Wu *et al. Critical Care* (2025) 29:531
https://doi.org/10.1186/s13054-025-05762-4

Critical Care

RESEARCH Open Access

Longitudinal assessment of automated gray-white matter ratio for outcome prediction after cardiac arrest

Min Wu[1†], Weiwei Liu[2†], Sheng Kuang[3], Jiajia Zhou[1], Xujian He[2], Jia Hu[2], Yongjian Deng[1], Huiying Lin[1], Jie Zhang[4], Chenyang Zhao[5], Meiqi Zeng[5], Hanxiao Wang[1], Meng Wu[6], Wangxiao Bao[7], Tong Li[2*†], Benyan Luo[1*†] and Kang Wang[1*†]

Abstract

Background The gray-white matter ratio (GWR) on head CT is a well-established marker of hypoxic-ischemic brain injury after cardiac arrest, but its prognostic performance may vary with the timing of imaging. We aimed (i) to evaluate the prognostic value of GWR across serial CT scans within the same comatose patients, and (ii) to determine whether the longitudinal changes of GWR provide additional prognostic information beyond single time-point measurements.

Methods We prospectively recruited 123 comatose patients with cardiac arrest admitted to three intensive care units. All patients underwent serial non-contrast head CT at three predefined time windows (< 24 h, 24–96 h, and 96–168 h after cardiac arrest). GWR values were automatically calculated using an atlas-based approach. Neurological outcome at 3 months was assessed with the Cerebral Performance Category score (CPC) and dichotomized into good (CPC 1–2) or poor (CPC 3–5). GWR values and their progression were compared between outcome groups. Prognostic accuracy of GWR at each time window was assessed using receiver operating characteristic (ROC) analysis.

Results GWR was consistently lower in patients with poor outcomes compared to those with good outcomes across all time windows (for all $p < 0.001$). In poor-outcome patients, GWR declined after the first 24 h, whereas it was stable in good-outcome patients. The prognostic performance of GWR improved with later imaging, with an AUC of 0.72 (95% CI 0.62–0.81) at < 24 h, 0.78 (95% CI 0.69–0.86) at 24–96 h, and 0.81 (95% CI 0.72–0.88) at 96–168 h after cardiac

†Min Wu and Weiwei Liu contributed equally as first authors.

†Kang Wang, Benyan Luo and Tong Li contributed equally as senior authors.

*Correspondence:
Tong Li
drli@zju.edu.cn
Benyan Luo
luobenyan@zju.edu.cn
Kang Wang
fcwangk1@zju.edu.cn

Full list of author information is available at the end of the article

arrest. Incorporating longitudinal changes in GWR slightly improved prediction, with the AUC increasing from 0.81 to 0.83 at 96–168 h.

Conclusions Automated GWR is a useful predictor of outcome after cardiac arrest, with higher accuracy on delayed CT (> 24 h). The different GWR progression trajectories between patients with poor and good outcomes suggest that longitudinal CT assessments may provide additional prognostic information.

Keywords Coma, Prognosis, Computed tomography, Longitudinal imaging, Gray-white matter ratio

Introduction

Accurate prognostication in comatose patients after cardiac arrest remains a major clinical and ethical challenge in the intensive care units (ICU) [1–4]. Bedside neurological examination is commonly used but is often confounded by sedation and mechanical intubation. Current guidelines from European Resuscitation Council (ERC) and the European Society for Intensive Care Medicine (ESICM) recommend brain imaging as part of multimodal assessments for prognosis after cardiac arrest [5].

Head computed tomography (CT) is widely available and rapidly performed, making it well suited for the ICU setting. After cardiac arrest, hypoxic-ischemic brain injury is frequently visualized on non-contrast CT as diffuse cerebral edema with loss of gray-white matter differentiation [4, 6–8]. Several qualitative radiological signs, such as the pseudo-subarachnoid hemorrhage sign and the white cerebellum sign, have been associated with poor prognosis [9, 10]. To improve objectivity and reduce inter-rater variability, the gray-white matter ratio (GWR) was introduced as a quantitative marker of edema after hypoxic-ischemic brain injury [8, 11]. A lower GWR reflects a decrease in the normally hyperdense gray matter toward white matter levels. Compared to manual calculation of GWR, automated GWR calculation demonstrated superior accuracy and reproducibility [12, 13].

Several previous studies have shown that the prognostic accuracy of GWR depends strongly on the timing of CT [14–16]. Specifically, discrimination is poorer in the early hours after arrest and improves with CTs obtained later. However, most prior evidence focused on the first 24 h [12, 16–18], and is based on cross-sectional analyses, with each patient contributing only one CT scan at a given time point, making it difficult to disentangle timing effects from inter-patient variability. Although a few studies have explored GWR changes within the same patients using two consecutive CTs [14, 19, 20], the time-dependent effect of GWR requires further evidence from longitudinal data. In particular, it remains unclear whether the progression of GWR within individual patients provides prognostic information beyond static GWR and short-term outcomes such as in-hospital mortality. Our study addressed these gaps by evaluating serial CTs within individual patients and evaluating their prognostic value for 3-month neurological outcomes.

In this study, we prospectively enrolled comatose patients after cardiac arrest from three ICUs who underwent serial head CTs at three predefined intervals (< 24 h, 24–96 h, and 96–168 h after cardiac arrest). We aimed to (i) compare the prognostic value of GWR across these time windows within the same patients, and (ii) test whether longitudinal, intra-individual progression of GWR is informative for outcomes.

Methods

Participants

This prospective observational study was conducted in three ICUs (the Zhijiang and Yuhang campuses of the First Affiliated Hospital, Zhejiang University School of Medicine, and Quzhou People's Hospital) between March 2021 and June 2025. Comatose patients admitted to the intensive care unit (ICU) after non-traumatic cardiac arrest were screened. Inclusion criteria were as follows: age ≥ 18 years; in a coma defined as a Glasgow Coma Scale (GCS) score < 8 after return of spontaneous circulation (ROSC); and availability of non-contrast head CT scans at all three predefined time windows (< 24 h, 24–96 h, and 96–168 h after cardiac arrest). Patients were excluded if the cardiac arrest was due to a primary neurological etiology or if CT image quality was inadequate or showed a large structural brain lesion. All patients received targeted normothermia. The decision on withdrawal of life-sustaining therapy (WLST) was made collaboratively by the multidisciplinary team discussion and patient's representatives, while the team remained blinded to the GWR results of the study. The study complied with the Declaration of Helsinki, was approved by the local Research Ethics Committee, and written informed consent was obtained from legally authorized representatives.

