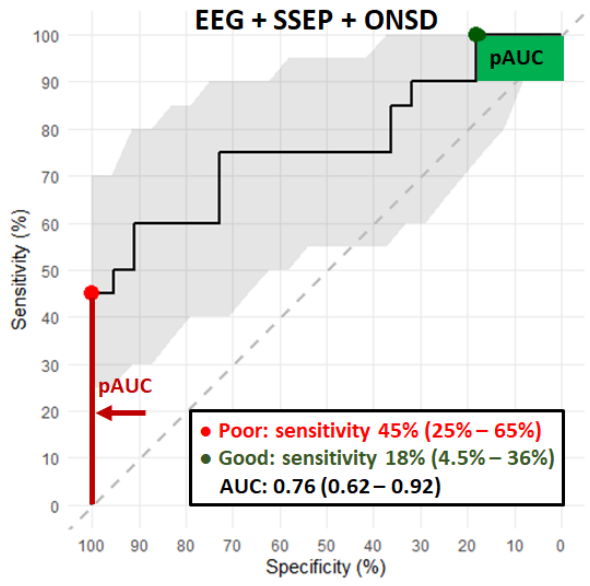
Figure 10: ROC curve of the mixed-effects logistic regression model predicting neurological outcome by electroencephalography (EEG) pattern, somatosensory evoked potentials (SSEP) results and optic nerve sheath diameter (ONSD) measurements. The red point marks the probability threshold for predicting a poor neurological outcome. The green point marks the probability threshold for predicting a good neurological outcome. The red area is the pAUC for predicting poor outcome with a specificity > 99%, and the green area the pAUC for predicting a good neurological outcome with a specificity > 90%.

3.4 Objective 2: Semi-automatic ONSD Measurements

3.4.1 Optimal Hyperparameter Combination

The hyperparameter combination of the active contour model α = 0.08, γ = 1, δ = 0.05, w_{lines} = -0.02, and w_{edges} = 1 gave the smallest MSE of 4.15 mm^2 and had a feasibility of 86.1% (**Figure E3 in Appendix E**). When applying this hyperparameter combination, a good estimation of the ONSD with an absolute error of ≤ 0.5 mm was obtained in 17% of sonograms. An over- or underestimation of the ONSD by >0.5 mm occurred in 46% and 23% of sonograms, respectively. In 14% of sonograms, the method was not able to measure any ONSD from the segmentations.

3.4.2 Visual Inspection of Segmentations and Automatic ONSD Measurements

When visually inspecting the sonograms with an absolute error ≤ 0.5 mm, the ONSD was measured correctly in about 50% of sonograms (**Figure 7, panel A**), but in the other 50%, the segmentation did not overlap well with the optic nerve and its sheath, for example in the case of a horizontal shifting (**Figure 7, panel B**). An underestimation was in most cases caused by: (1) inaccurate detection of the retina because of hyperechoic structures within the hypoechoic eye globe, which led to ONSD measurements too anteriorly, (2) attraction towards retrobulbar fat instead of the optic nerve sheath (**Figure 7, panel C**), resulting in too small segmentations, and (3) insufficient inflation of the contour towards the optic nerve sheath (**Figure 7, panel D**). An overestimation occurred mostly when the contour grew outside the edges between the optic nerve sheath and the retrobulbar fat (**Figure 7, panels E and F**). No estimation of the ONSD was available when the retina was detected too anteriorly,

for example in case of noise (**Figure 7, panel G**), artefacts (**Figure 7, panel H**), or text at the position of the eye globe.

3.3.3 Agreement between Manual and Automatic ONSD Measurements

We found no correlation between the manual and automatic ONSD measurements (r_s = -0.029, p = 0.514) (**Figure 8**). The Bland-Altman plots showed a bias of -0.267 mm and 95% limits agreements between -4.230 and 3.696 mm (**Figure 9**). A negative trend was evident along the Bland-Altman plot (y = 9.05 − 1.42x, R^2 = 0.49, F(1,611) = 507, p < 0.001). **Figure E4 in Appendix E** gives a supplementary Bland-Altman plot with the manual ONSD measurements on the horizontal axis.

3.3.4 Contrast and Histograms

On visual inspection, violin plots of contrast and average histograms did not differ between the transorbital sonograms in which the ONSD was estimated automatically with an absolute error ≤ 0.5 mm, and sonograms in which the ONSD was under- or overestimated by > 0.5 mm. **Appendix F** presents a statistical analysis.
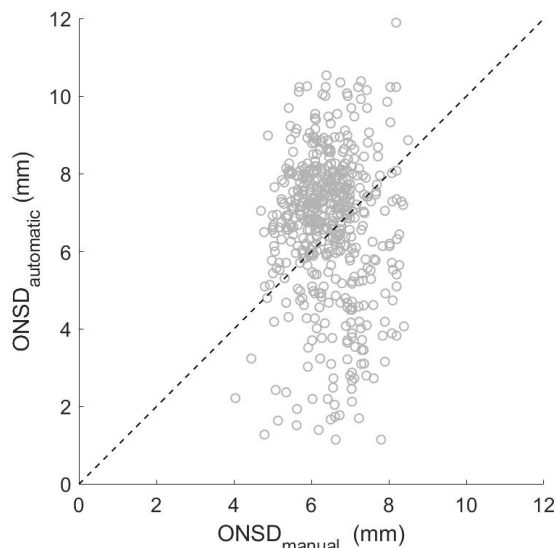

Figure 8: Scatter plot of the manual versus automatic ONSD measurements obtained with the hyperparameter combination of the active contour model (α = 0.08, γ = 1, δ = 0.05, w_{lines} = -0.02, and w_{edges} = 1) giving the smallest mean square error (4.15 mm^2). There was no correlation between manual and automatic ONSD measurements (r_s = -0.029, p = 0.514).
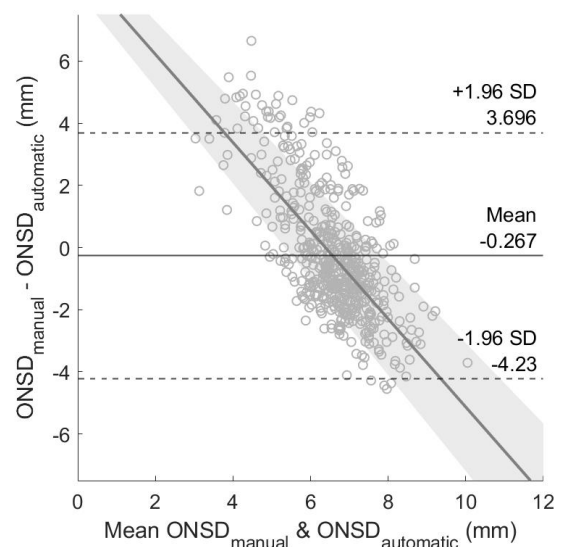
Figure 9: Bland-Altman plots for the manual and automatic ONSD measurements. The grey line represents the regression line (y = 9.05 (8.23 to 9.88) − 1.42 (± 0.12) x). It shows a bias of -0.267 mm and 95% limits agreements between -4.230 and 3.696 mm.
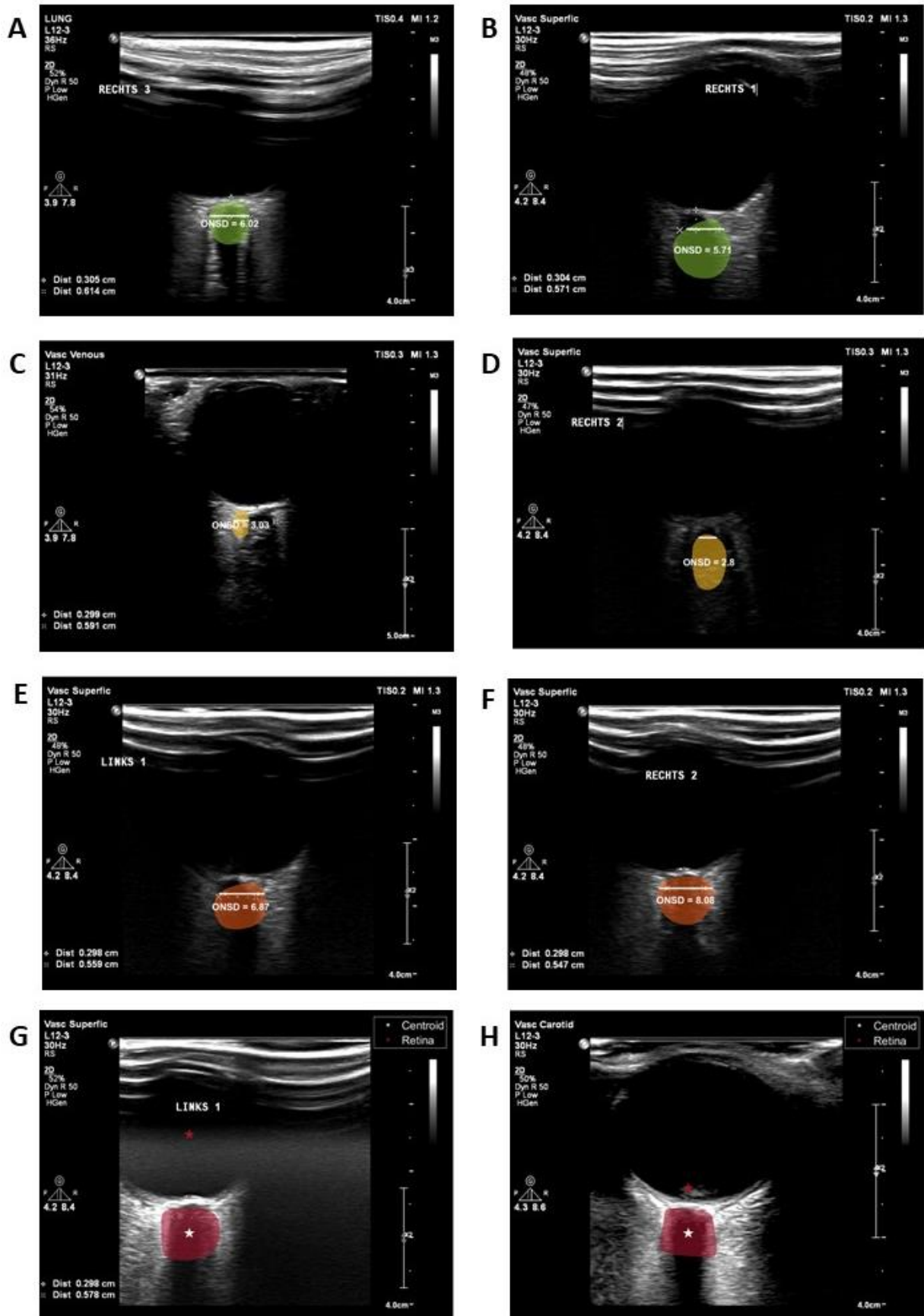
Figure 7: Examples of transorbital sonograms where the ONSD was estimated relatively well with an absolute error of ≤ 0.5 mm (green, A and B), underestimated by > 0.5 mm (yellow, C and D), overestimated by > 0.5 mm (orange, E and F), and no ONSD was estimated (red, G and H).

