



**HAL**  
open science

# Standardised and automated assessment of head computed tomography reliably predicts poor functional outcome after cardiac arrest: a prospective multicentre study

Margareta Lang, Martin Kenda, Michael Scheel, Juha Martola, Matthew Wheeler, Stephanie Owen, Mikael Johnsson, Martin Annborn, Josef Dankiewicz, Nicolas Deye, et al.

## ► To cite this version:

Margareta Lang, Martin Kenda, Michael Scheel, Juha Martola, Matthew Wheeler, et al.. Standardised and automated assessment of head computed tomography reliably predicts poor functional outcome after cardiac arrest: a prospective multicentre study. *Intensive Care Medicine*, 2024, 50 (7), pp.1096-1107. 10.1007/s00134-024-07497-2 . hal-04671911

**HAL Id: hal-04671911**

**<https://hal.science/hal-04671911v1>**

Submitted on 16 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.




Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

ORIGINAL



# Standardised and automated assessment of head computed tomography reliably predicts poor functional outcome after cardiac arrest: a prospective multicentre study

Margareta Lang<sup>1,2</sup>, Martin Kenda<sup>3,4</sup>, Michael Scheel<sup>5</sup>, Juha Martola<sup>6</sup>, Matthew Wheeler<sup>7</sup>, Stephanie Owen<sup>7</sup>, Mikael Johnsson<sup>2</sup>, Martin Annborn<sup>1,8</sup>, Josef Dankiewicz<sup>1,9</sup>, Nicolas Deye<sup>10</sup>, Joachim Düring<sup>1,11</sup>, Hans Friberg<sup>1,11</sup>, Thomas Halliday<sup>12</sup>, Janus Christian Jakobsen<sup>13,14</sup>, Jean-Baptiste Lascarrou<sup>15</sup>, Helena Levin<sup>1,16</sup>, Gisela Lilja<sup>1,17</sup>, Anna Lybeck<sup>1,18</sup>, Peter McGuigan<sup>19,20</sup>, Christian Rylander<sup>21</sup>, Victoria Sem<sup>22</sup>, Matthew Thomas<sup>23</sup>, Susann Ullén<sup>24</sup>, Johan Undén<sup>1,25</sup>, Matt P. Wise<sup>26</sup>, Tobias Cronberg<sup>1,17</sup>, Johan Wassélius<sup>1,27</sup>, Niklas Nielsen<sup>1,8</sup>, Christoph Leithner<sup>4</sup> and Marion Moseby-Knappe<sup>1,17,28\*</sup> 

© 2024 The Author(s)

## Abstract

**Purpose:** Application of standardised and automated assessments of head computed tomography (CT) for neuro-prognostication after out-of-hospital cardiac arrest.

**Methods:** Prospective, international, multicentre, observational study within the Targeted Hypothermia versus Targeted Normothermia after out-of-hospital cardiac arrest (TTM2) trial. Routine CTs from adult unconscious patients obtained > 48 h ≤ 7 days post-arrest were assessed qualitatively and quantitatively by seven international raters blinded to clinical information using a pre-published protocol. Grey–white-matter ratio (GWR) was calculated from four (GWR-4) and eight (GWR-8) regions of interest manually placed at the basal ganglia level. Additionally, GWR was obtained using an automated atlas-based approach. Prognostic accuracies for prediction of poor functional outcome (modified Rankin Scale 4–6) for the qualitative assessment and for the pre-defined GWR cutoff < 1.10 were calculated.

**Results:** 140 unconscious patients were included; median age was 68 years (interquartile range [IQR] 59–76), 76% were male, and 75% had poor outcome. Standardised qualitative assessment and all GWR models predicted poor outcome with 100% specificity (95% confidence interval [CI] 90–100). Sensitivity in median was 37% for the standardised qualitative assessment, 39% for GWR-8, 30% for GWR-4 and 41% for automated GWR. GWR-8 was superior to GWR-4 regarding prognostic accuracies, intra- and interrater agreement. Overall prognostic accuracy for automated GWR (area under the curve [AUC] 0.84, 95% CI 0.77–0.91) did not significantly differ from manually obtained GWR.

\*Correspondence: marion.moseby\_knappe@med.lu.se

<sup>28</sup> Department of Rehabilitation, Skåne University Hospital, 22185 Lund, Sweden

Full author information is available at the end of the article

Margareta Lang and Martin Kenda have contributed equally to this work. Christoph Leithner and Marion Moseby-Knappe have contributed equally to this work.

**Conclusion:** Standardised qualitative and quantitative assessments of CT are reliable and feasible methods to predict poor functional outcome after cardiac arrest. Automated GWR has the potential to make CT quantification for neuroprognostication accessible to all centres treating cardiac arrest patients.

**Keywords:** Cardiac arrest, Computed tomography, Prognosis, Hypoxic–ischaemic encephalopathy, GWR

## Introduction

According to the European Resuscitation Council (ERC) and the European Society for Intensive Care Medicine (ESICM), “diffuse and extensive anoxic injury” on neuroimaging is predictive of poor functional outcome after cardiac arrest [1]. Head computed tomography (CT) is widely available and is frequently used for neuroprognostication [2–4]. Recent meta-analyses conclude that the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) level of evidence for CT to predict outcome after cardiac arrest is very low [5–7]. Signs of “diffuse and extensive anoxic injury” seen as a reduced differentiation between grey and white matter and/or effacement of the cerebral sulci on CT correlate well with elevated levels of neuronal injury markers [8] and histopathological severity of hypoxic–ischaemic encephalopathy (HIE) [9].

In clinical practice, CTs are usually assessed qualitatively using a non-standardised approach [10]. Some specialised centres use non-standardised quantitative methods such as grey–white-matter ratio (GWR) placing regions of interest at the basal ganglia and/or in (sub)cortical regions to quantify oedema [7]. Manual assessments carry the risk of interrater variability and standardisation and/or automation may be necessary to ensure a safe translation from research to clinical routine [11–14]. In stroke imaging, automated quantification of non-contrast CTs is routinely used [15]. This has not yet been achieved for neuroprognostication and a few single-centre retrospective studies with automated quantification of GWR after cardiac arrest have been published [16, 17].

Based on our retrospective studies with adult out-of-hospital cardiac arrest patients, we established Standard Operating Procedures (SOPs) for qualitative (visual interpretation) and quantitative (GWR) CT assessments for clinical use [13, 18–20]. In line with the previous studies [18, 19, 21–24], our recent pilot study from the Target Temperature Management after Out-of-hospital Cardiac Arrest (TTM)-trial confirmed that GWR was most accurate at the basal ganglia level where a cutoff < 1.10 predicted poor functional outcome with 100% specificity [18].