CT acquisition and image analysis

Non-contrast head CTs from all types of scanners and software platforms were permitted (Fig. S2). Technical prerequisites included the availability of axial slices with 5 mm thickness obtained with a tube voltage of 120 kV. CT scans with severe artefacts or structural brain lesions were excluded by two experienced neurologists. Patients with small focal infarctions outside the basal ganglia were retained, provided that the lesions did not distort

or overlap with the regions of interest used for GWR measurement.

Image analysis followed previously published approaches [13, 21]. Each CT was co-registered to a freely available CT template [22] using the FMRIB Linear Image Registration Tool (FLIRT; settings: 256 histogram bins, correlation ratio cost function, 12 degrees of freedom, trilinear interpolation) and subsequently the FMRIB Nonlinear Image Registration Tool (FNIRT; settings: membrane energy regularization). Registration quality was manually verified with *slicesdir* in FSL by a senior neurologist (JJZ) with over 10 years of experience in neuroimaging, overlaying the template on the registered images. Inverse transformation fields were then calculated using *invwarp* to allow atlas-based regions to be projected into individual CT spaces.

A putamen probability map thresholded at 60% was derived from the Harvard-Oxford subcortical structural atlas, and a posterior limb of the internal capsule (PLIC) mask was obtained from the ICBM-DTI-81 white matter atlas. Both atlases are distributed with FSL and matched the template anatomy. Using the inverse transformations, atlas maps were warped into each patient's CT space, and mean Hounsfield units (HU) were bilaterally extracted from the putamen and PLIC with *fslstats*. The automated gray-white matter ratio was then computed for each CT scan as:

$$GWR = \frac{Putamen}{PLIC}$$

Outcome assessment

Neurological outcome was assessed at 3 months after cardiac arrest using the Cerebral Performance Category (CPC) scale (1 = good cerebral performance; 2 = moderate disability; 3 = severe disability; 4 = coma/vegetative state; 5 = brain death) [23]. Outcomes were obtained by structured interview either in person or by telephone if the patients had been discharged from hospital. The interviewer conducting the outcome assessments was blinded to the results of prior examinations, including routine clinical evaluations, and was unaware of the CT results. Outcomes were dichotomized as good (CPC 1–2) and poor (CPC 3–5) [4, 24].

Statistics

Continuous variables were assessed for normality using the Shapiro-Wilk test. Normally distributed variables were reported as mean ± standard deviation (SD), and non-normally distributed variables were reported as median with interquartile range (IQR). Categorical variables were summarized as counts and percentages.

Between-group comparisons of GWR (good vs. poor outcome) were performed with the Mann-Whitney U test. Within-subject changes of GWR across predefined time windows were analyzed with linear mixed-effects models (LMMs) using the *lme4* package by including time as a fixed effect and subject as a random intercept. Post hoc multiple comparisons were adjusted using the false discovery rate (FDR) [25].

To evaluate prognostic performance, receiver operating characteristic (ROC) curves were constructed for GWR at each time window. Areas under the ROC curves (AUCs) were estimated with 95% confidence intervals (CIs) using 10,000 bootstrap resamples. A fixed GWR threshold of 1.10 was used following previous literature [13, 21]. Sensitivity and specificity were calculated, and 95% CIs were estimated using Wilson's method.

To test whether progression of GWR improved prognostic accuracy beyond single time points, we further constructed nested logistic regression models including GWR and longitudinal changes in GWR (changes from < 24 h to later windows). AUC comparisons were performed by DeLong test. A p-value < 0.05 was considered statistically significant. Analyses were performed in MATLAB (R2020a, MathWorks), Python 3.10, and R (version 4.3.2).

Results

Patient enrollment and demographics

357 comatose patients with non-traumatic cardiac arrest were assessed for eligibility (Fig. 1). One hundred and sixteen patients were excluded because they had no head CT, had primary neurologic etiology, or had inadequate CT quality. Among the remaining 241 patients with at least one CT within 168 h (7 days), 118 patients did not complete the full three-scan sequence within the predefined windows. The final primary analysis cohort consisted of 123 patients with serial CTs at < 24 h, 24–96 h, and 96–168 h post-arrest (Fig. 1). The median age was 57 years (IQR 44–67), and 85 (69.1%) were male. The time from arrest to imaging was 12.4 h (IQR 6.7–15.5) for the first window (< 24 h), 62.3 h (IQR 47.1–79.5) for the second window (24–96 h), and 134.3 h (IQR 120.5–154.1) for the third window (96–168 h). At 3 months, 91 patients (74.0%) had poor functional outcome (CPC 3–5) and 34 of 59 deceased patients (57.6%) died following WLST. Baseline characteristics and outcomes for both included and excluded patients are summarized in Table 1.

GWR comparison between poor and good outcomes

The gray-white matter ratio was consistently lower in patients with poor outcomes compared to those with good outcomes across all three time windows. At < 24 h, the median GWR was 1.13 (IQR 1.10–1.17) in patients with poor outcome and 1.18 (IQR 1.15–1.20) in patients with good outcome (Mann-Whitney U test,