4. Discussion

4.1 Objective 1: Additional Predictive Value of ONSD

In our cohort of comatose patients after cardiac arrest, we found that sensitivity for a poor neurological outcome increased from 25% to 45% and sensitivity for a good neurological outcome from 8% to 18% when adding ultrasonographic ONSD measurements to continuous EEG and SSEP recordings. Thus, ONSD measurements hold potential to have an additional value in neurological outcome prediction in comatose patients after cardiac arrest.

Our logistic regression model based on established predictors predicted a good neurological outcome with an unexpectedly low sensitivity of 8%, while 26% of all patients with a good neurological outcome had a continuous EEG pattern within 12 hours. This discrepancy was caused by the low incidence of 8% of a continuous EEG pattern in patients with a good neurological outcome within our validation set, despite shuffling our data before splitting data into a training and validation set. The logistic regression model predicted a poor neurological outcome with a sensitivity of 25%, which was comparable with the 29% incidence of an absent SSEP N20 response and a suppressed EEG pattern after 24 hours in all our patients with a poor neurological outcome.

The ONSD measurements were not significantly associated with the odds of a good or poor neurological outcome. Only a continuous EEG pattern was significantly associated with lower odds of a poor neurological outcome in both the logistic regression model and mixed-effects logistic regression model. We expect this to be the result of too little statistical power, as we found significantly larger ONSD in patients with a poor neurological outcome than in those with a good neurological outcome.

As far as we know, this is the first study to investigate the predictive value of ultrasonographic ONSD patients in addition to established predictors in comatose patients after cardiac arrest. Previous studies focused on the predictive value of ONSD measurements alone or in combination with predictors not included in current guidelines [11–15]. These studies reported mean ultrasonographic ONSD measurements between 3.8 and 6.5 mm in patients with a good neurological outcome, and between 5.2 and 7.2 mm in those with a poor neurological outcome. We report median ONSD values of 6.25 and 6.40 mm in patients with good and poor neurological outcome, respectively, at day 1 after cardiac arrest. These ONSD values are comparable with other studies, although our ONSD values were rather large. An explanation for this discrepancy might be differences in ultrasound marker placement between studies [41]. A systematic review reported a pooled sensitivity for poor neurological outcome of 60% (95% CI: 45-73%), and a specificity of 94% (95% CI: 83–98%) [17]. This would imply that ONSD measurements solely can reach higher predictive values than the values that we found by combining ONSD measurements with EEG and SSEP recordings. As our sensitivity increased by 20%, we presume the pooled sensitivity to be an overestimation. We accepted no false positive predictions of a poor neurological outcome, which might have lowered our sensitivity. Furthermore, the sample size within these studies was small, making their results less reliable.

The potential additional predictive value of ONSD measurements can be explained from a pathophysiological perspective. EEG rhythm and SSEP responses reflect synaptic functioning [50]. On the other hand, the ONSD is an indirect measure of intracranial pressure [10,18–25], which is mostly caused by cerebral oedema, which arises within hours after cardiac arrest and is most prominent after approximately 3 to 5 days [51,52]. However, we only found significant differences in ONSD between patients with a good and poor neurological outcome at day 1. An explanation for this finding might be a lack of statistical power, as our primary analysis included 62 patients on day 1, but only 50 and 36 patients on days 2 and 3, respectively. These ONSD measurements were most often missing because patients with a good neurological outcome regained consciousness, and those with a poor outcome had died. As a result, the group size was larger and the range of severity of post-anoxic encephalopathy was probably wider on day 1 compared to days 2 and 3.

4.2 Objective 2: Semi-automatic ONSD Measurements

Our semi-automatic measurements of the ONSD based on active contours did not have good agreement with the manual ONSD measurements. A good estimation of the ONSD was obtained in only a minor portion (17%) of transorbital sonograms. Moreover, the correlation plot revealed a lack of correlation between manual and automatic measurements. This was confirmed by the Bland-Altman plot, which showed wide limits of agreements of approximately ± 4 mm, severely exceeding our a priori set acceptable limits of ± 0.5 mm. So, our semi-automatic method based on active contours is not suited for automatic ONSD measurements.

Besides wide limits of agreements, the Bland-Altman plot also exhibited a noteworthy negative trend. This trend can be explained by the poor performance of the semi-automatic method: overestimated ONSD ($ONSD_{automatic} > ONSD_{manual}$) tended to have a higher mean of the manual and automatic measurements and underestimated ONSD ($ONSD_{automatic} < ONSD_{manual}$) tended to have a lower mean of the manual and automatic measurements.

Our semi-automatic methods might not be appropriate for segmenting the optic nerve and its sheath due to the two different echoic appearances typical for transorbital sonograms. A part of our sonograms displayed the optic nerve as a linear hypoechoic structure surrounded by retrobulbar fat (**Figure 7B**), the other part shows hyperechoic striped bands within the hypoechoic region, which are presumably the subarachnoid space, and a hypoechoic line in between the retrobulbar fat and these bands, most likely being the dura mater (**Figure 7A**) [41]. The aetiology of the presence or absence of these bands is still unclear [41]. In our large dataset, we observed both echoic appearances. These different characteristics might have complicated the grid search to find an optimal hyperparameter combination, as the two types of appearances might require different hyperparameter settings. This idea is supported by the observation that a good segmentation could be obtained when manually adjusting the hyperparameter combinations for individual sonograms.

To the best of our knowledge, five studies have been published on fully-automated algorithms for ultrasonographic ONSD measurements implementing different segmentation methods [35–39]. Stevens et al. [35] developed an algorithm which estimated the ONSD in 39 of 42 sonograms with a mean difference of only −0.08 ± 0.45 and −0.05 ± 0.41 mm in two observers. Their algorithm detected edges between the optic nerve sheath using signed asymmetry features [53], after which they used an active contour model to segment the optic nerve and its sheath. Meiburger et al. [36] implemented a dual snake model, where the mean error between manual and automatic measurements was 0.06 ± 0.52 mm in 71 sonograms. Rajajee et al. [39] used an intensity-based clustering approach to segment the optic nerve and its sheath and obtained a mean difference between manual and automatic measurements of 0.012 ± 0.046 mm in 88 sonograms. Moore et al. [38] segmented the optic nerve and its sheath of an ocular phantom using the watershed algorithm. They reported absolute errors of maximally 1 mm. Gerber et al. [37] used a binary thresholding algorithm on 23 sonograms, that gave correlations of 0.85. To conclude, these authors reported good agreement between manual and automatic ONSD measurements, in contrast to our results. However, their algorithms were tested in only a few dozen sonograms, which might have resulted in less variability within their datasets. Moreover, these studies might have encountered only one type of echoic appearance due to their smaller datasets, hampering the generalisability of their methods to larger datasets including the two types of echoic behaviours. The algorithm proposed by Stevens et al. [35] might be able to overcome the challenges of two different echoic appearances of the optic nerve and its sheath. Their search for edges within the signed asymmetry features could be adapted, for example by starting the search for edges in the retrobulbar fat instead of the middle of the cropped sonogram.

4.3 Limitations and Strengths

Strengths of our study include its prospective design and the broad selection of comatose patients after cardiac arrest, ensuring generalisability to a broad range of clinical settings. Furthermore, we measured ONSD on three consecutive days after cardiac arrest, while most other studies performed measurements only once on day 1. As cerebral oedema typically peaks on day 3 after cardiac arrest [51], testing later measurements seems appropriate.