We here present the results of a prospective observational substudy of an international multicentre trial in which we applied our previously published criteria for

## Take-home message

In a prospective international trial, blinded and standardised qualitative, manual and automated quantitative assessments of CT reliably predicted poor functional outcome in unconscious patients after out-of-hospital cardiac arrest

standardised qualitative and quantitative CT assessment as well as an atlas-based automated GWR method for neuroprognostication after cardiac arrest [14]. Our main hypotheses were that “Definite signs of severe HIE”, by standardised assessment and manually or automatically obtained  $GWR < 1.10$ , would predict poor functional outcome without false positives in CTs performed 48 h–7 days after cardiac arrest [13].

## Method

### Study design

This prospective international multicentre observational study (Clinicaltrials.gov NCT03913065) was a substudy of the Targeted Hypothermia versus Targeted Normothermia after out-of-hospital cardiac arrest (TTM2) trial [25] (Clinicaltrials.gov NCT02908308). The design and statistical analysis plan of this substudy has previously been published [13].

### Patient selection and ethics

Between November 2017 and January 2020, the TTM2-trial consecutively screened unconscious patients  $\geq 18$  years admitted to hospital after out-of-hospital cardiac arrest of a presumed cardiac or unknown cause [25]. Approval was waived/obtained from the appropriate ethics committees. The trial was performed in accordance with the ethical standards laid down in the Declaration of Helsinki and its later amendments [26]. Consent was obtained from legal representatives and/or patients according to local legislation.

Thirteen sites from Sweden, Germany, France, and United Kingdom that routinely use CT for neuroprognostication in patients unconscious > 48 h post-arrest participated (electronic supplementary material [ESM] Table E1). Unconsciousness was defined as not obeying verbal commands. Included patients were managed according to the TTM2-trial protocol regarding randomisation, clinical management, neurological

prognostication, decisions on withdrawal of life-sustaining therapy, and follow-up [25, 27–29].

#### **Data collection and technical requirements**

All types of scanners and software were permitted. Technical prerequisites were availability of axial slices of 4–5 mm thickness obtained with a tube voltage of 120 kV.

#### **CT assessments**

CTs with artefacts or structural lesions interfering with reliable evaluation were excluded. Five radiologists and two neurologists with 3–15 years of experience of CTs after cardiac arrest from four countries evaluated images individually using a virtual private network (VPN) secured platform (Human Observer Net) [30] (ESM Table E2). Raters were blinded to all information except the patients age, since brain volume may decrease with age, and thus, this information was considered necessary for assessing the extent of cerebral oedema. The raters received approximately 30 min of training for the software used for evaluations, but unrelated to the actual rating of images. Raters were encouraged to have the SOP accessible during ratings.

#### **Standardised operating procedures for qualitative assessments**

Axial images were evaluated at four levels; brain stem and cerebellum, basal ganglia, frontoparietal cortex at corona radiata level, and at high convexity cortex (ESM Fig. E1A) [13]. The raters confirmed or declined; “Are there definite signs of severe HIE defined as complete or near complete loss of grey–white-matter differentiation at the basal ganglia level and in the frontoparietal cortex with additional evidence of brain swelling/sulcal effacement?”

#### **Standardised operating procedures for quantitative assessments**

Circular 0.1 cm<sup>2</sup> regions of interest were manually placed at the basal ganglia level in the putamen, the caudate nucleus (caput), the posterior limb of the internal capsule, and the genu corpus callosum bilaterally (ESM Fig. E1B) [13].

#### **Automated measurements**

The software pipeline for automated GWR determinations has been published [17]. Images were co-registered to a freely available standard brain atlas and mean Hounsfield Units were quantified in each individual CT space using inversely transferred probabilistic tissue maps [31, 32] (ESM Figs. E2–E3).

#### **GWR calculations**

GWR was calculated as the sum of the radiodensity of the grey matter regions of interest divided by the sum of the radiodensity of the white matter regions of interest (ESM Fig. E1B). The GWR-8 model included all eight regions of interest. The GWR-4 model and the automated GWR only included the measurements in the putamen and in the posterior limb of the internal capsule.

#### **Outcome assessment**

Functional outcome by the modified Rankin Scale (mRS) was assessed by a trained outcome assessor at a structured face-to-face or telephone follow-up, at six months after randomisation. Functional outcome was dichotomised into good (mRS 0–3) and poor (mRS 4–6) [25, 27, 33].

#### **Statistical analysis**

The results are reported according to the Standards of Reporting Diagnostic Studies [34] and the Standards for Studies of Neurological Prognostication in Comatose Survivors of Cardiac Arrest [35]. Continuous variables are reported as median (interquartile range, IQR) or means ( $\pm$  standard deviation) and categorical variables in numbers (percentages). Sensitivities and specificities for prediction of poor functional outcome, and negative and positive predictive values are presented with 95% confidence intervals (CI) calculated with Wilson’s method. Results from the manual standardised assessments are presented separately for each rater and as median (min–max) of all raters. For GWR, we decided to apply the cutoff  $< 1.10$ , since this yielded a 100% specificity for poor outcome prediction in our pilot study [18]. In addition, we analysed the pre-specified GWR cutoff  $< 1.15$ . The overall prognostic performance of GWR for good versus poor functional outcome was assessed by the area under the receiver-operating characteristic curve (AUC) with 95% CI. AUC was classified as;  $< 0.60 =$  failure,  $0.60–0.70 =$  poor,  $0.70–0.80 =$  fair,  $0.80–0.90 =$  good, and  $0.90–1.00 =$  excellent [36]. The mean AUC for manual GWR was compared to the automated GWR using DeLong.

The interrater agreement between the blinded raters was calculated with Fleiss’ kappa. Intra-rater agreements for 20% of the images re-evaluated by each rater (identical for all raters) were analysed with Cohen’s kappa. The strength of the agreement was classified as kappa ( $\kappa$ );  $< 0.20 =$  poor,  $0.21–0.40 =$  fair,  $0.41–0.60 =$  moderate,  $0.61–0.80 =$  good, and  $0.81–1.00 =$  very good for Fleiss’ and Cohen’s kappa [37–39].

CTs performed 48 h–7 days post-arrest were included in our prospective cohort. To assess the accuracy of automated GWR, we included all patients with CTs performed  $\leq 7$ d in a post hoc cohort. We

assessed the impact of timing on prognostic accuracies for automated GWR < 1.10 in time-windows < 2 h, 2–6 h, > 6–48 h, > 48–96 h, and > 96–168 h after cardiac arrest.

Sensitivities and specificities with 95% CI are also presented separately for patients randomised to hypothermia and normothermia within the prospective cohort for each rater for the standardised qualitative assessment, GWR-8 and GWR-4 at cutoff < 1.10 and for automated GWR < 1.10.

We further examined whether "severe HIE" by qualitative rating and GWR-8 cutoff < 1.10 evaluated by four or more raters corresponded with pathological findings of routine prognostic methods.

Statistical analyses were performed with IBM SPSS Statistics (SPSS Statistics for Windows, Version 29.0.0.0 Armonk, NY: IBM Corp) and R, version 4.0.4 (The R Foundation for Statistical Computing).