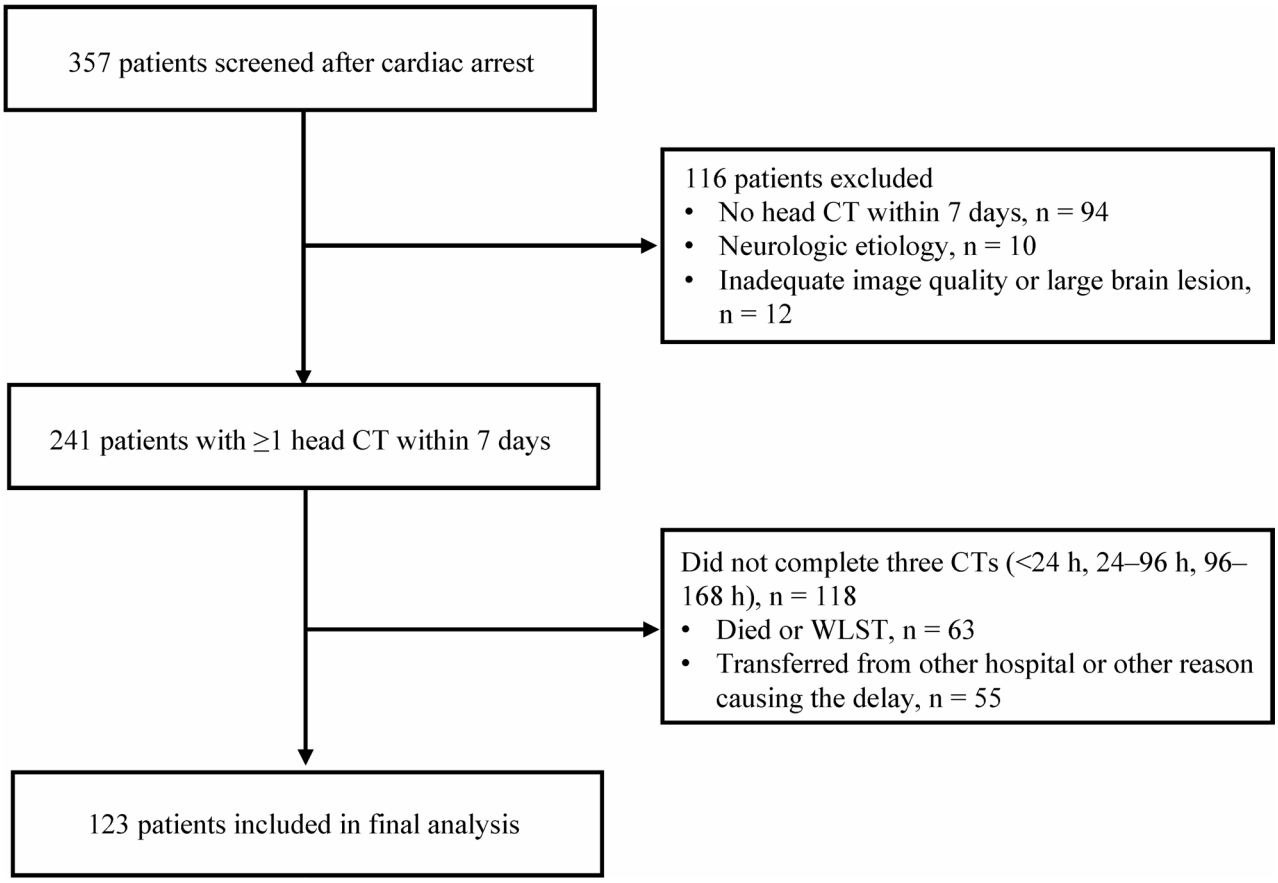

Fig. 1 Flowchart of patient selection and exclusion

$p = 1.94 \times 10^{-4}$, FDR-corrected). At 24–96 h, GWR was 1.11 (IQR 1.06–1.16) and 1.18 (IQR 1.15–1.20) for poor and good outcomes, respectively ($p = 4.11 \times 10^{-6}$, FDR-corrected). At 96–168 h, GWR remained lower in poor-outcome patients (median 1.13, IQR 1.07–1.17) compared with good-outcome patients (median 1.20, IQR 1.17–1.22; $p = 8.27 \times 10^{-7}$, FDR-corrected, Fig. 2). The difference in median GWR between the two groups increased after 24 h (0.05 at < 24 h, 0.07 at 24–96 h and 96–168 h).

Reduction in GWR over time

In the poor-outcome group, a linear mixed-effects model showed a significant main effect of time window ($F(2,180) = 6.90$, $p = 0.001$). Post-hoc comparisons indicated a decline from < 24 h to 24–96 h ($t(180) = 3.69$, $p = 9.00 \times 10^{-4}$, FDR corrected) and from < 24 h to 96–168 h ($t(180) = 2.18$, $p = 0.04$, FDR corrected). No significant difference was observed between 24 and 96 h and 96–168 h ($t(180) = -1.51$, $p = 0.13$, FDR corrected).

In contrast, in the good-outcome group, the mixed-effects model did not show a significant main effect of time windows ($F(2,62) = 2.56$, $p = 0.09$). Pairwise comparisons were all non-significant (< 24 h vs. 24–96 h: $t(62) = 0.41$, $p = 0.68$; <24 h vs. 96–168 h: $t(62) = -1.72$, $p = 0.14$; 24–96 h vs. 96–168 h: $t(62) = -2.13$, $p = 0.11$; all FDR-corrected). These results indicated relative stability of GWR over time in patients with good outcomes. An example of the GWR change is shown in Fig. 3.

Outcome prediction using GWR

Receiver-operating-characteristic analysis showed that GWR discrimination increased over time. The AUC was 0.72 (95% CI 0.62–0.81) at < 24 h, 0.78 (95% CI 0.69–0.86) at 24–96 h, and 0.81 (95% CI 0.72–0.88) at 96–168 h (Table 2, Fig. S1). DeLong tests confirmed that these increases were statistically significant (< 24 h vs. 24–96 h, $p = 1.13 \times 10^{-5}$; <24 h vs. 96–168 h, $p = 3.21 \times 10^{-8}$; 24–96 h vs. 96–168 h, $p = 1.49 \times 10^{-5}$). Using a fixed threshold of 1.10 (i.e., poor outcome: GWR < 1.10), specificity remained high while sensitivity was low: at < 24 h, sensitivity was 0.25 (95% CI 0.17–0.35) and specificity was 0.94 (95% CI 0.80–0.98; poor outcome was incorrectly predicted in two patients, Fig. S3); at 24–96 h, sensitivity was 0.44 (95% CI 0.34–0.55) and specificity was 0.91 (95% CI 0.76–0.97); and at 96–168 h, sensitivity was 0.37 (95% CI 0.28–0.48) and specificity was 1.00 (95% CI 0.90–1.00). Since the initial < 24 h window was broad