We acknowledge some limitations. First, even though we included 95 patients, our sample size, on day 3 especially, is small (n = 36). As a result, our CI are wide, making confident conclusions about the additional predictive value of ONSD measurements impossible. Second, the wide time range of measurements for each day (e.g., 2 or 23 hours after cardiac arrest) could have influenced our results, because of the strong time-dependency of cerebral oedema during the first days after cardiac arrest [52]. Third, transorbital sonography was performed by nine sonographers. Despite careful training, underestimations of the ONSD were observed during a visual inspection, especially within the first included patients. This is most likely because of the steep learning curve of ONSD measurements [54]. We did not exclude these inaccurate measurements because of time constraints. As these incorrect measurements were underestimations of the ONSD, they would not lead to false positive predictions of a poor neurological outcome. Fourth, the influence of a self-fulfilling prophecy cannot be fully excluded, which plagues literature on neurological outcome prediction after cardiac arrest in general. However, ONSD measurements were never taken into account in decisions on WLST. Fifth, manual ONSD measurements were considered the ground truth for tuning the active contour model, but the true anatomical ONSD in these patients was obviously unknown. Sixth, we selected predictors that had non-statistically significant coefficients for the logistic regression models, which would generally lead to overfitting. We however decided to still include these predictors, as the relationship between a suppressed EEG pattern and an absent SSEP and a poor neurological outcome is straightforward. The logistic regression model based on these established predictors predicted neurological outcome as would be done by current guidelines [5].

4.4 Future Perspectives

ONSD measurements hold potential to add predictive value to neurological outcome predictions in comatose patients after cardiac arrest. We recommend increasing the sample size for the validation of our results. Our research group aims to include another 65 patients. Cut-off ONSD values for predicting a poor (at 100% specificity) and good (at 90% specificity) neurological outcome should be defined and validated in an external dataset. As of now, cut-off values between 5.11 and 5.75 mm [11–15] are reported, but these were calculated in only small datasets and would lead to many false positive predictions of a poor neurological outcome in our study population. Due to the possibly low intra- and interobserver reliability of ultrasonographic ONSD measurements [27–34], we recommend further developing a fully-automated method to measure the ONSD from transorbital sonograms to eliminate human error. We suggest adapting the asymmetry features-based algorithm developed by Stevens et al. [35] such that it would be applicable to the two different echoic appearances of the optic nerve and its sheath. If testing their algorithm on our dataset, we recommend removing the sonograms with incorrect manual ONSD measurements, which were mentioned before, and cropping sonograms in which the field depth was not set at 4 cm properly. It would also be convenient for clinical applications to incorporate such an automated algorithm in ultrasound machines while performing transorbital sonography, which would require the involvement of manufacturers. If these hurdles are overcome, more patients with poor and good neurological outcome could be identified using a readily available, non-invasive, bedside, fast, inexpensive, and radiation-free tool for intracranial pressure. In this way, the intensity of care in cases of delayed awakening or multi-organ failure can be guided in more comatose patients after cardiac arrest and more relatives can be provided with correct information. Needless to say, even when ONSD measurement would prove to have an additional predictive value, sensitivities of multimodal prediction of neurological outcome after cardiac arrest would remain low. This highlights the ongoing need for the development of additional prognostic tools.

5. Conclusion

In conclusion, ultrasonographic ONSD measurements on days 1 to 3 after cardiac arrest provide a fast and bedside method that holds potential to add predictive value for neurological outcome in addition to EEG and SSEP recordings. Our semi-automatic method based on active contours is not suited for automatic ONSD measurements.

6. References

[1]   Gräsner J-T, Herlitz J, Tjelmeland IBM, Wnent J, Masterson S, Lilja G, et al. European Resuscitation Council Guidelines 2021: Epidemiology of cardiac arrest in Europe. Resuscitation 2021;161:61–79. https://doi.org/10.1016/j.resuscitation.2021.02.007.

[2]   Sandroni C, D'Arrigo S, Nolan JP. Prognostication after cardiac arrest. Critical Care 2018;22:150. https://doi.org/10.1186/s13054-018-2060-7.

[3]   Dragancea I, Wise MP, Al-Subaie N, Cranshaw J, Friberg H, Glover G, et al. Protocol-driven neurological prognostication and withdrawal of life-sustaining therapy after cardiac arrest and targeted temperature management. Resuscitation 2017;117:50–7. https://doi.org/10.1016/j.resuscitation.2017.05.014.

[4]   Sandroni C, D'Arrigo S, Cacciola S, Hoedemaekers CWE, Kamps MJA, Oddo M, et al. Prediction of poor neurological outcome in comatose survivors of cardiac arrest: a systematic review. Intensive Care Medicine 2020;46:1803–51. https://doi.org/10.1007/s00134-020-06198-w.

[5]   Nolan JP, Sandroni C, Böttiger BW, Cariou A, Cronberg T, Friberg H, et al. European Resuscitation Council and European Society of Intensive Care Medicine guidelines 2021: post-resuscitation care. Intensive Care Medicine 2021;47:369–421. https://doi.org/10.1007/s00134-021-06368-4.

[6]   Moseby-Knappe M, Westhall E, Backman S, Mattsson-Carlgren N, Dragancea I, Lybeck A, et al. Performance of a guideline-recommended algorithm for prognostication of poor neurological outcome after cardiac arrest. Intensive Care Medicine 2020;46:1852–62. https://doi.org/10.1007/s00134-020-06080-9.

[7]   Bongiovanni F, Romagnosi F, Barbella G, Di Rocco A, Rossetti AO, Taccone FS, et al. Standardized EEG analysis to reduce the uncertainty of outcome prognostication after cardiac arrest. Intensive Care Medicine 2020;46:963–72. https://doi.org/10.1007/s00134-019-05921-6.

[8]   Ruijter BJ, Tjepkema-Cloostermans MC, Tromp SC, van den Bergh WM, Foudraine NA, Kornips FHM, et al. Early electroencephalography for outcome prediction of postanoxic coma: A prospective cohort study. Annals of Neurology 2019;86:203–14. https://doi.org/10.1002/ana.25518.

[9]   Zhou SE, Maciel CB, Ormseth CH, Beekman R, Gilmore EJ, Greer DM. Distinct predictive values of current neuroprognostic guidelines in post-cardiac arrest patients. Resuscitation 2019;139:343–50. https://doi.org/10.1016/j.resuscitation.2019.03.035.

[10]   Cardim D, Griesdale DE, Ainslie PN, Robba C, Calviello L, Czosnyka M, et al. A comparison of non-invasive versus invasive measures of intracranial pressure in hypoxic ischaemic brain injury after cardiac arrest. Resuscitation 2019;137:221–8. https://doi.org/10.1016/j.resuscitation.2019.01.002.

[11]   Ertl M, Weber S, Hammel G, Schroeder C, Krogias C. Transorbital Sonography for Early Prognostication of Hypoxic-Ischemic Encephalopathy After Cardiac Arrest. Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging 2018;28:542–8. https://doi.org/10.1111/jon.12528.

[12]   Chelly J, Deye N, Guichard JP, Vodovar D, Vong L, Jochmans S, et al. The optic nerve sheath diameter as a useful tool for early prediction of outcome after cardiac arrest: A prospective pilot study. Resuscitation 2016;103:7–13. https://doi.org/10.1016/j.resuscitation.2016.03.006.

[13]   Ueda T, Ishida E, Kojima Y, Yoshikawa S, Yonemoto H. Sonographic Optic Nerve Sheath Diameter: A Simple and Rapid Tool to Assess the Neurologic Prognosis After Cardiac Arrest.

Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging 2015;25:927–30. https://doi.org/10.1111/jon.12246.

[14]   You Y, Park J, Min J, Yoo I, Jeong W, Cho Y, et al. Relationship between time related serum albumin concentration, optic nerve sheath diameter, cerebrospinal fluid pressure, and neurological prognosis in cardiac arrest survivors. Resuscitation 2018;131:42–7. https://doi.org/10.1016/j.resuscitation.2018.08.003.

[15]   Park JS, Cho Y, You Y, Min JH, Jeong W, Ahn HJ, et al. Optimal timing to measure optic nerve sheath diameter as a prognostic predictor in post-cardiac arrest patients treated with targeted temperature management. Resuscitation 2019;143:173–9. https://doi.org/10.1016/j.resuscitation.2019.07.004.

[16]   Lee SH, Jong Yun S. Diagnostic performance of optic nerve sheath diameter for predicting neurologic outcome in post-cardiac arrest patients: A systematic review and meta-analysis. Resuscitation 2019;138:59–67. https://doi.org/https://doi.org/10.1016/j.resuscitation.2019.03.004.

[17]   Zhang YW, Zhang S, Gao H, Li C, Zhang MX. Prognostic Role of Optic Nerve Sheath Diameter for Neurological Outcomes in Post-Cardiac Arrest Patients: A Systematic Review and Meta-Analysis. BioMed Research International 2020;2020:5219367. https://doi.org/10.1155/2020/5219367.

[18]   Dubourg J, Javouhey E, Geeraerts T, Messerer M, Kassai B. Ultrasonography of optic nerve sheath diameter for detection of raised intracranial pressure: a systematic review and meta-analysis. Intensive Care Medicine 2011;37:1059–68. https://doi.org/10.1007/s00134-011-2224-2.

[19]   Jeon JP, Lee SU, Kim S-E, Kang SH, Yang JS, Choi HJ, et al. Correlation of optic nerve sheath diameter with directly measured intracranial pressure in Korean adults using bedside ultrasonography. PloS One 2017;12:e0183170. https://doi.org/10.1371/journal.pone.0183170.

[20]   Kimberly HH, Shah S, Marill K, Noble V. Correlation of optic nerve sheath diameter with direct measurement of intracranial pressure. Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine 2008;15:201–4. https://doi.org/10.1111/j.1553-2712.2007.00031.x.