## Results

### Patient demographics

387/635 (60.9%) patients underwent at least one head CT  $\leq 7$  days after cardiac arrest. Forty-six patients were excluded due to unmet technical requirements or the presence of other intracranial pathologies (Fig. 1).  $N=140$  patients were unconscious and examined with CT 48 h–7 days after cardiac arrest, thereby meeting the inclusion criteria for our prospective cohort. Together with the prospective cohort, further  $N=201$  patients examined  $\leq 48$  h post-arrest were included in a post hoc analysis using automated GWR [13]. The median age of our prospective cohort was 68 years, 76% were male, and the median time from cardiac arrest to CT was 84 h (IQR 66–109) (Table 1). Poor functional outcome was more frequent among included than excluded patients (75% versus 50%). Prognostic accuracies for predicting poor functional outcome by other routine prognostic tests used within the prospective cohort are displayed in ESM Table E3. Life-sustaining therapy was withdrawn due to poor neurological prognosis in 66/140 (47%), at median 121 h (IQR 98–156) post-arrest.

### Prediction of functional outcome

#### Blinded qualitative assessment

The standardised qualitative assessment of "definite signs of severe HIE" predicted poor outcome with 100% specificity and 100% positive predictive value in all seven raters (980 ratings overall) (Table 2, ESM Table E4). The median sensitivity of all raters was 37% (min–max 11–61%). Inter-rater agreement between raters was moderate ( $\kappa=0.6$ ) (ESM Table E5). Intra-rater agreement

ranged from fair ( $\kappa=0.33$ ) to very good ( $\kappa=0.93$ ) (ESM Table E6).

#### Blinded quantitative assessment

Median GWR was significantly lower in poor outcome patients compared to good outcome patients,  $p < 0.0001$  (ESM Table E7). In both models, GWR < 1.10 predicted poor functional outcome with 100% specificity and 100% positive predictive value in all 980 ratings (Table 2, ESM Table E4, ESM Fig. E4A–B). A cutoff < 1.15 yielded two false-positive predictions of poor outcome for GWR-8 (median specificity 100%, min–max 94–100%) and five false-positive predictions for GWR-4 (median specificity 100%, min–max 91–100%) (ESM Table E4, ESM Table E8, ESM Fig. E4A–B). Median sensitivities were higher for GWR-8 than GWR-4 at cutoff < 1.10 (39% versus 30%) and at cutoff < 1.15 (48% versus 38%). Median AUC of all seven raters was 0.86 for GWR-8 and 0.81 for GWR-4 (Fig. 2A–C).

The interrater agreement for GWR-8 ranged from good ( $\kappa=0.72$ ) for < 1.10 to very good ( $\kappa=0.83$ ) for < 1.15 (ESM Table E5). The interrater agreement for GWR-4 was moderate ( $\kappa=0.57$ ) at both cutoffs. Examples of CTs with low, moderate, and high interrater agreement are found in the ESM Fig. E5A–C. Intra-rater agreements for GWR-8 ranged from moderate to very good ( $\kappa=0.58$ –1) and from fair to good for GWR-4 ( $\kappa=0.26$ –0.75) (ESM Table E9).

#### Effect of targeted temperature management

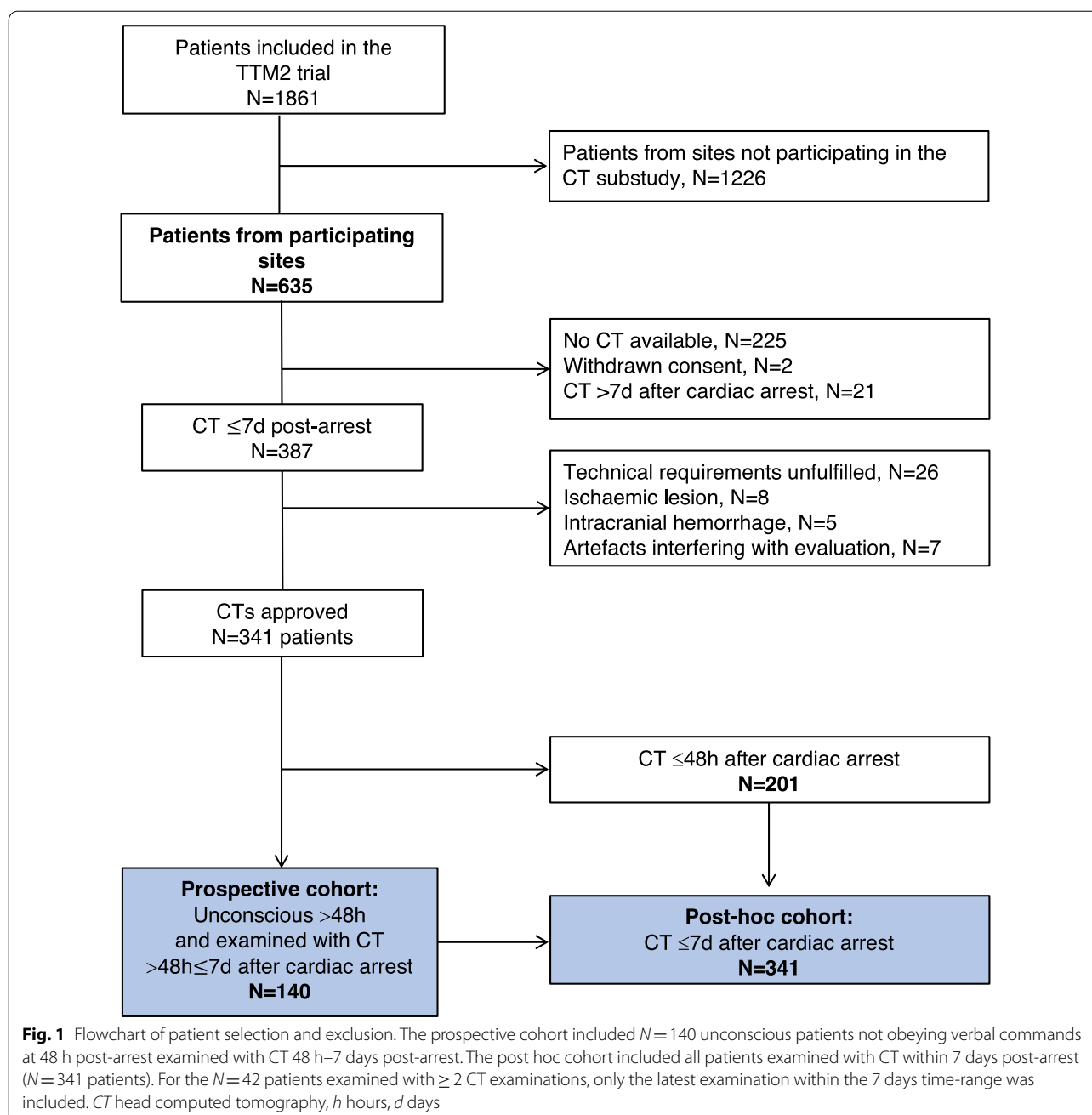
ESM Table E10 depicts prognostic accuracy measures for hypothermia and normothermia groups. We found no clinically relevant effect of temperature allocation on prognostic accuracies for prediction of poor functional outcome.