Table 1 Characteristics of included and excluded patients

Variable	Included			Excluded
	All (*n* = 123)	Good outcome (*n* = 32)	Poor outcome (*n* = 91)	All (*n* = 234)
Demographics				
Age, years (median [IQR])	57 [44–67]	47 [37–63]	59 [49–69]	58 [53–69]
Sex, male — no. (%)	85 (69.1%)	22 (68.8%)	63 (69.2%)	168 (71.8%)
Medical history, no. (%)				
Hypertension	48 (39.0%)	11 (34.4%)	37 (40.7%)	67 (28.6%)
Diabetes	35 (28.5%)	7 (21.9%)	28 (30.8%)	69 (29.5%)
Heart failure	25 (20.3%)	4 (12.5%)	21 (23.1%)	59 (25.2%)
Prior stroke	7 (5.7%)	0 (0%)	7 (7.7%)	21 (9.0%)
Renal failure	6 (4.9%)	2 (6.3%)	4 (4.4%)	22 (9.4%)
COPD	6 (4.9%)	1 (3.1%)	5 (5.5%)	4 (1.7%)
Cause of cardiac arrest, no. (%)				
Acute coronary syndrome	49 (39.8%)	15 (46.9%)	34 (37.4%)	84 (35.9%)
Primary arrhythmia	20 (16.3%)	3 (9.4%)	17 (18.7%)	41 (17.5%)
Cardiomyopathy	9 (7.3%)	4 (12.5%)	5 (5.5%)	19 (8.1%)
Myocarditis	3 (2.4%)	2 (6.3%)	1 (1.1%)	13 (5.6%)
Cardiogenic shock/Heart failure	0 (0%)	0 (0%)	0 (0%)	3 (1.3%)
Valvular disease	3 (2.4%)	1 (3.1%)	2 (2.2%)	11 (4.7%)
Asphyxia/Respiratory failure	20 (16.3%)	2 (6.3%)	18 (19.8%)	30 (12.8%)
Toxic/Metabolic cause	2 (1.6%)	1 (3.1%)	1 (1.1%)	5 (2.1%)
Other/Unknown cause	17 (13.8%)	4 (12.5%)	13 (14.3%)	28 (12.0%)
EEG patterns	(*n* = 107)			(*n* = 109)
Benign, no. (%)	42 (39.2%)	19 (73.1%)	23 (28.4%)	36 (33.0%)
Malignant, no. (%)	34 (31.8%)	5 (19.2%)	29 (35.8%)	43 (39.5%)
Highly malignant, no. (%)	31 (29.0%)	2 (7.7%)	29 (35.8%)	30 (27.5%)
Resuscitation/timing				
Time to ROSC, min (median [IQR])	30.0 [13.0–60.0]	17.0 [6.0–33.0]	29.0 [16.5–61.5]	32.5 [15.0–64.0]
Timing of 1 st CT, h (median [IQR] (range))	12.4 [6.7–15.5] (1–23.7)	12.8 [8.9–15.2] (3–23.7)	11.8 [5.8–15.8] (1–23.5)	–
Timing of 2nd CT, h	62.3 [47.1–79.5] (25–92)	60.2 [38.5–80.7] (25–87.8)	62.6 [53.8–78.8] (29.0–92)	–
Timing of 3rd CT, h	134.3 [120.5–154.1] (99.5–165)	127.2 [108.9–134.5] (101.2–156.1)	135.3 [121.8–158.1] (99.5–165)	–
Outcome at 3 months (CPC)				
Good (CPC 1–2), no. (%)	32 (26.0%)	32 (100%)	0 (0%)	43 (18.4%)
Poor (CPC 3–5), no. (%)	91 (74.0%)	0 (0%)	91 (100%)	191 (81.6%)
CPC = 1	16 (13.0%)	16 (50%)	0 (0%)	20 (8.6%)
CPC = 2	16 (13.0%)	16 (50%)	0 (0%)	23 (9.8%)
CPC = 3	14 (11.4%)	0 (0%)	14 (15.4%)	18 (7.7%)
CPC = 4	18 (14.6%)	0 (0%)	18 (19.8%)	22 (9.4%)
CPC = 5	59 (48.0%)	0 (0%)	59 (64.8%)	151 (64.5%)
Withdrawal of life-sustaining therapy, no. (%)	34 (27.6%)	0 (0%)	34 (37.4%)	77 (32.9%)

COPD: chronic obstructive pulmonary disease. ROSC: return of spontaneous circulation; IQR: interquartile range; CPC: Cerebral Performance Category

and less comparable to prior studies on ultra-early imaging [13, 16], we further evaluated the prognostic performance of GWR in the ultra-early subgroup (< 6 h after cardiac arrest). In this subgroup, the AUC was 0.61 (95% CI 0.35–0.83), with sensitivity of 0.13 (95% CI 0.05–0.32), and specificity of 1.00 (95% CI 0.57–1.00). However, given the small number of patients (n = 28), these results should be interpreted with caution.

To test whether GWR progression (changes from the < 24 h baseline) improved prediction beyond static GWR, we fitted nested logistic models with and without changes of GWR from < 24 h to later windows. Adding the change from < 24 h yielded a slight increase of AUC from 0.81 (95% CI 0.72–0.88) to 0.82 (95% CI 0.74–0.89) at 96–168 h (p = 0.073, DeLong test). When both changes (< 24 to 24–96 h and < 24 to 96–168 h) were included, the

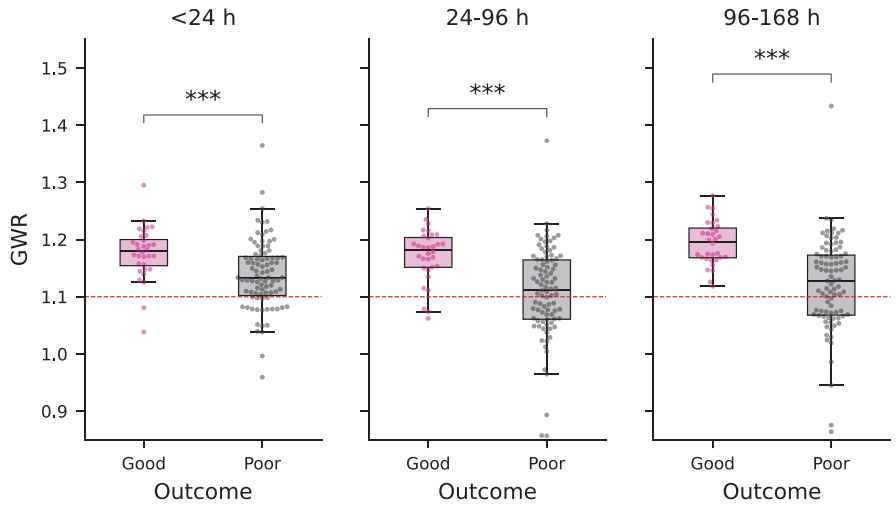

Fig. 2 Gray-white matter ratio (GWR) across time windows in patients after cardiac arrest. Patients with poor outcome (gray) had significantly lower GWR compared to those with good outcome (magenta). Dots show individual patients across all time windows. Box plots show the median and interquartile range (IQR). The red dashed line indicates the threshold of 1.1. *** $p < 0.001$