[21]   Soldatos T, Karakitsos D, Chatzimichail K, Papathanasiou M, Gouliamos A, Karabinis A. Optic nerve sonography in the diagnostic evaluation of adult brain injury. Critical Care (London, England) 2008;12:R67–R67. https://doi.org/10.1186/cc6897.

[22]   Ali MA, Hashmi M, Shamim S, Salam B, Siraj S, Salim B. Correlation of Optic Nerve Sheath Diameter with Direct Measurement of Intracranial Pressure through an External Ventricular Drain. Cureus 2019;11:e5777–e5777. https://doi.org/10.7759/cureus.5777.

[23]   Robba C, Cardim D, Tajsic T, Pietersen J, Bulman M, Donnelly J, et al. Ultrasound non-invasive measurement of intracranial pressure in neurointensive care: A prospective observational study. PLoS Medicine 2017;14:e1002356–e1002356. https://doi.org/10.1371/journal.pmed.1002356.

[24]   Robba C, Santori G, Czosnyka M, Corradi F, Bragazzi N, Padayachy L, et al. Optic nerve sheath diameter measured sonographically as non-invasive estimator of intracranial pressure: a systematic review and meta-analysis. Intensive Care Medicine 2018;44:1284–94. https://doi.org/10.1007/s00134-018-5305-7.

[25]   Geeraerts T, Merceron S, Benhamou D, Vigué B, Duranteau J. Non-invasive assessment of

intracranial pressure using ocular sonography in neurocritical care patients. Intensive Care Medicine 2008;34:2062–7. https://doi.org/10.1007/s00134-008-1149-x.

[26] Sandroni C, Cronberg T, Sekhon M. Brain injury after cardiac arrest: pathophysiology, treatment, and prognosis. Intensive Care Medicine 2021;47:1393–414. https://doi.org/10.1007/s00134-021-06548-2.

[27] Oberfoell S, Murphy D, French A, Trent S, Richards D. Inter-rater Reliability of Sonographic Optic Nerve Sheath Diameter Measurements by Emergency Medicine Physicians. Journal of Ultrasound in Medicine : Official Journal of the American Institute of Ultrasound in Medicine 2017;36:1579–84. https://doi.org/10.7863/ultra.16.05055.

[28] Cimilli Ozturk T, Demir H, Yorulmaz R, Ozdemir S, Isat G, Ecmel Onur O. Assessment of intra-interobserver reliability of the sonographic optic nerve sheath diameter measurement. The Kaohsiung Journal of Medical Sciences 2015;31:432–6. https://doi.org/10.1016/j.kjms.2015.06.004.

[29] Lochner P, Coppo L, Cantello R, Nardone R, Naldi A, Leone MA, et al. Intra- and interobserver reliability of transorbital sonographic assessment of the optic nerve sheath diameter and optic nerve diameter in healthy adults. Journal of Ultrasound 2016;19:41–5. https://doi.org/10.1007/s40477-014-0144-z.

[30] Bäuerle J, Lochner P, Kaps M, Nedelmann M. Intra- and interobsever reliability of sonographic assessment of the optic nerve sheath diameter in healthy adults. Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging 2012;22:42–5. https://doi.org/10.1111/j.1552-6569.2010.00546.x.

[31] Moretti R, Pizzi B, Cassini F, Vivaldi N. Reliability of optic nerve ultrasound for the evaluation of patients with spontaneous intracranial hemorrhage. Neurocritical Care 2009;11:406–10. https://doi.org/10.1007/s12028-009-9250-8.

[32] Ballantyne SA, O'Neill G, Hamilton R, Hollman AS. Observer variation in the sonographic measurement of optic nerve sheath diameter in normal adults. European Journal of Ultrasound 2002;15:145–9. https://doi.org/https://doi.org/10.1016/S0929-8266(02)00036-8.

[33] Shah S, Kimberly H, Marill K, Noble VE. Ultrasound techniques to measure the optic nerve sheath: is a specialized probe necessary? Medical Science Monitor 2009;15:63–8. https://doi.org/https://dx.doi.org/.

[34] Jennings JB, Oliva C, Joyce M, Vitto MJ, Tozer J, Taylor LA, et al. Inter-rater reliability of optic nerve sheath diameter measurement using real-time ultrasonography. The Ultrasound Journal 2022;14:6. https://doi.org/10.1186/s13089-021-00255-1.

[35] Stevens RRF, Huberts W, Gommer ED, Ertl M, Aries M, Mess WH, et al. An Automated Algorithm for Optic Nerve Sheath Diameter Assessment from B-mode Ultrasound Images. Journal of Neuroimaging 2021;31:724–32. https://doi.org/https://doi.org/10.1111/jon.12851.

[36] Meiburger KM, Naldi A, Michielli N, Coppo L, Fassbender K, Molinari F, et al. Automatic Optic Nerve Measurement: A New Tool to Standardize Optic Nerve Assessment in Ultrasound B-Mode Images. Ultrasound in Medicine and Biology 2020;46:1533–44. https://doi.org/10.1016/j.ultrasmedbio.2020.01.034.

[37] Gerber S, Jallais M, Greer H, McCormick M, Montgomery S, Freeman B, et al. Automatic Estimation of the Optic Nerve Sheath Diameter from Ultrasound Images BT - Imaging for Patient-Customized Simulations and Systems for Point-of-Care Ultrasound. In: Cardoso MJ, Arbel T, Tavares JMRS, Aylward S, Li S, Boctor E, et al., editors., Cham: Springer International

Publishing; 2017, p. 113–20.

[38] Moore BT, Montgomery SP, Niethammer M, Greer H, Aylward SR. Automatic Optic Nerve Sheath Measurement in Point-of-Care Ultrasound BT - Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis. In: Hu Y, Licandro R, Noble JA, Hutter J, Aylward S, Melbourne A, et al., editors., Cham: Springer International Publishing; 2020, p. 23–32.

[39] Rajajee MBBS V, Soroushmehr PhD R, Williamson MD, MS CA, Najarian PhD K, Gryak PhD J, Awad MD A, et al. Novel Algorithm for Automated Optic Nerve Sheath Diameter Measurement Using a Clustering Approach. Military Medicine 2021;186:496–501. https://doi.org/10.1093/milmed/usaa231.

[40] Hansen HC, Helmke K. The subarachnoid space surrounding the optic nerves. An ultrasound study of the optic nerve sheath. Surgical and Radiologic Anatomy 1996;18:323–8. https://doi.org/10.1007/BF01627611.

[41] Stevens RRF, Gommer ED, Aries MJH, Ertl M, Mess WH, Huberts W, et al. Optic nerve sheath diameter assessment by neurosonology: A review of methodologic discrepancies. Journal of Neuroimaging 2021;31:814–25. https://doi.org/https://doi.org/10.1111/jon.12906.

[42] Stead GA, Cresswell F V, Jjunju S, Oanh PKN, Thwaites GE, Donovan J. The role of optic nerve sheath diameter ultrasound in brain infection. ENeurologicalSci 2021;23:100330. https://doi.org/https://doi.org/10.1016/j.ensci.2021.100330.

[43] Sheather SJ. 6.4.1 Multicollinearity and Variance Inflation Factors. A modern approach to regression analysis. 1st ed., New York: Springer Science+Business Media; 2009, p. 203. https://doi.org/10.1007/978-0-387-09608-7.

[44] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. International Journal of Computer Vision 1988;1:321–31. https://doi.org/10.1007/BF00133570.

[45] Jain AK. Image Enhancement. Fundamentals of Digital Image Processing, Englewood Cliffs, New Yersey: Prentice Hall; 1989, p. 235–7.

[46] Kroon D-J. Snake: Active Contour. MATLAB Central File Exchange 2022. https://nl.mathworks.com/matlabcentral/fileexchange/28149-snake-active-contour (accessed 21 March 2022).

[47] Digital_Signal_Processing_Committee_of_the_IEEE_Acoustics, _Speech _and_ Signal_Processing_Society. Programs for Digital Signal Processing. New York: IEEE Press; 1979.

[48] Xu C, Prince JL. Gradient vector flow: A new external force for snakes. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1997;2:66–71. https://doi.org/10.1109/cvpr.1997.609299.

[49] Bowman AW, Azzalini A. Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations. 1st ed. Oxford University Press; 1997.

[50] Glimmerveen AB, Ruijter BJ, Keijzer HM, Tjepkema-Cloostermans MC, van Putten MJAM, Hofmeijer J. Association between somatosensory evoked potentials and EEG in comatose patients after cardiac arrest. Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology 2019;130:2026–31. https://doi.org/10.1016/j.clinph.2019.08.022.

[51] Keijzer HM, Hoedemaekers CWE, Meijer FJA, Tonino BAR, Klijn CJM, Hofmeijer J. Brain imaging in comatose survivors of cardiac arrest: Pathophysiological correlates and prognostic properties. Resuscitation 2018;133:124–36.

https://doi.org/10.1016/j.resuscitation.2018.09.012.

[52]   Mlynash M, Campbell DM, Leproust EM, Fischbein NJ, Bammer R, Eyngorn I, et al. Temporal and spatial profile of brain diffusion-weighted MRI after cardiac arrest. Stroke 2010;41:1665–72. https://doi.org/10.1161/STROKEAHA.110.582452.

[53]   Kovesi P. Invariant Measures of Image Features from Phase Information. University of Western Australia, 1996.

[54]   Shrestha GS, Upadhyay B, Shahi A, Jaya Ram KC, Joshi P, Poudyal BS. Sonographic Measurement of Optic Nerve Sheath Diameter: How Steep is the Learning Curve for a Novice Operator? Indian Journal of Critical Care Medicine : Peer-Reviewed, Official Publication of Indian Society of Critical Care Medicine 2018;22:646–9. https://doi.org/10.4103/ijccm.IJCCM_104_18.