#### Congruence of HIE with routine prognostic methods

In patients with "severe HIE" in qualitative assessment or a GWR-8 < 1.10 diagnosed by the majority of raters, 96% of patients had high levels of neuron-specific enolase (NSE) (> 60 ng/ml), 44–53% had bilaterally absent somatosensory evoked potentials (SSEP) N20 potentials and 66–68% had a highly malignant electroencephalogram (EEG) (ESM Tables E11A–B). The lowest congruence was seen for bilaterally absent corneal and pupillary reflexes (29–37%).

#### Automated GWR

Automated GWR < 1.10 predicted poor outcome with 100% specificity (95% CI 90–100) and 41% sensitivity (95% CI 32–51) (Table 2, Fig. 3, ESM Table E12). AUC was 0.84 (95% CI 0.77–0.91) with no significant difference



compared to the average of manually determined GWR-8 or GWR-4 (Fig. 2C). Automated GWR was significantly lower in poor outcome patients, regardless of temperature management,  $p < 0.001$ , ESM Fig. E6.

#### Post hoc analysis of automated GWR

GWR was significantly lower in poor outcome patients compared to good outcome patients at all timepoints except for CT examinations performed  $< 2$  h post-arrest (ESM Table E13). In CTs obtained  $\leq 7$  days after

cardiac arrest ( $N=341$ ), overall specificity of automated GWR  $< 1.10$  was 99% (95% CI 96–100) (ESM Table E12). Poor outcome was incorrectly predicted in one patient, probably due to a lacunar infarction and enlarged perivascular spaces (ESM Fig. E7). Sensitivity for automated GWR  $< 1.10$  increased gradually from 9% (95% CI 2–27%) for CTs performed  $< 2$  h post-arrest to a peak of 48% (95% CI 37–59%) for examinations performed  $> 48$ –96 h (Fig. 3, ESM Table E12).



**Table 1 Characteristics of included and excluded patients**

	Post-hoc cohort CT ≤ 7 days (N = 341)		Excluded (N = 294)
	Prospective CT 48 h–7days (N = 140)	Early CT ≤ 48 h (N = 201)	
<b>Demographics characteristics</b>			
Age (years)	68 (59–76)	64 (56–73)	67 (55–74)
Male	106 (76)	158 (79)	229 (78)
<b>CT</b>			
Hours from cardiac arrest to CT	84 (66–109)	3 (2–4)	–
<b>Characteristics of cardiac arrest</b>			
Minutes from cardiac arrest to ROSC	30 (20–45)	26 (15–40)	28 (18–41)
First monitored rhythm on ECG shockable	94 (67)	141 (70)	218 (74)
Bystander witnessed arrest	136 (97)	190 (95)	269 (92)
Bystander CPR performed	106 (76)	163 (81)	230 (78)
Cardiac arrest at the place of residence	86 (61)	109 (54)	159 (54)
<b>Characteristics on hospital admission</b>			
Corneal reflexes bilaterally absent	58/84 (69)	56/91 (62)	94/139 (68)
Pupillary reflexes bilaterally absent	47/120 (39)	61/156 (39)	78/226 (35)
ST-segment elevation myocardial infarction	58 (41)	46/151 (30)	110/290 (38)
Circulatory shock	37 (26)	55 (27)	81 (28)
Randomized to hypothermia	70 (50)	98 (49)	145 (49)
Arterial lactate level mmol/L	6.21 (± 3.96) <sup>a</sup> 139	6.54 (± 4.63) <sup>a</sup> 193	5.77 (± 4.35) <sup>a</sup> 280
<b>Medical history</b>			
Hypertension	57/132 (43)	75 (37)	101/281 (36)
Dementia	0 (0)	3 (1)	7 (2)
Cerebrovascular disease	9 (6)	10 (5)	23 (8)
Hemiplegia	2 (1)	2 (1)	7 (2)
Diabetes	26 (19)	33 (16)	57 (20)
Myocardial infarction	23 (16)	21 (10)	45 (15)
Heart failure NYHA III or IV	17 (12)	12 (6)	35 (12)
Renal failure	6 (4)	9 (4)	15 (5)
Severe liver failure	0 (0)	2 (1)	4 (1)
Charlson comorbidity index	3 (2–4)	3 (1–4)	3 (2–5)
<b>Withdrawal of treatment</b>			
Withdrawal for any reason	85 (61)	66 (33)	116 (39)
Withdrawal poor neurological prognosis	66 (47)	26 (13)	42 (14)
Hours from cardiac arrest to withdrawal due to poor neurological prognosis	121 (98–156)	110 (92–139)	106 (106–143)
<b>Functional outcome after 6 months</b>			
Good, modified Rankin Scale 0–3	35 (25)	100 (50)	134 (46)
Poor, modified Rankin Scale 4–6	105 (75)	101 (50)	159 (54)
mRS 0	13/137 (9)	37/189 (20)	39/269 (14)
mRS 1	8/137 (6)	18/189 (10)	24/269 (9)
mRS 2	12/137 (9)	25/189 (13)	45/269 (17)
mRS 3	2/137 (1)	11/189 (6)	6/269 (2)
mRS 4	0/137 (0)	5/189 (3)	4/269 (1)
mRS 5	5/137 (4)	0/189 (0)	3/269 (1)
mRS 6	97/137 (71)	93/189 (49)	148/269 (55)

Results are presented in numbers (percentages), in mean (± standard deviation) or in median (interquartile range). The low denominator for some of the variables is due to missing data

CT head computed tomography, NYHA New York Heart Association, ROSC return of spontaneous circulation, ECG electrocardiogram, CPR cardiopulmonary resuscitation, mRS modified Rankin Scale