AUC increased to 0.83 (95% CI 0.75–0.90) and the difference reached statistical significance ($p = 0.032$, DeLong test).

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value; TP: true positive; TN: true negative; FP: false positive; FN: false negative.

Discussion

Our study confirms that the gray-white matter ratio is a valuable prognostic marker after cardiac arrest and that its performance is time-dependent. Later CT scans provided higher predictive accuracy than those obtained within the first 24 h. We further observed distinct GWR trajectories with values remaining relatively stable in patients with good outcomes but declining after the first 24 h in patients with poor outcomes. Incorporating longitudinal GWR change into prediction models slightly improved prognostic performance.

We found that a simplified automated GWR is informative for outcome prediction after cardiac arrest. Previous studies have shown that a simplified GWR is often sufficient because the putamen can be easily identified on CT and is particularly vulnerable to hypoxic-ischemic injury, while the internal capsule remains unaffected in the acute and early subacute periods [26–28]. This four-regions of interest (ROI) method (putamen and internal capsule bilaterally) has been reported to achieve similar or even higher prognostic accuracy compared with more complex 16-ROI approaches [27, 29]. More recently, Lang et al. demonstrated that attenuation measurements for GWR at the level of the basal ganglia–using both a simplified manual four-ROI method and an eight-ROI method–could predict poor outcomes with very high specificity [13]. Furthermore, they presented an objective automated method including four ROIs at basal ganglia with almost the same predictive performance.

The AUCs observed in our study were lower than those reported by Lang et al. despite using similar atlas-based automated GWR methods [13]. This discrepancy may be attributed to several methodological differences between the two studies. First, our study used the CPC for outcome assessment, whereas Lang et al. used the modified Rankin Scale (mRS). Although both are validated measures of functional outcome, they are not directly equivalent [30]. Second, our inclusion criteria were stricter as we included only comatose patients defined by a GCS < 8, while Lang et al. included unconscious patients defined by lack of command following.

Our results confirm the time dependence of GWR performance and are consistent with prior studies [13, 29]. Within the first 24 h, GWR showed limited discriminative power with low sensitivity, whereas specificity remained high. At 96–168 h, sensitivity had improved and specificity was near-perfect. This highlights that the clinical value of GWR lies primarily in its very high specificity for predicting poor neurological outcomes. Consistently, earlier studies reported AUCs around 0.70 in the < 6 h and 6–24 h windows and showed that a threshold of GWR < 1.10 identified only ~ 10–17% of poor outcomes while maintaining high specificity [14, 15]. A large prospective study similarly found increasing sensitivity with later imaging (up to ~ 48% at 48–96 h) while specificity remained ~ 99–100% [13]. Taken together, these findings emphasize two key points: (1) a low GWR can serve as a valuable marker in WLST decisions when false positives must be avoided, and (2) later imaging provides greater