[55]   J. Ivins, Porrill J. Everything you always wanted to know about snakes (but were afraid to ask). 2000.

[56]   Kroon   D.   Region   Growing.   MATLAB   Central   File   Exchange   2022. https://nl.mathworks.com/matlabcentral/fileexchange/19084-region-growing (accessed 25 April 2022).

[57]   Canny J. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and   Machine   Intelligence   1986;PAMI-8:679–98. https://doi.org/10.1109/TPAMI.1986.4767851.

Appendix A: Tables

Table A1: EEG categories classified at 6, 12, 24, 36, 48, and 72 hours after cardiac arrest. EEG patterns were classified as suppressed with or without superimposed synchronous activity (associated with poor neurological outcome > 24 hours after cardiac arrest), continuous (associated with good neurological outcome < 12 hours after cardiac arrest), or other patterns (association with neurological outcome unclear).

Electroencephalography patterns
Suppressed with or without superimposed synchronous activity - Suppressed (maximum amplitude < 10 µV) - Burst-suppression (synchronous) with generalized, abrupt-onset bursts or identical bursts with suppressed background activity - Generalized periodic discharges with suppressed background activity
Continuous or nearly continuous activity: maximum amplitude ≥ 20 µV, < 10% suppressions* - Delta (dominant frequency < 4 Hz) - Theta (dominant frequency 4-8 Hz) - Alpha (dominant frequency > 8 Hz)
Other patterns - Low-voltage (maximum amplitude 10-20 µV) - Epileptiform on other background - Burst-suppression (heterogeneous, ≥ 50% suppressions*) - Discontinuous (10-49% suppressions*)

* Suppressions are defined as segments with amplitude < 10 µV, or segments with amplitude ≥ 10 µV, but <50% of background/burst voltage.

Table A2: Mean binocular optic nerve sheath diameters (ONSD) measured at days 1,2, and 3 in comatose patients after cardiac arrest with a good and poor neurological outcome. Data are presented as median with lower and upper quartiles.

	Good neurological outcome (n = 42)	Poor neurological outcome (n = 41)	p-value
ONSD$_{Day1}$ (mm)	6.25 [5.68 – 6.63]	6.40 [6.15 – 6.88]	p = 0.023*
ONSD$_{Day2}$ (mm)	6.32 [5.84 – 5.56]	6.51 [6.06 – 6.88]	p = 0.130
ONSD$_{Day3}$ (mm)	6.45 [5.76 – 6.89]	6.67 [6.37 – 6.92]	p = 0.490

Appendix B: Secondary Analysis Including All Patients

Mean binocular optic nerve sheath diameter (ONSD) measured on day 1 (6.25 [5.68 – 6.63] mm) versus 6.26 [6.10 – 6.87], p = 0.092), day 2 (6.32 [5.84 – 5.56] mm versus 6.50 [6.04 – 6.84] mm, p = 0.169) and 3 (6.45 [5.76 – 6.89] mm versus 6.56 [6.16 – 6.90] mm, p = 0.862) did not differ between patients with good and poor outcomes, respectively (**Table B1 and Figure B1**).

Table B1: Mean binocular optic nerve sheath diameters (ONSD) measured at days 1,2, and 3 in comatose patients after cardiac arrest who had a good and poor outcome. Data are presented as median with lower and upper quartiles.

	Good outcome (n = 42)	Poor outcome (n = 53)	p-value
$ONSD_{Day1}$ (mm)	6.25 [5.68 – 6.63]	6.26 [6.10 – 6.87]	p = 0.092
$ONSD_{Day2}$ (mm)	6.32 [5.84 – 5.56]	6.50 [6.04 – 6.84]	p = 0.169
$ONSD_{Day3}$ (mm)	6.45 [5.76 – 6.89]	6.56 [6.16 – 6.90]	p = 0.862
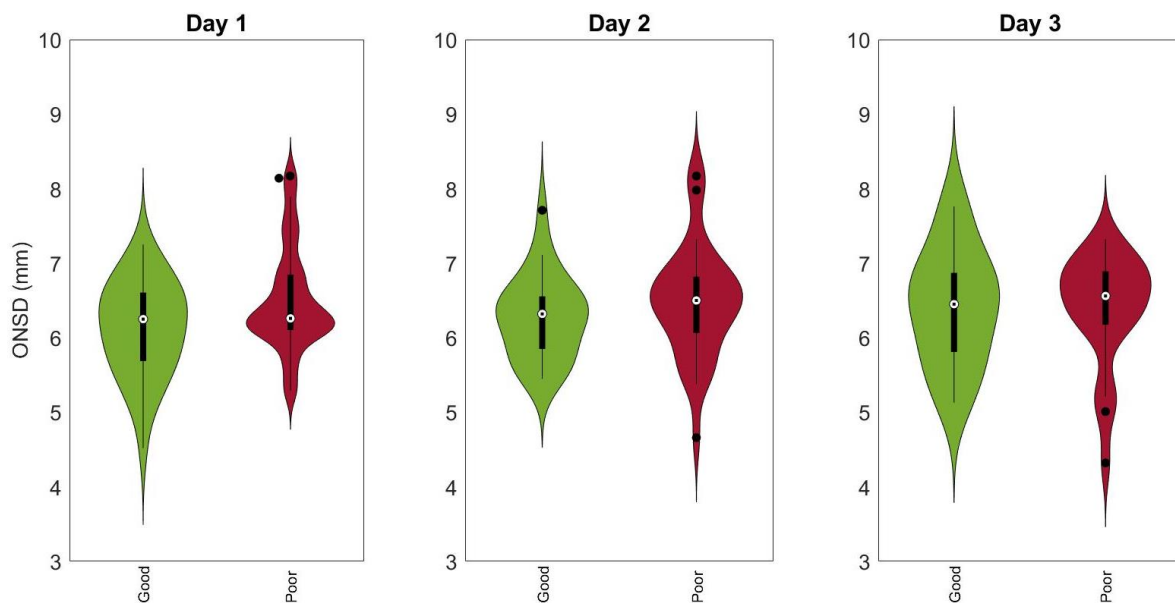
Figure B1: Violin plots of the mean binocular optic nerve sheath diameter (ONSD) measured at day 1 (n = 72), 2 (n = 57), and 3 (n = 39) in patients with a good (in green) and poor (in red) outcome after cardiac arrest. ONSD on days 1, 2 and 3 did not differ between patients with good and poor neurological outcomes. Boxplots were drawn according to Tukey's method.

Of the established predictors of neurological outcome, electroencephalography (EEG) pattern and somatosensory evoked potentials (SSEP) results were associated with outcome (both p < 0.001), and the pupillary light reflex (PLR) was not (p > 0.999). The variance inflation factors (VIF) between ONSD on days 1 to 3 and EEG and SSEP were 1.08, 1.11, and 1.22 respectively, indicating low collinearity.

When predicting outcomes using the EEG pattern and SSEP recordings with a logistic regression model, a continuous EEG pattern was significantly associated with lower odds of a poor outcome (odds ratio (OR) = 0.146, 95% confidence interval (CI): 0.0208 - 0.634, p = 0.021) (**Table B2**). Similarly, a continuous EEG pattern was significantly associated with lower odds of a poor outcome (OR = 0.196, 95% CI: 0.0480 - 0.800, p = 0.023) when predicting outcomes with a mixed-effects logistic regression model using the ONSD measurements on days 1 to 3, EEG pattern and SSEP recordings (**Table B3**). A suppressed EEG (p = 0.995), an absent N20 response (p = 0.994), and ONSD measurements (p = 0.740) were not significantly associated with outcome.

Table B2: Summary of logistic regression model predicting outcome using EEG pattern and SSEP results.

Predictor	Estimated coefficient	Odds ratio	95% CI for odds ratio		p-value
			Lower	Upper	
Intercept	0.314	1.36	0.761	2.50	0.299
Suppressed EEG	18.3	8.45e+7	7.62e-155	NA	0.995
Continuous EEG	-1.92	0.146	0.0208	0.634	**0.021***
Absent SSEP	18.3	8.45e+7	5.09e-124	NA	0.995

EEG: electroencephalography, SSEP: somatosensory evoked potentials, CI: Confidence interval. Significance is indicated by *.

Table B3: Summary of mixed-effects logistic regression model predicting outcome using EEG pattern, SSEP results and ONSD measurements on days 1 to 3.

Predictor	Estimated coefficient	Odds ratio	95% CI for odds ratio		p-value
			Lower	Upper	
Intercept	0.234	1.26	0.609	2.62	0.530
Suppressed EEG	17.1	2.65e+7	0	∞	0.995
Continuous EEG	-1.63	0.196	0.0480	0.800	**0.023***
Absent SSEP	17.2	2.95e+7	0	∞	0.994
Day 2	0.114	1.12	0.480	3.08	0.826
Day 3	0.443	1.56	0.505	4.81	0.441
ONSD (mm)	0.0939	1.10	0.631	1.91	0.740

EEG: electroencephalography, SSEP: somatosensory evoked potentials, CI: Confidence interval. Significance is indicated by *.