<sup>a</sup> Indicates the number of available lactate samples

**Table 2 Prediction of poor functional outcome by individual raters**

Method rater	Specificity (95% CI)	Sensitivity (95% CI)	TP	TN	FP	FN	PPV (95% CI)	NPV (95% CI)	N
<b>Qualitative assessment</b>									
1	100 (90–100)	11 (7–19)	12	35	0	93	100 (76–100)	27 (20–36)	140
2	100 (90–100)	37 (29–47)	39	35	0	66	100 (91–100)	35 (26–44)	140
3	100 (90–100)	36 (28–46)	38	35	0	67	100 (91–100)	34 (26–44)	140
4	100 (90–100)	61 (51–70)	64	35	0	41	100 (94–100)	46 (35–57)	140
5	100 (90–100)	35 (27–45)	37	35	0	68	100 (91–100)	34 (26–44)	140
6	100 (90–100)	57 (48–66)	60	35	0	45	100 (94–100)	44 (33–55)	140
7	100 (90–100)	40 (31–50)	42	35	0	63	100 (92–100)	36 (27–46)	140
<b>GWR-8 &lt; 1.10</b>									
1	100 (90–100)	31 (23–41)	33	35	0	72	100 (90–100)	33 (25–42)	140
2	100 (90–100)	34 (26–44)	36	35	0	69	100 (90–100)	34 (25–43)	140
3	100 (90–100)	38 (29–48)	40	35	0	65	100 (91–100)	35 (26–45)	140
4	100 (90–100)	46 (37–55)	48	35	0	57	100 (93–100)	38 (29–48)	140
5	100 (90–100)	47 (37–56)	49	35	0	56	100 (93–100)	38 (29–49)	140
6	100 (90–100)	39 (30–49)	41	35	0	64	100 (91–100)	35 (27–45)	140
7	100 (90–100)	46 (37–55)	48	35	0	46	100 (93–100)	43 (33–54)	140
<b>GWR-4 &lt; 1.10</b>									
1	100 (90–100)	30 (22–39)	31	35	0	74	100 (89–100)	32 (24–41)	140
2	100 (90–100)	18 (12–27)	19	35	0	86	100 (83–100)	29 (22–38)	140
3	100 (90–100)	22 (15–31)	23	35	0	82	100 (86–100)	30 (22–39)	140
4	100 (90–100)	27 (19–36)	28	35	0	77	100 (88–100)	31 (23–40)	140
5	100 (90–100)	31 (23–41)	33	35	0	72	100 (90–100)	33 (25–42)	140
6	100 (90–100)	28 (20–37)	29	35	0	76	100 (88–100)	32 (24–41)	140
7	100 (90–100)	37 (29–47)	39	35	0	66	100 (91–100)	35 (26–44)	140
<b>Automated GWR &lt; 1.10</b>	100 (90–100)	41 (32–51)	43	35	0	62	100 (92–100)	36 (27–46)	140

Prognostic accuracies for prediction of poor functional outcome (mRS 4–6) at 6 months post-arrest in  $N=140$  patients included in the prospective cohort examined 48 h–7 days post-arrest. Results are presented with 95% confidence intervals for each of the seven raters individually for the standardised qualitative assessment, manual GWR-8 and manual GWR-4 at pre-specified cutoff < 1.10, and for the automated GWR at pre-specified cutoff < 1.10

GWR grey–white-matter ratio, TP true positive (pathological CT and poor functional outcome), TN true negative (normal CT and good functional outcome), FP false positive (pathological CT and good functional outcome), FN false negative (normal CT and poor functional outcome)

## Discussion

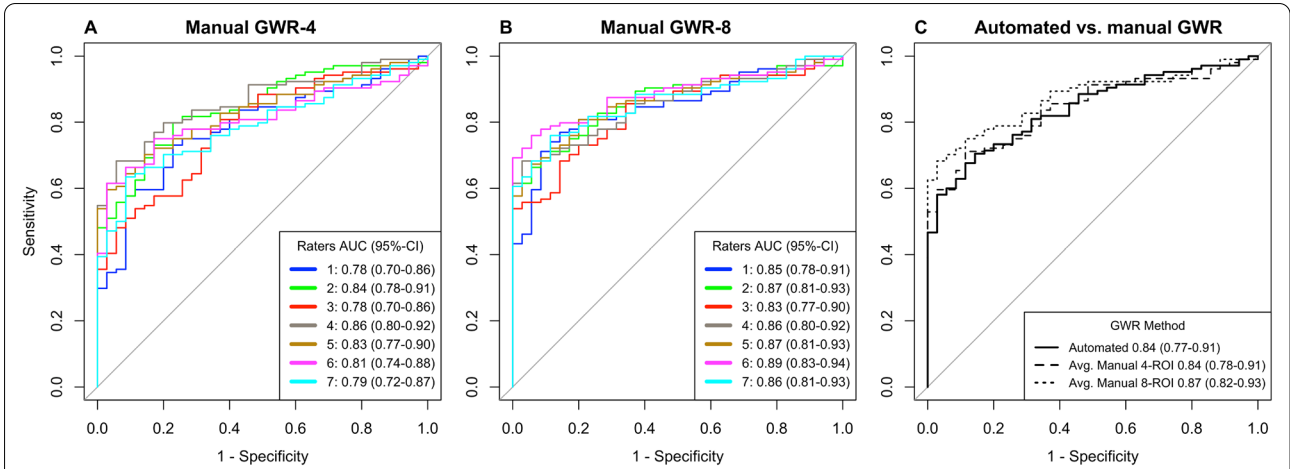
In this prospective multicentre study, evaluating three different methods of diagnosing severe hypoxic ischaemic injury on CT for prediction of poor functional outcome after cardiac arrest, we validate pre-published standardised criteria and evaluate GWR cutoff < 1.10 for manual and automated assessments [13]. We conclude that CT is a highly specific prognostic tool for neuroprognostication, regardless of assessment method, with highest sensitivities for poor outcome prediction when performed 48–96 h post-arrest. GWR determination at the basal ganglia level < 1.10 performed either manually or automated offer a more objective measure of HIE with reduced interrater variability.

CT is a guideline-recommended predictor of outcome after cardiac arrest with very low quality of evidence [1, 5, 7, 40]. The main concerns raised by ERC/ESICM include the lack of multicentre validation and standardised

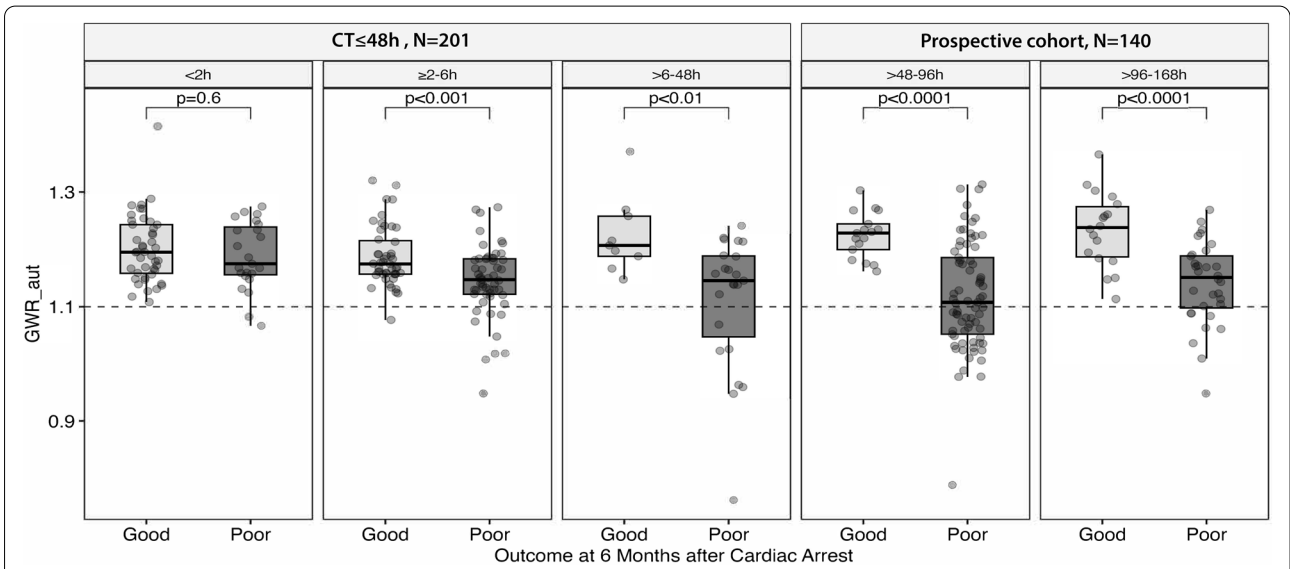
assessments of both qualitative and quantitative methods [1]. Our study provides a framework that is easy to use in clinical practice and addresses several concerns raised in recent publications [5–7, 10, 41].