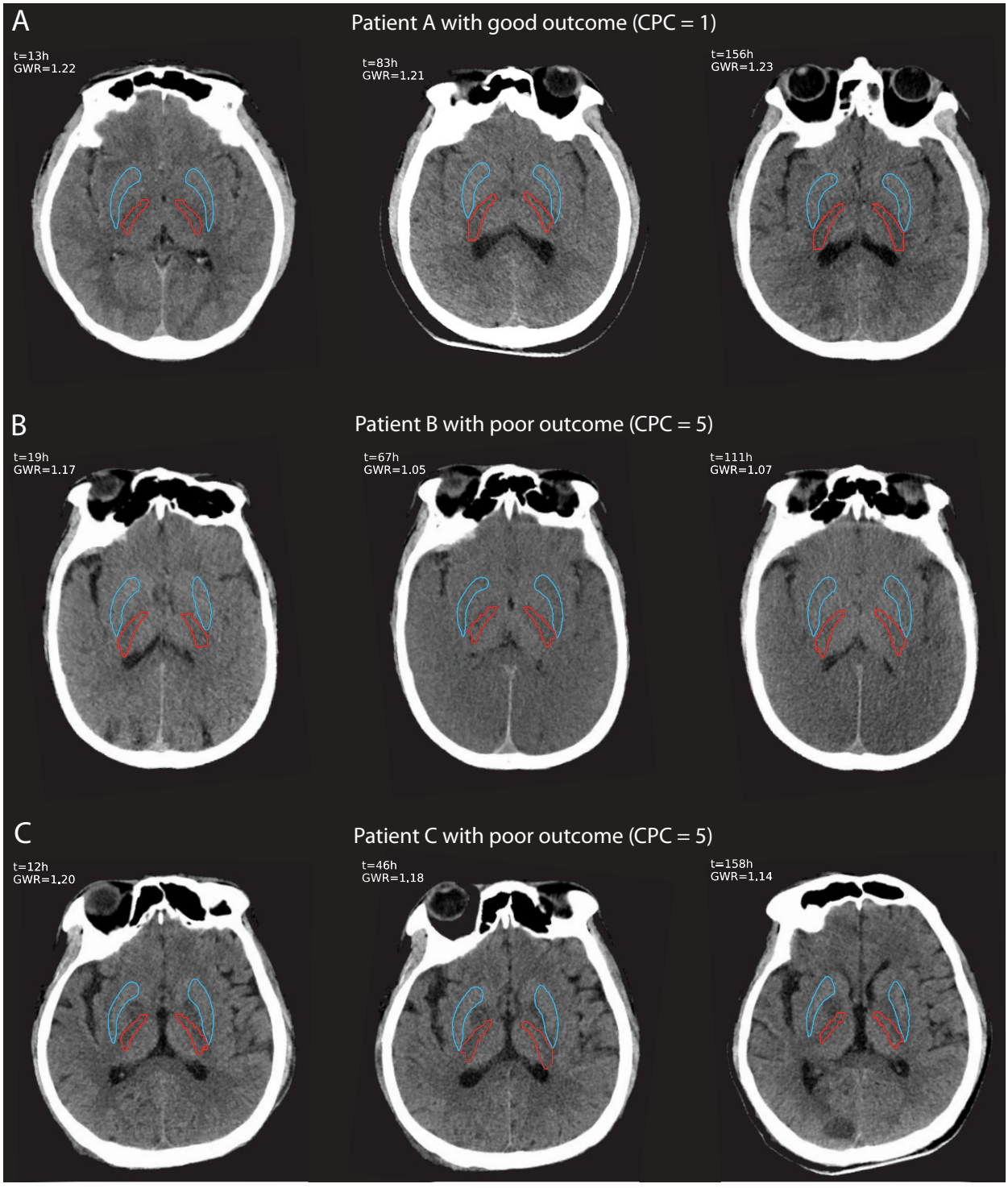

Fig. 3 Example of longitudinal CT scans showing gray-white matter ratio (GWR) progression in comatose patients after cardiac arrest. (A) Patient with good outcome (CPC = 1) showing stable GWR and preserved gray-white matter discrimination across time. (B) Patient with poor outcome (CPC = 5) with decreasing GWR and progressive loss of gray-white matter discrimination. (C) Patient with poor outcome (CPC = 5) with preserved gray-white matter discrimination but a progressive decline of GWR. Regions of interest used for GWR calculation are outlined (blue = putamen, red = posterior limb of internal capsule).

Table 2 Prognostic performance of gray–white matter ratio (GWR < 1.10) at different post–cardiac arrest time windows

Timing after arrest	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	TP	TN	FP	FN
< 24 h	0.72 (0.62–0.81)	0.25 (0.17–0.35)	0.94 (0.80–0.98)	0.92 (0.75–0.98)	0.31 (0.22–0.40)	23	30	2	68
24–96 h	0.78 (0.69–0.86)	0.44 (0.34–0.55)	0.91 (0.76–0.97)	0.93 (0.81–0.98)	0.36 (0.27–0.47)	40	29	3	51
96–168 h	0.81 (0.72–0.88)	0.37 (0.28–0.48)	1.00 (0.90–1.00)	1.00 (0.90–1.00)	0.36 (0.27–0.46)	34	32	0	57

prognostic reliability, supporting current guideline recommendations to rely on later imaging for prognostication [5].

Following previous studies, we set a threshold at GWR < 1.10, which achieved high specificity for poor neurological outcome. However, alternative cutoffs have been explored using higher thresholds [15, 29, 31, 32]. For example, Tsai refined GWR using the Youden index (1.204) and showed higher sensitivity at the cost of specificity [29].

We also observed distinct GWR trajectories between outcome groups. In patients with good outcomes, GWR remained relatively stable across time windows, consistent with limited progression of brain edema. In contrast, patients with poor outcomes showed a decline in GWR after the first 24 h, indicating progressive loss of gray-white contrast, followed by a plateau at 96–168 h. This trajectory is consistent with the pattern reported by Lang et al. and resembles the course of acute ischemic stroke [13], where cerebral edema typically peaks 3–5 days after injury and independently predicts worse outcomes [33, 34].

Although direct prior evidence on longitudinal GWR in patients after cardiac arrest is limited, a related study reported that a >8% increase in net water uptake in deep nuclei from < 6 h to >24 h predicted poor outcome with 100% specificity and 43% sensitivity [19]. Taken together, these results support the concept that longitudinal, within-patient change carries prognostic information.

Furthermore, our nested logistic models showed that adding GWR changes from < 24 h improved outcome prediction at 96–168 h. These findings reinforce the idea that longitudinal, within-patient change provides prognostic information beyond a single time point. This may be particularly relevant in patients with near-threshold GWR values (i.e., close to but above 1.1), where a continued decline may indicate progressive injury and a higher risk of poor outcome. However, for most patients, a single CT scan obtained at the later window already provides 100% specificity which is essential for WLST decision. It is not routinely recommended to perform three CT scans for all patients after cardiac arrest; rather, repeat imaging should be considered selectively for patients with inconclusive or borderline imaging findings.

There are some limitations to our study. First, although we included patients from three ICUs, larger multicenter studies with external validation are needed to confirm these results. Second, our analysis focused on prediction of poor prognosis and achieved high specificity, but its sensitivity remained low. This indicates that CT alone may not be sufficient and should be combined with other approaches such as electrophysiology (EEG), blood biomarkers, and clinical examination to support decision-making. Third, we did not separate deaths due to WLST from those due to natural progression after cardiac arrest. However, there is no formal policy for WLST in China, and treatment withdrawal was typically guided by clinicians but the final decision was made together with families. Fourth, our study specifically included comatose survivors of cardiac arrest who completed serial CT scans. This selection likely excluded patients who died early, underwent WLST before completing the imaging sequence, or were transferred from other hospitals with delayed scanning. Consequently, the most severely injured patients may have been excluded, limiting the generalizability of our findings. However, the results remained consistent when the analysis was repeated in an extended dataset that included all patients with CT scans (Table S1). Fifth, the use of different CT scanners may have introduced systematic variability in GWR measurements (see Fig. S2) and should be considered as a potential confounder in further studies. Finally, although automated GWR measurement improves reproducibility and minimizes observer bias, relying solely on automated assessments may occasionally lead to errors. Accordingly, visual inspection may serve as a useful complementary quality-control step.

Conclusions

Automated GWR is a useful predictor of outcome after cardiac arrest, with higher accuracy on delayed CT (> 24 h). The different GWR progression trajectories between patients with poor and good outcomes suggest that longitudinal CT assessments may provide additional prognostic information. However, further validation is needed in future studies.