Performance measures of the mixed-effects logistic regression and logistic regression are listed in **Table B4**. ROC curves are given in **Figures B2** and **B3**. When adding ONSD measurements to neurological outcome predictions, sensitivity for a poor outcome increased from 19% (95% CI: 0% – 38%) to 27% (95% CI: 11% – 46%). Sensitivity for a good outcome raised from 7.6% (95% CI: 0%– 23%) to 14% (95% CI: 0% – 29%). AUC did not change by adding ONSD measurements. The pAUC for predicting a poor neurological outcome with a specificity > 99% increased from 0.0019 (95% CI: 0.0005 – 0.0037) to 0.0027 (95% CI: 0.0012 – 0.0046), proportional to the sensitivity for poor neurological outcome. The pAUC curve for predicting a good neurological outcome with a specificity > 90% increased from 0.0061 (95% CI: 0 – 0.027) to 0.011 (95% CI: 0 – 0.029). As expected, sensitivities and (p)AUCs were overall higher when excluding patients who died because of non-neurological causes.

Table B4: Performance measures of the logistic regression model, predicting outcome based on EEG pattern and SSEP results, and the mixed-effects logistic regression model, predicting outcome based on EEG pattern, SSEP results, and ONSD on days 1-3. When adding ONSD measurements to outcome predictions, AUC and sensitivity for both a good and poor outcome increased.

	EEG + SSEP	EEG + SSEP + ONSD
AUC	0.59 (0.46 – 0.72)	0.59 (0.43 – 0.76)
Poor outcome		
pAUC	0.0019 (0.0005 – 0.0037)	0.0027 (0.0012 – 0.0046)
Sensitivity (%)	19 (0 – 38)	27 (11 – 46)
Specificity (%)	100 (100 – 100)	100 (100 – 100)
Good outcome		
pAUC	0.0061 (0 – 0.027)	0.011 (0 – 0.029)
Sensitivity (%)	7.6 (0 – 23)	14 (0 – 29)
Specificity (%)	94 (81 – 100)	92 (81 – 100)

Performance measures are presented with 95% confidence intervals. EEG: electroencephalography, SSEP: somatosensory evoked potentials, ONSD: optic nerve sheath diameter, AUC: area under curve.
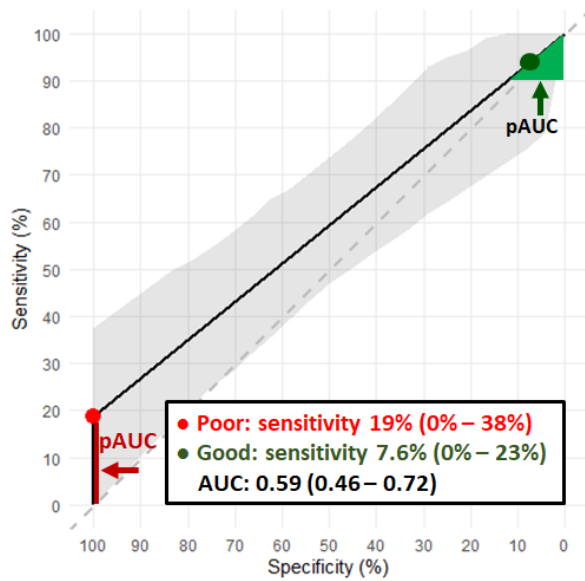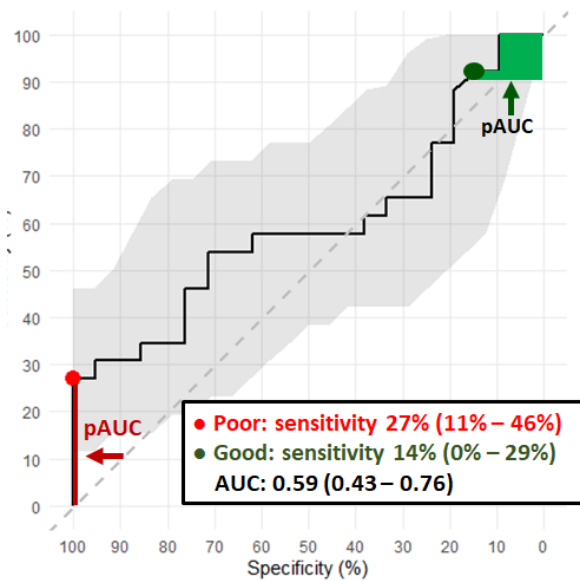
Figure B2: ROC curve of the logistic regression model predicting neurological outcomes by the electroencephalography (EEG) pattern and somatosensory evoked potentials (SSEP) results. The red point marks the threshold for predicting a poor neurological outcome. The green point marks the threshold for predicting a good neurological outcome. The red area is the pAUC for predicting poor outcome with a specificity > 99%, and the green area the pAUC for predicting a good neurological outcome with a specificity > 90%.

Figure B3: ROC curve of the mixed-effects logistic regression model predicting neurological outcome by electroencephalography (EEG) pattern, somatosensory evoked potentials (SSEP) results and optic nerve sheath diameter (ONSD) measurements. The red point marks the threshold for predicting a poor neurological outcome. The green point marks the threshold for predicting a good neurological outcome. The red area is the pAUC for predicting poor outcome with a specificity > 99%, and the green area the pAUC for predicting a good neurological outcome with a specificity > 90%.

Appendix C: Excluded segmentation algorithms

Besides the active contour models, two other segmentation algorithms were developed and evaluated, being the region growing and edge detection methods.

C1. Region-Growing

A region-growing algorithm was downloaded and adapted from MATLAB Central File Exchange [56]. The region-growing algorithm performed segmentation from a single seed point, which was manually placed inside the optic nerve. The region was iteratively grown by comparing the intensity values of the unallocated neighbouring pixels to the mean intensity of the region and adding the pixel with the smallest difference. The iterative process stopped when the intensity difference between the mean intensity of the region and a new neighbouring pixel was larger than a set threshold. The region-growing algorithm was evaluated on four transorbital sonograms. The stopping criterion was varied to study its effect on segmentation.

An example of the segmentation results obtained using the region-growing algorithm is shown in **Figure C1**. The optic nerve and its sheath were separated from the surrounding tissues, but the region had grown outwards towards the eye globe and other surrounding black structures. When the stopping criterion was lowered, the optic nerve and its sheath are not segmented properly. Similar results were obtained for the other three sonograms. Further analysis of the segmented region to ultimately measure the ONSD would thus be difficult and was therefore not initiated.
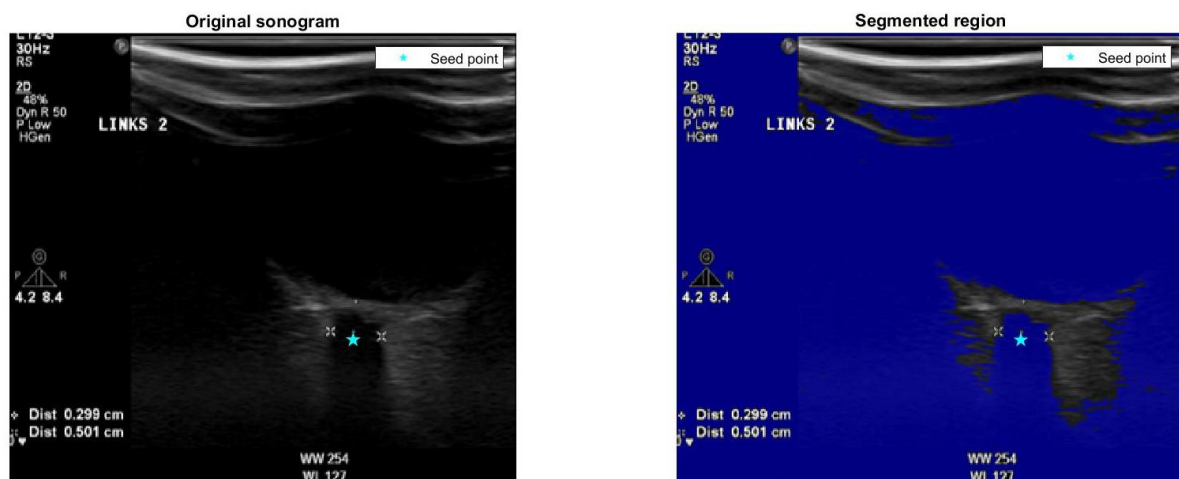

Figure C1: Results of the region growing algorithm performed on a transorbital sonogram. The region has grown outwards towards the eye globe and other surrounding black structures.

C2. Edge Detection

Edge detection using Canny's method was applied to four sonograms. Canny's method finds edges by searching for the local maxima of the image gradient and applies two thresholds to detect strong and weak edges [57]. The result of the application of Canny's method to a transorbital sonogram is shown in **Figure C2**. Canny's approach was not sensitive enough to detect the boundaries between the optic nerve sheath and the retrobulbar fat, as the contrast and therefore the image gradient between these tissues was too low. The method was able to detect the boundaries in the other three sonograms. However, the edge detection method was discarded due to its low sensitivity.
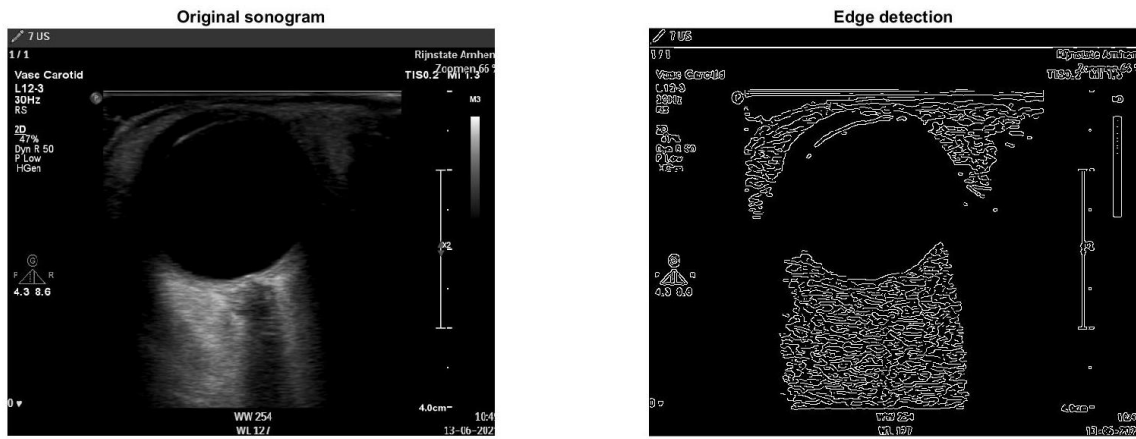
Figure C2: Results of edge detection using Canny's approach performed on a transorbital sonogram. Canny's approach was not sensitive enough to detect the boundaries between the optic nerve sheath and the retrobulbar fat.