Our standardised qualitative criteria define signs of severe HIE as a “complete or nearly complete loss of grey-white matter differentiation in the basal ganglia and in the frontoparietal cortex with additional evidence of brain swelling/sulcal effacement” [13]. A visual interpretation according to a checklist with mandatory evaluation at several levels of the brain had to be completed before reaching a conclusion. This qualitative assessment predicted poor outcome with 0% false-positive rate in 980 blinded ratings overall. In line with the previous qualitative studies, sensitivities for individual raters ranged between 11 and 61% for imaging performed 48 h–7 days post-arrest [18, 20, 41].





**Fig. 2 A–C** Manual and automated GWR assessments for overall prediction of functional outcome. Comparison of manual and automated GWR for prediction of good versus poor functional outcome (modified Rankin Scale 0–3 versus 4–6) at 6 months post-arrest in  $N = 140$  patients included in the prospective cohort. The area under the receiver-operating characteristics curve (AUC) is presented with 95% confidence intervals. Results are presented separately for all raters for the manual GWR-4 (**A**) and the manual GWR-8 (**B**). **C** Displays the prognostic accuracy for the automated measurements versus mean of the seven raters' manual assessments of GWR-4 and GWR-8, respectively. In **C**, the AUC for automated GWR did not differ significantly between mean manual GWR-8 ( $p = 0.10$ ) or between mean manual GWR-4 ( $p = 0.84$ )



**Fig. 3** Automated GWR for prediction of functional outcome. Automated GWR and functional outcome (modified Rankin Scale 0–3 versus 4–6) at 6 months post-arrest grouped by timing of CT acquisition in hours after cardiac arrest. Examinations are evaluated in the following time-windows after cardiac arrest;  $< 2$  h:  $N = 67$ ,  $\geq 2-6$  h:  $N = 102$ ,  $> 6-48$  h:  $N = 32$ ,  $> 48-96$  h:  $N = 90$ ,  $> 96$  h–168 h:  $N = 50$ . Automated GWR was significantly lower in poor outcome patients than in good outcome patients at all timepoints  $\geq 2$  h post-arrest. *ns* not significant,  $**p > 0.01$ ,  $***p < 0.001$ ,  $****p < 0.0001$ . Overall, automated GWR  $< 1.10$  predicted poor outcome with 1 false pathological prediction at 2 h post-arrest, with an overall specificity of 99% (95% CI 96–100) and 30% sensitivity (95% CI 24–36). Exact prognostic accuracies and a contingency table for each timepoint in this figure can be seen in ESM Table E12. Median GWR values for good and poor outcome patients are displayed in ESM Table E13. The CT image and case description of the 1 false-positive patient is displayed in ESM Fig. E7

Both the ERC/ESICM and the American Neurocritical Care Society recommendations use similar, undefined terminology to describe signs of severe HIE on CT; "diffuse", "extensive" or "bilaterally across vascular

territories", with a "loss of grey-white-matter differentiation" [1, 5]. While our standardised qualitative criteria may offer a more precise definition of severe HIE than those given in the current guidelines, it achieved only

moderate interrater reliability. The CT evaluation in our study was mostly performed by experienced raters (3–15 years with CTs of cardiac arrest patients). Rater experience may impact both sensitivity and specificity of CT evaluation using our SOP. This should be kept in mind when implementing our CT analysis in clinical routine. Future improvements to improve interrater reliability are necessary and may include a better standardisation of windowing during visual analysis, standards regarding decision in case of residual grey–white differentiation and awareness of the effects from residual contrast agents. In contrast to clinical practice, our raters only had one CT available for analysis and did not have access to pre-cardiac arrest CTs. We plan a subsequent study using serial CTs to evaluate whether an analysis of changes in grey–white-matter differentiation and brain volume over time improves prognostic accuracy.

GWR is the only guideline-recommended method to quantify the extent of HIE on CT and can be applied with routine radiological software, but there is no consensus on the number, size, and exact location of regions of interest [1, 10, 22, 42–44]. Based on previous investigations and our retrospective pilot study, we chose to validate manually placed 0.1 cm<sup>2</sup> regions of interest at the basal ganglia level [18]. Importantly, we included the instruction to place the regions of interest in a sub-region with a radiodensity representative of the entire anatomical target region and to avoid potential confounders (artefacts, calcifications, lacunar infarcts, etc.). We confirmed that both manual GWR models had a maximal specificity at GWR < 1.10, which is in accordance with the previous studies [6, 7, 18]. As expected, sensitivities increased at cutoff < 1.15 at cost of a slightly decreased specificity. None of the false positives through quantitative measurements fulfilled criteria for "severe HIE" with the qualitative assessment, underlining the potential value of combining both approaches. As in our pilot study, GWR-8 was consistently superior to GWR-4 concerning prognostic accuracies, intra- and interrater agreements [18]. A possible explanation for the higher accuracy of GWR-8 is the reduction of noise due to the larger number of regions of interest [18, 45]. GWR-8 was superior to the qualitative assessment for some raters and the interrater reliability for GWR-8 was superior to that of qualitative assessments—highlighting a potential advantage of quantification. We presume that the interrater reliability of manual GWR could be further improved by applying stricter instructions for measurements within anatomical regions or using non-circular and/or larger regions of interest.

Automated atlas-based GWR measurements offer an alternative to manual measurements unaffected by interrater variability and could increase the availability

of GWR for hospitals without on-site neuroradiologic expertise [17]. A few previous studies evaluate automated GWR quantification and they are limited by single-centre, retrospective designs, and early assessment of functional outcome [16, 17]. The prognostic accuracy of automated GWR < 1.10 in our prospective cohort was as good as the manually assessed GWR with 40% sensitivity at 100% specificity. This performance is also in the range of manual and automated GWR from CTs performed > 24 h post-arrest in the previous studies [18, 20, 24, 46], routine EEG, and SSEP [7]. Except for one study on early CTs, 1.10 was the lowest reported GWR cut-off with 100% specificity thus far [47]. Overall, automated GWR < 1.10 performed within 7 days post-arrest had one false-positive prediction of poor outcome in a patient with a subcortical low attenuating area close to putamen, most likely an old lacunar infarction or perivascular space, but with intact overall grey–white-matter differentiation. The use of automated GWR relies on anatomic landmarks and its use must include a quality check for co-registration and exclusion of artefacts or acute brain pathologies potentially interfering with measurements [10]. Future studies on larger cohorts should investigate whether machine learning can predict outcome from CTs after cardiac arrest with superior accuracy compared to our human rater-based approach [48].