Abbreviations

AUCs Areas under the curves
CPC Cerebral performance category score

CT	Computed tomography
CIs	Confidence intervals
ERC	European resuscitation council
ESICM	European society for intensive care medicine
FDR	False discovery rate
ROI	Four-regions of interest
GCS	Glasgow coma scale
GWR	Gray-white matter ratio
HU	Hounsfield units
ICU	Intensive care units
IQR	Interquartile range
LMMs	Linear mixed-effects models
PLIC	Posterior limb of the internal capsule
ROC	Receiver operating characteristic
ROSC	Return of spontaneous circulation
SD	standard deviation
WLST	Withdrawal of life-sustaining therapy

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13054-025-05762-4.

Supplementary Material 1

Acknowledgements
We would like to thank all patients and their families for their participation.

Author contributions
MW, WWL, TL, BYL, and KW conceived and designed the study, interpreted the results, and drafted the manuscript. MW and SK performed the data analyses. MW, SK, and JJZ prepared the tables and figures. JJZ, XJH, JH, YJD, HYL, JZ, CYZ, MQZ, HXW, MW, WXB were responsible for patient enrollment and data collection. YJD and BYL were responsible for data storage. WXB and BYL obtained funding. All authors revised and approved the final version of the manuscript.

Funding
This work was supported by the China Brain Project (2022ZD0208905 [BYL]), the National Natural Science Foundation of China (U22A20293 [BYL], 82302852 [WXB]), and the Fundamental Research Funds for the Central Universities (2023ZFJH01-01, 2024ZFJH01-01 [BYL]).

Data availability
The imaging data used in this study cannot be publicly shared due to Chinese data protection rules. Requests for data access can be sent as a formal proposal specifying the recipient and the purpose of the data transfer to the appropriate data protection authority. The data analysis scripts will be shared via the sharing platform OSF ([https://osf.io/283yz/]) on publication.

Declarations

Ethics approval and consent to participate
This study was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (20211163, 20240265). Quzhou People's Hospital participated as a sub-study site under this approval; no additional local ethics number was required. Written informed consent was obtained from the patients' legal representatives.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Neurology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
[2]Intensive Care Unit, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
[3]The D-Lab, Department of Precision Medicine, GROW - Research Institute of Oncology and Reproduction, Maastricht University, 6229 ER, Maastricht, The Netherlands
[4]Center for Rehabilitation Medicine, Rehabilitation and Sports Medicine Research Institute of Zhejiang Province, Department of Rehabilitation Medicine, Zhejiang Provincial People's Hospital, Hangzhou Medical College, 310014, Hangzhou, China
[5]Department of Information Engineering, Quzhou People's Hospital, 324000, Quzhou, China
[6]Intensive Care Unit, Quzhou People's Hospital, 324000, Quzhou, China
[7]Department of Rehabilitation Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, 310003, Hangzhou, China

Received: 7 October 2025 / Accepted: 9 November 2025
Published online: 29 December 2025

References

1. Rossetti AO, Rabinstein AA, Oddo M. Neurological prognostication of outcome in patients in coma after cardiac arrest. Lancet Neurol. 2016;15:597–609. https://doi.org/10.1016/S1474-4422(16)00015-6.
2. Greer DM, Rosenthal ES, Wu O. Neuroprognostication of hypoxic–ischaemic coma in the therapeutic hypothermia era. Nat Rev Neurol Nat Publishing Group. 2014;10:190–203. https://doi.org/10.1038/nrneurol.2014.36.
3. Rohaut B, Calligaris C, Hermann B, Perez P, Faugeras F, Raimondo F, et al. Multimodal assessment improves neuroprognosis performance in clinically unresponsive critical-care patients with brain injury. Nat Med. 2024. https://doi.org/10.1038/s41591-024-03019-1.
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 Med. 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 Med. 2021;47:369–421. https://doi.org/10.1007/s00134-021-06368-4.
6. Chae MK, Ko E, Lee JH, Lee TR, Yoon H, Hwang SY, et al. Better prognostic value with combined optic nerve sheath diameter and grey-to-white matter ratio on initial brain computed tomography in post-cardiac arrest patients. Resuscitation. 2016;104:40–5. https://doi.org/10.1016/j.resuscitation.2016.04.001.
7. Lee BK, Jeung KW, Song KH, Jung YH, Choi WJ, Kim SH, et al. Prognostic values of Gray matter to white matter ratios on early brain computed tomography in adult comatose patients after out-of-hospital cardiac arrest of cardiac etiology. Resuscitation. 2015;96:46–52. https://doi.org/10.1016/j.resuscitation.2015.07.027.
8. Torbey MT, Selim M, Knorr J, Bigelow C, Recht L. Quantitative analysis of the loss of distinction between Gray and white matter in comatose patients after cardiac arrest. Stroke Am Heart Association. 2000;31:2163–7. https://doi.org/10.1161/01.STR.31.9.2163.
9. Yuzawa H, Higano S, Mugikura S, Umetsu A, Murata T, Nakagawa A, et al. Pseudo-subarachnoid hemorrhage found in patients with postresuscitation encephalopathy: characteristics of CT findings and clinical importance. AJNR Am J Neuroradiol. 2008;29:1544–9. https://doi.org/10.3174/ajnr.A1167.
10. Badejo OA, Nwafuluaku EC, Olatunji RB, Balogun JA. WHITE cerebellum sign as a dark prognostic indicator of cerebral injury: a case report. Ann Ib Postgrad Med. 2024;22:108–11.
11. Na MK, Kim W, Lim TH, Jang B, Cho Y, Choi K-S, et al. Gray matter to white matter ratio for predicting neurological outcomes in patients treated with target temperature management after cardiac arrest: a systematic review and meta-analysis. Resuscitation. 2018;132:21–8. https://doi.org/10.1016/j.resuscitation.2018.08.024.
12. Hanning U, Sporns PB, Lebiedz P, Niederstadt T, Zoubi T, Schmidt R, et al. Automated assessment of early hypoxic brain edema in non-enhanced CT predicts outcome in patients after cardiac arrest. Resuscitation. 2016;104:91–4. https://doi.org/10.1016/j.resuscitation.2016.03.018.
13. Lang M, Kenda M, Scheel M, Martola J, Wheeler M, Owen S, et al. Standardised and automated assessment of head computed tomography reliably predicts poor functional outcome after cardiac arrest: a prospective multicentre study. Intensive Care Med. 2024;50:1096–107. https://doi.org/10.1007/s00134-024-07497-2.