Appendix D: Active Contour Models

Active contour models, also called 'snakes', are energy-minimising curves that deform to align with image features [41]. Each element **x** along the contour depend on two parameters:

$$x(s,t) = \begin{cases} s = \text{space parameter} \\ t = \text{time parameter} \end{cases}$$

The contour deforms over a series of iterations and its energy is affected by internal, image, and external constraint forces:

$$E_{\text{snake}} = \int_0^1 E_{\text{internal}}(x)ds + \int_0^1 E_{\text{external}}(x)\, ds + \kappa \int_0^1 E_{\text{image}}(x)\, ds,$$

where E_{snake}, E_{internal}, E_{external}, and E_{image} are the snake energy, the internal and external energy and the image energy respectively. κ affects the weight given to the image force.

The internal forces control the shape of the model. The internal energy of the snake is defined as:

$$E_{\text{internal}}(x) = \alpha \left| x'^{(s)} \right|^2 + \beta \left| x''(s) \right|^2,$$

where dashes indicate derivates.

The internal energy contains a first-order term controlled by $\alpha(s)$ and a second-order term controlled by $\beta(s)$. The first-order term gives the snake tension, making it contract. The second-order term creates stiffness, producing a snake smooth in outline.

The model is driven towards features of interest by the image forces. These image forces are generated by processing the image, for example by convolving the image with a Gaussian filter. The image energy is a weighted combination of three energy functionals:

$$E_{\text{image}} = w_{\text{lines}}E_{\text{line}} + w_{\text{edges}}E_{\text{edge}} + w_{\text{term}}E_{\text{term}}.$$

The energy functional for attracting towards lines can be set as:

$$E_{line} = I(x),$$

where $I(x)$ is the image intensity. Depending on the sign of w_{lines}, the snake will attract to light or dark lines. Edges can be found by the following energy functional:

$$E_{edge} = -|\nabla I(x)|^2,$$

Terminations of line segments and corners are found using:

$$E_{term} = \frac{C_{yy}C_x^2 - 2C_{xy}C_xC_y + C_{xx}C_y^2}{(C_x^2 + C_x^2)^{3/2}},$$

where $C(x) = G_\sigma(x) * I(x)$ with $G_\sigma(x)$ the Gaussian with standard deviation σ.

Snakes need to be initialised at a starting position. These external constraints come from high-level sources, such as human operators or automatic initialisation procedures.

The nearest local minimum of the snake energy is found using gradient descent, by moving along the negative gradient of the snakes energy with a controlled step size γ:

$$x \leftarrow x - \gamma \nabla E_{\text{snake}}(x),$$

Gradient vector flow is applied to encourage convergence to concave boundaries by extending the gradient vectors of the image forces throughout the total image [48].