Data from our current and previous studies do not suggest that CT can predict good outcome/absence of severe HIE. Future studies, for example using analysis of serial CTs, should re-investigate this issue.

Our results on optimal timing support guideline recommendations [1, 5] that CTs performed 48–96 h have higher sensitivity for predicting poor outcome than examinations performed within the first hours post-arrest [5, 17, 18, 20, 23, 24, 49]. Examinations performed on hospital admission are often routinely used to exclude cerebral causes of unconsciousness and may be too early to detect HIE for most patients. The increase in sensitivity within the first days corresponds with developing HIE. The higher sensitivity of later examinations is in line with previous observations and supports the notion that an optimal time window of a few days exists for neuroprognostic CT's [17]. We found no clinically relevant effect of temperature allocation on prognostic accuracies for prediction of poor functional outcome. When performed at an optimal timepoint and analysed using standardised interpretation, combining CT with other prognostic methods with higher sensitivities such as EEG or NSE could increase the number of correctly identified poor outcome patients.

## Strengths and limitations

Strengths of our study include the prospective, multi-centre design with standardised criteria for neuroprognostication and withdrawal of life-sustaining therapy and a structured assessment of functional outcome at 6 months. CT's were prospectively performed in unconscious patients at a timepoint clinically most relevant for neuroprognostication. Radiological assessments were blinded and performed by multiple raters from different countries according to a pre-published protocol using standardised radiological criteria and pre-defined cutoffs. A comparison with automated GWR within the same cohort further strengthens our results.

Our main limitation is imprecision due to sample size [5]. A substantial proportion of patients were examined before the pre-specified time point, reasonably as part of clinical practice, and thus reported as part of a post hoc cohort examined  $\leq 7$  days. Additional patients did not receive CT  $> 48$  h, because they underwent magnetic resonance imaging rather than CT, used other prognostic methods or because CT could not be performed for logistical reasons.

In contrast to clinical practice, to standardise the protocol within a clinical trial, our raters only had axial CT images available, separately performed qualitative and quantitative assessments, their rating were final, and they did not have the possibility to discuss their results with colleagues.

Patients included in the TTM2-trial had a presumed cardiac or unknown cause of cardiac arrest and results may differ from other causes of arrest [25]. The conservative approach to prognostication within the TTM2-trial was designed to limit the risk of self-fulfilling prophecies, reflected in the longer times to withdrawal of life-sustaining therapy in our prospective cohort [25, 28, 29]. Nonetheless, despite the blinded CT evaluations in this study the risk of self-fulfilling prophecies cannot be entirely excluded, since local radiologists CT reports were available to treating physicians. Our results should be validated in a cohort where withdrawal of treatment was not performed.

## Conclusion

The combination of a structured qualitative assessment of severe HIE with a quantitative assessment at the basal ganglia level (GWR-8 or automated GWR  $< 1.10$ ) allows the prediction of poor functional outcome after cardiac arrest with high specificity and moderate sensitivity. CT should be considered in patients unconscious later than 48 h after cardiac arrest using a multimodal approach to neuroprognostication. Automated GWR could help avoid errors compared to manual ratings and make head CT quantification more accessible.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s00134-024-07497-2>.

## Author details

<sup>1</sup> Department of Clinical Sciences Lund, Lund University, Lund, Sweden. <sup>2</sup> Department of Radiology, Helsingborg Hospital, Helsingborg, Sweden. <sup>3</sup> Berlin Institute of Health at Charité-Universitätsmedizin Berlin, BIH Biomedical Innovation Academy, BIH Charité Junior Digital Clinician Scientist Program, Universitätsmedizin Berlin, Berlin, Germany. <sup>4</sup> Department of Neurology and Experimental Neurology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität and Humboldt-Universität Zu Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. <sup>5</sup> Department of Neuroradiology, Charité-Universitätsmedizin Berlin, Berlin, Germany. <sup>6</sup> HUS Medical Imaging Center, Radiology, University of Helsinki, Helsinki University Hospital, Helsinki, Finland. <sup>7</sup> University Hospital of Wales, Cardiff and Vale University Health Board, Cardiff, Wales, UK. <sup>8</sup> Department of Anaesthesia and Intensive Care, Helsingborg Hospital, Helsingborg, Sweden. <sup>9</sup> Department of Cardiology, Skåne University Hospital, Lund, Sweden. <sup>10</sup> Department of Medical and Toxicological Intensive Care Unit, Inserm UMR-S 942, Assistance Publique des Hopitaux de Paris, Lariboisière University Hospital, Paris, France. <sup>11</sup> Department of Anaesthesia and Intensive Care, Skåne University Hospital, Malmö, Sweden. <sup>12</sup> Department of Operation and Intensive Care, Linköping University Hospital, Linköping, Sweden. <sup>13</sup> Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. <sup>14</sup> Copenhagen Trial Unit, Centre for Clinical Intervention Research, Capital Region of Denmark, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. <sup>15</sup> Médecine Intensive Réanimation, Movement-Interactions-Performance, Nantes Université, CHU Nantes, MIP, UR 4334, 44000 Nantes, France. <sup>16</sup> Department of Research and Education, Skåne University Hospital, Lund, Sweden. <sup>17</sup> Department of Neurology, Skåne University Hospital, Lund, Sweden. <sup>18</sup> Department of Anaesthesia and Intensive Care, Skåne University Hospital, Lund, Sweden. <sup>19</sup> Regional Intensive Care Unit, Royal Victoria Hospital, Belfast, UK. <sup>20</sup> Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK. <sup>21</sup> Anaesthesia and Intensive Care, Department of Surgical Sciences, Uppsala University, Uppsala, Sweden. <sup>22</sup> Department of Anaesthesia and Intensive Care, Central Hospital of Karlstad, Karlstad, Sweden. <sup>23</sup> Intensive Care Unit, University Hospitals Bristol and Weston, Bristol, UK. <sup>24</sup> Clinical Studies Sweden-Forum South, Skåne University Hospital, Lund, Sweden. <sup>25</sup> Department of Operation and Intensive Care, Hallands Hospital Halmstad, Halmstad, Sweden. <sup>26</sup> Adult Critical Care, University Hospital of Wales, Cardiff, UK. <sup>27</sup> Department of Medical Imaging and Physiology, Skåne University Hospital, Lund, Sweden. <sup>28</sup> Department of Rehabilitation, Skåne University Hospital, 22185 Lund, Sweden.