14. Streitberger KJ, Endisch C, Ploner CJ, Stevens R, Scheel M, Kenda M, et al. Timing of brain computed tomography and accuracy of outcome prediction after cardiac arrest. Resuscitation. 2019;145:8–14. https://doi.org/10.1016/j.resuscitation.2019.09.025.
15. Kirsch K, Heymel S, Günther A, Vahl K, Schmidt T, Michalski D, et al. Prognostication of neurologic outcome using gray-white-matter-ratio in comatose patients after cardiac arrest. BMC Neurol. 2021;21:456. https://doi.org/10.1186/s12883-021-02480-6.
16. Case N, Coppler PJ, Mettenburg J, Ratay C, Tam J, Faiver L, et al. Time-dependent association of grey-white ratio on early brain CT predicting outcomes after cardiac arrest at hospital discharge. Resuscitation. 2025. https://doi.org/10.1016/j.resuscitation.2024.110440.
17. Metter RB, Rittenberger JC, Guyette FX, Callaway CW. Association between a quantitative CT scan measure of brain edema and outcome after cardiac arrest. Resuscitation. 2011;82:1180–5. https://doi.org/10.1016/j.resuscitation.2011.04.001.
18. Hannawi Y, Muschelli J, Mulder M, Sharrock M, Storm C, Leithner C, et al. Postcardiac arrest neurological prognostication with quantitative regional cerebral densitometry. Resuscitation. 2020;154:101–9. https://doi.org/10.1016/j.resuscitation.2020.06.023.
19. Kenda M, Lang M, Nee J, Hinrichs C, Dell'Orco A, Salih F, et al. Regional brain net water uptake in computed tomography after cardiac Arrest – A novel biomarker for neuroprognostication. Resuscitation. 2024;200:110243. https://doi.org/10.1016/j.resuscitation.2024.110243.
20. Wang G, Zhang Z, Chen W, Xu X, Zhang J. Timing of brain computed tomography for predicting neurological prognosis in comatose cardiac arrest survivors: a retrospective observational study. World J Emerg Med. 2022;13:349–54. https://doi.org/10.5847/wjem.j.1920-8642.2022.080.
21. Kenda M, Scheel M, Kemmling A, Aalberts N, Guettler C, Streitberger KJ, et al. Automated assessment of brain CT after cardiac arrest—an observational derivation/validation cohort study. Crit Care Med. 2021;49:e1212. https://doi.org/10.1097/CCM.0000000000005198.
22. Rorden C, Bonilha L, Fridriksson J, Bender B, Karnath H-O. Age-specific CT and MRI templates for spatial normalization. Neuroimage. 2012;61:957–65. https://doi.org/10.1016/j.neuroimage.2012.03.020.
23. Cummins RO, Chamberlain DA, Abramson NS, Allen M, Baskett PJ, Becker L, et al. Recommended guidelines for uniform reporting of data from out-of-hospital cardiac arrest: the Utstein style. A statement for health professionals from a task force of the American heart Association, the European resuscitation Council, the heart and stroke foundation of Canada, and the Australian resuscitation Council. Circulation. 1991;84:960–75. https://doi.org/10.1161/01.cir.84.2.960.
24. Kiehl EL, Parker AM, Matar RM, Gottbrecht MF, Johansen MC, Adams MP, et al. C-GRApH: a validated scoring system for early stratification of neurologic outcome after out-of-hospital cardiac arrest treated with targeted temperature management. J Am Heart Assoc. 2017;6:e003821. https://doi.org/10.1161/JAHA.116.003821.
25. Benjamini Y, Hochberg Y, Royal Statistical Society, Oxford University Press. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57:289–300.
26. Arbelaez A, Castillo M, Mukherji SK. Diffusion-weighted MR imaging ofglobal cerebral anoxia. AJNR Am J Neuroradiol. 1999;20:999–1007.
27. Gentsch A, Storm C, Leithner C, Schroeder T, Ploner CJ, Hamm B, et al. Outcome prediction in patients after cardiac arrest: a simplified method for determination of Gray–White matter ratio in cranial computed tomography. Clin Neuroradiol. 2015;25:49–54. https://doi.org/10.1007/s00062-013-0281-3.
28. Huang BY, Castillo M. Hypoxic-ischemic brain injury: imaging findings from birth to adulthood. Radiographics. 2008;28:417–39. https://doi.org/10.1148/rg.282075066. quiz 617.
29. Tsai H, Chi C-Y, Wang L-W, Su Y-J, Chen Y-F, Tsai M-S, et al. Outcome prediction of cardiac arrest with automatically computed gray-white matter ratio on computed tomography images. Crit Care. 2024;28:118. https://doi.org/10.1186/s13054-024-04895-2.
30. Rittenberger JC, Raina K, Holm MB, Kim YJ, Callaway CW. Association between cerebral performance Category, modified Rankin Scale, and discharge disposition after cardiac arrest. Resuscitation. 2011;82:1036–40. https://doi.org/10.1016/j.resuscitation.2011.03.034.
31. Cristia C, Ho M-L, Levy S, Andersen LW, Perman SM, Giberson T, et al. The association between a quantitative computed tomography (CT) measurement of cerebral edema and outcomes in post-cardiac arrest-a validation study. Resuscitation. 2014;85:1348–53. https://doi.org/10.1016/j.resuscitation.2014.05.022.
32. Zhou F, Wang H, Jian M, Wang Z, He Y, Duan H, et al. Gray-white matter ratio at the level of the basal ganglia as a predictor of neurologic outcomes in cardiac arrest survivors: a literature review. Front Med (Lausanne). 2022. https://doi.org/10.3389/fmed.2022.847089.
33. Battey TWK, Karki M, Singhal AB, Wu O, Sadaghiani S, Campbell BCV, et al. Brain edema predicts outcome after nonlacunar ischemic stroke. Stroke. 2014;45:3643–8. https://doi.org/10.1161/STROKEAHA.114.006884.
34. Cheng-Mei S, Alvord EC JR, Berry RG. Swelling of the brain following ischemic infarction with arterial occlusion. Arch Neurol. 1959;1:161–77. https://doi.org/10.1001/archneur.1959.03840020035006.

Publisher's note