Appendix E: Figures

Original image

Stretching with n = 2

Stretching with n = 3

Stretching with n = 4
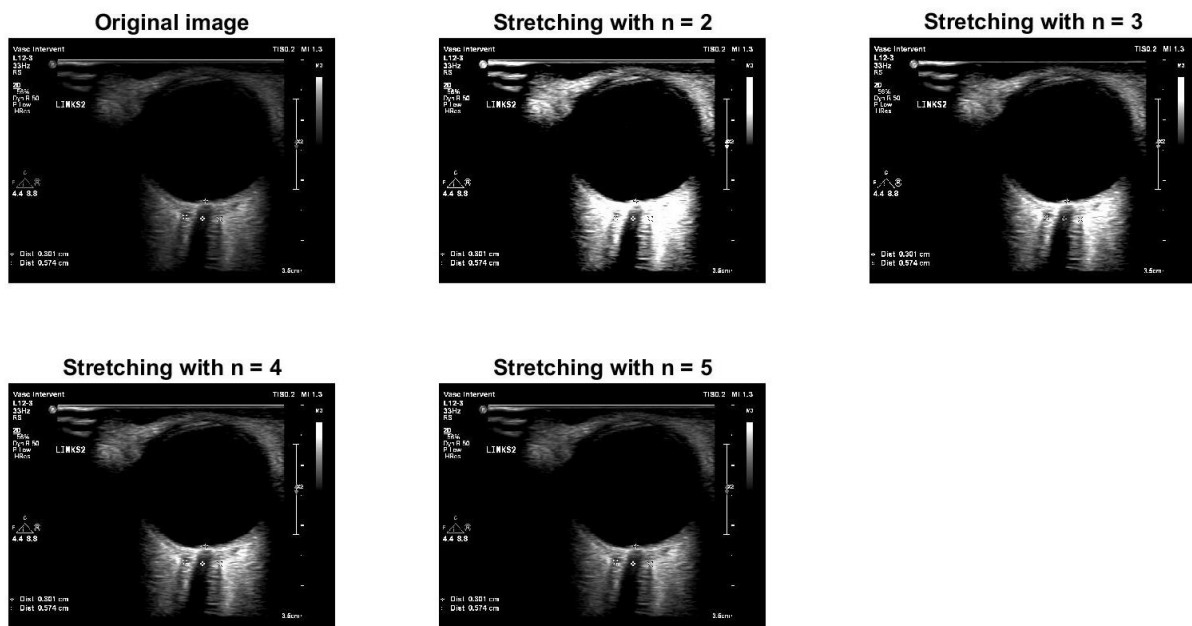

Stretching with n = 5

Figure E1: Enhanced transorbital sonograms by performing standard deviation based image stretching using different values of n when mapping intensity values between 0 and $\mathrm{mean}(\boldsymbol{I}) + n\,\mathrm{std}(\boldsymbol{I})$ to 0 and 1.
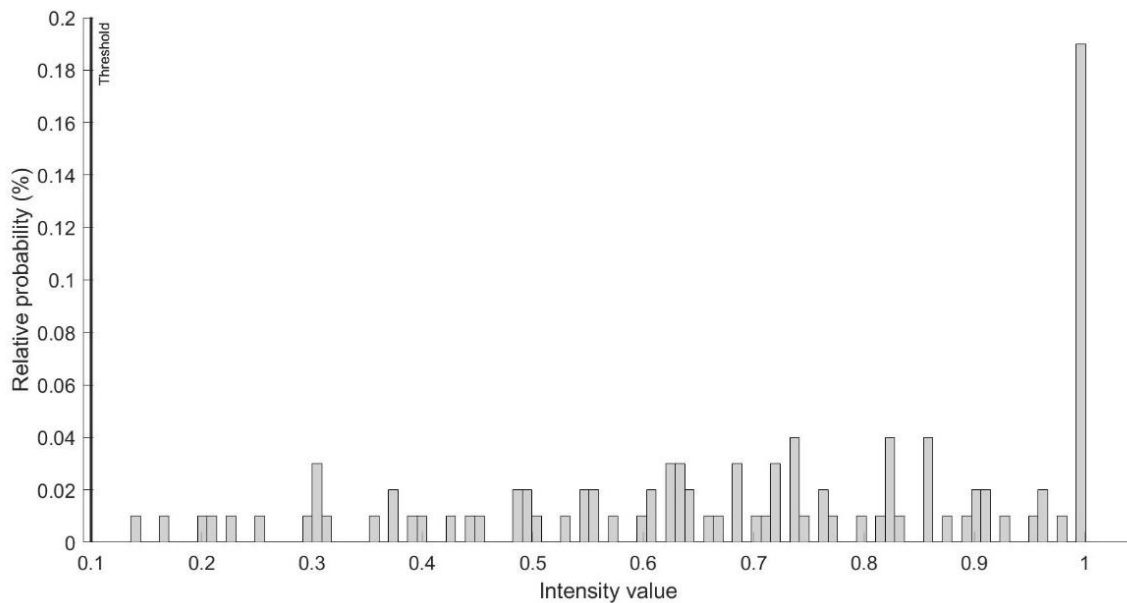
Figure E2: Intensity values of the retrobulbar fat in one hundred randomly selected transorbital sonograms, when a pixel representing the surrounding tissue was selected.
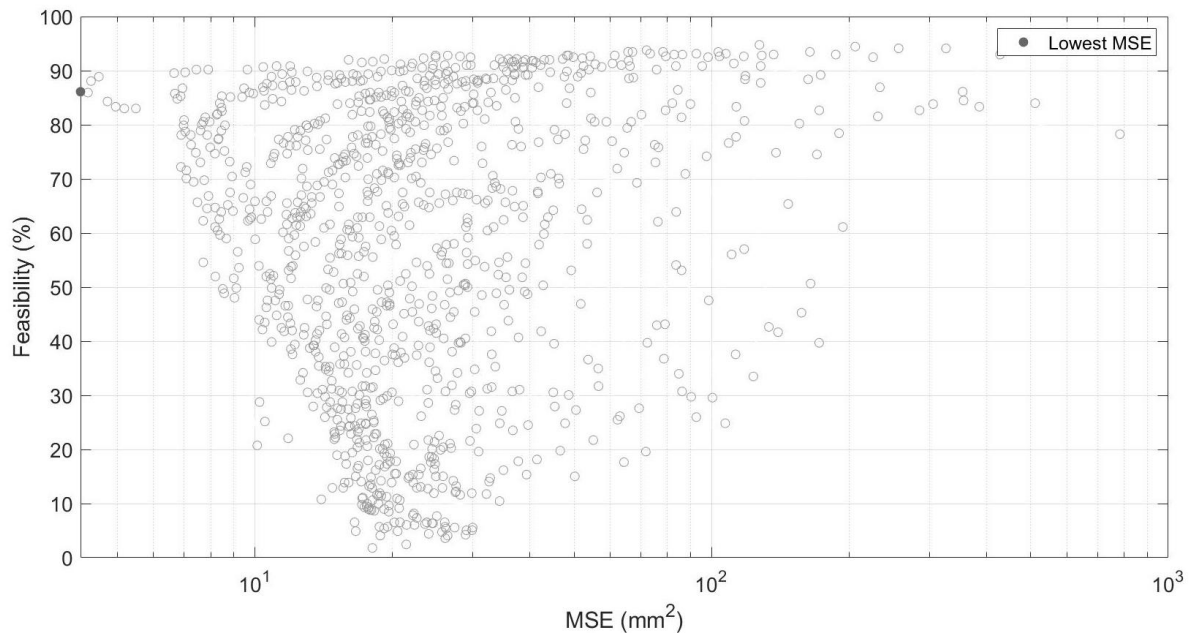
Figure E3: Mean square error (MSE) and feasibility of the 1024 hyperparameter combinations to automatically measure the optic nerve sheath diameter (ONSD) in transorbital sonograms. The x-axis is plotted on a base-10 logarithmic scale. The hyperparameter combination α = 0.08, γ = 1, δ = 0.05, w_{lines} = -0.02, and w_{edges} = 1 with the smallest error (4.15 mm^2) and feasibility of 90.2% is highlighted as the filled circle.
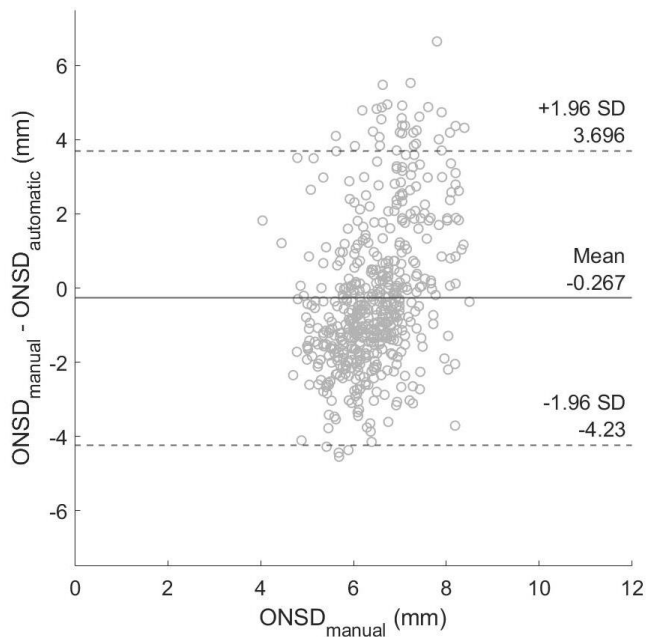

Figure E4: Bland-Altman plot for the manual and automatic ONSD measurements, with the manual ONSD measurements on the x-axis. It shows a bias of -0.267 mm and 95% limits agreements between -4.230 and 3.696 mm.

Appendix F: Contrast and Histograms of Sonograms

Sonograms in which the ONSD was overestimated had significantly higher contrast (0.47 [0.36 – 0.62]) than sonograms with an absolute error ≤ 0.5 mm, which had a contrast of 0.40 [0.28 – 0.61], (p = 0.006), and than sonograms in which the ONSD was underestimated (0.44 [0.28 – 0.59]) (p = 0.012) (**Figure F1**). Average histograms of sonograms with an absolute error ≤ 0.5 mm between manual and automatic ONSD measurements, sonograms in which the ONSD was under- or overestimated by >0.5 mm were dissimilar (p < 0.001) (**Figure F2**). Visual inspection of the violin plots and average histograms however showed similar distributions.
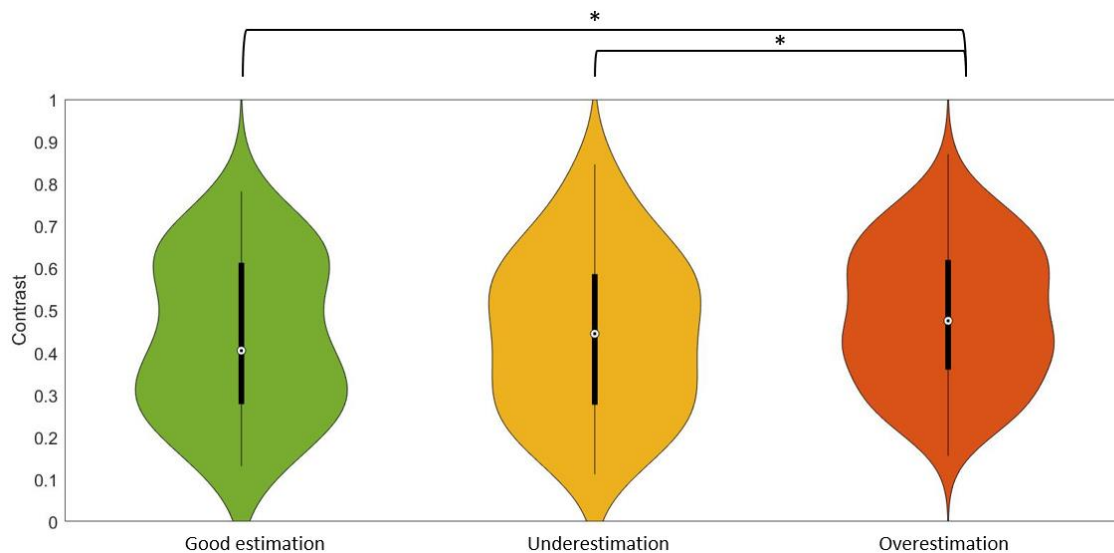
Figure F1: Violin plots of contrast in sonograms with an absolute error ≤ 0.5 mm between manual and automatic ONSD measurements (green), sonograms in which the ONSD was underestimated by > 0.5 mm (yellow), and overestimated by > 0.5 mm (orange). Sonograms in which the ONSD was overestimated had higher contrast (0.47 [0.36 – 0.62]) in comparison with sonograms with an absolute error ≤ 0.5 mm, which had a contrast of 0.40 [0.28 – 0.61], (p = 0.006), and in comparison with sonograms in which the ONSD was underestimated (0.44 [0.28 – 0.59]) (p = 0.012). Significant differences are indicated by *.
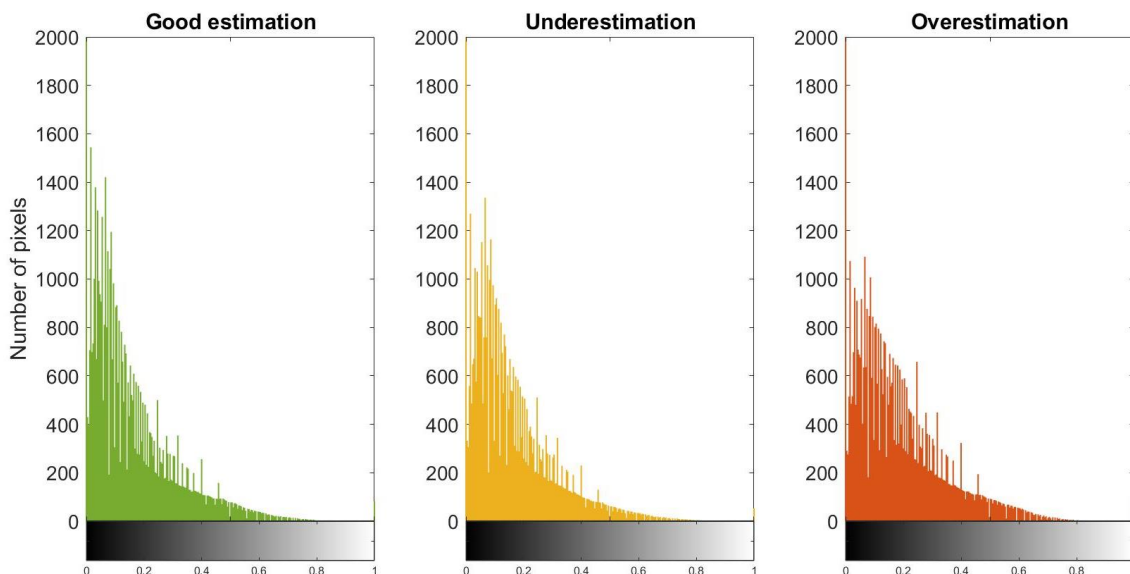

Figure F2: Average histograms of sonograms with an absolute error ≤ 0.5 mm between manual and automatic ONSD measurements (green), sonograms in which the ONSD was underestimated by > 0.5 mm (yellow), and overestimated by > 0.5 mm (orange).