## Author contributions

Concept and design: ML, MK, CL, MMK, TC, NN, and MS. Radiological evaluations: ML, MK, CL, JM, MW, SO, and MK. Acquisition, analysis, or interpretation of data: All authors. Drafting the manuscript: ML, MK, CL, and MMK. Critical revision of the manuscript for important intellectual content and approved the version to be published: All authors. Statistical analysis: ML, MK, and MMK. Obtained funding: MMK, NN, ML, and TC. Administrative, technical, or material support: The authors would like to thank the TTM2-trial investigators and the sponsors for their support, and Michael Borring from Skåne University Hospital, Malmö for help with data collection and storage. Supervision: CL, NN, JW, MMK, TC, and SU. Agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: All authors.

## Funding

Open access funding provided by Lund University. The TTM2-trial is supported by independent research grants from nonprofit or governmental agencies (the Swedish Research Council [Vetenskapsrådet]), Swedish Heart-Lung Foundation, Stig and Ragna Gorthon Foundation, Knutsson Foundation, Laerdal Foundation, Hans-Gabriel and Alice Trolle-Wachtmeister Foundation for Medical Research, and Regional Research Support in Region Skåne and Region Halland) and by governmental funding of clinical research within the Swedish National Health Service (ALF). In addition, the CT substudy is supported by grants from the Bundy Academy at Lund University, the Segerfalk Foundation, and the Elsa Schmitz Foundation. MK is supported by the Laerdal

Foundation and the Berlin Institute of Health Junior Digital Clinician Scientist Program. The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the manuscript, and its final contents.

#### Data availability

After publication, data may be shared upon reasonable request.

#### Declarations

#### Conflicts of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest related to the current manuscript.

#### Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 13 March 2024 Accepted: 21 May 2024

Published: 20 June 2024

#### References

- Nolan JP, Sandroni C, Bottiger BW, Cariou A, Cronberg T, Friberg H, Genbrugge C, Haywood K, Lilja G, Moulart VRM, Nikolaou N, Olsveengen TM, Skrifvars MB, Taccone F, Soar J (2021) European Resuscitation Council and European Society of Intensive Care Medicine guidelines 2021: post-resuscitation care. *Intensive Care Med* 47(4):369–421. <https://doi.org/10.1007/s00134-021-06368-4>
- Elmer J, Steinberg A, Callaway CW (2023) Paucity of neuroprognostic testing after cardiac arrest in the United States. *Resuscitation* 188:109762. <https://doi.org/10.1016/j.resuscitation.2023.109762>
- Jarpestam S, Martinell L, Rylander C, Lilja L (2023) Post-cardiac arrest intensive care in Sweden: a survey of current clinical practice. *Acta Anaesthesiol Scand* 67(9):1249–1255. <https://doi.org/10.1111/aas.14298>
- Zhou SE, Barden MM, Gilmore EJ, Pontes-Neto OM, Sampaio Silva G, Kurtz P, Oliveira-Filho J, Cougo-Pinto PT, Zampieri FG, Napoli NJ, Theriot JJ, Greer DM, Maciel CB (2021) Post-cardiac arrest neuroprognostication practices: a survey of Brazilian Physicians. *Crit Care Explor* 3(1):e0321. <https://doi.org/10.1097/CCE.0000000000000321>
- Rajajee V, Muehlschlegel S, Wartenberg KE, Alexander SA, Busl KM, Chou SHY, Creutzfeldt CJ, Fontaine GV, Fried H, Hocker SE, Hwang DY, Kim KS, Madzar D, Mahanes D, Mainali S, Meixensberger J, Montellano F, Sakowitz OW, Weimar C, Westermaier T, Varelas PN (2023) Guidelines for neuroprognostication in comatose adult survivors of cardiac arrest. *Neurocrit Care* 38(3):533–563. <https://doi.org/10.1007/s12028-023-01688-3>
- Keijzer HM, Hoedemaekers CWE, Meijer FJA, Tonino BAR, Klijn CJM, Hofmeijer J (2018) Brain imaging in comatose survivors of cardiac arrest: pathophysiological correlates and prognostic properties. *Resuscitation* 133:124–136. <https://doi.org/10.1016/j.resuscitation.2018.09.012>
- Sandroni C, D'Arrigo S, Cacciola S, Hoedemaekers CWE, Kamps MJA, Oddo M, Taccone FS, Di Rocco A, Meijer FJA, Westhall E, Antonelli M, Soar J, Nolan JP, Cronberg T (2020) Prediction of poor neurological outcome in comatose survivors of cardiac arrest: a systematic review. *Intensive Care Med* 46(10):1803–1851. <https://doi.org/10.1007/s00134-020-06198-w>
- Lagebrant A, Lang M, Nielsen N, Blennow K, Dankiewicz J, Friberg H, Hassager C, Horn J, Kjaergaard J, Kuiper MA, Mattsson-Carlgrén N, Pellis T, Rylander C, Sigmund R, Stammen P, Uden J, Zetterberg H, Wise MP, Cronberg T, Moseby-Knappe M (2023) Brain injury markers in blood predict signs of hypoxic ischaemic encephalopathy on head computed tomography after cardiac arrest. *Resuscitation* 184:109668. <https://doi.org/10.1016/j.resuscitation.2022.12.006>
- Endisch C, Westhall E, Kenda M, Streitberger KJ, Kirkegaard H, Stenzel W, Storm C, Ploner CJ, Cronberg T, Friberg H, Englund E, Leithner C (2020) Hypoxic-ischemic encephalopathy evaluated by brain autopsy and neuroprognostication after cardiac arrest. *JAMA Neurol.* <https://doi.org/10.1001/jamaneurol.2020.2340>
- Beekman R, Hirsch KG (2023) Brain imaging after cardiac arrest. *Curr Opin Crit Care* 29(3):192–198. <https://doi.org/10.1097/MCC.0000000000001032>
- Caraganis A, Mulder M, Kempainen RR, Brown RZ, Oswood M, Hoffman B, Prekker ME (2020) Interobserver variability in the recognition of hypoxic-ischemic brain injury on computed tomography soon after out-of-hospital cardiac arrest. *Neurocrit Care* 33(2):414–421. <https://doi.org/10.1007/s12028-019-00900-7>
- Beekman R, Maciel CB, Ormseth CH, Zhou SE, Galluzzo D, Miyares LC, Torres-Lopez VM, Payabvash S, Mak A, Greer DM, Gilmore EJ (2021) Early head CT in post-cardiac arrest patients: a helpful tool or contributor to self-fulfilling prophecy? *Resuscitation* 165:68–76. <https://doi.org/10.1016/j.resuscitation.2021.06.004>
- Lang M, Leithner C, Scheel M, Kenda M, Cronberg T, During J, Rylander C, Annborn M, Dankiewicz J, Deye N, Halliday T, Lascarrou JB, Matthew T, McGuigan P, Morgan M, Thomas M, Ullen S, Uden J, Nielsen N, Moseby-Knappe M (2022) Prognostic accuracy of head computed tomography for prediction of functional outcome after out-of-hospital cardiac arrest: rationale and design of the prospective TTM2-CT-substudy. *Resusc Plus* 12:100316. <https://doi.org/10.1016/j.resplu.2022.100316>
- Kenda M, Cheng Z, Guettler C, Storm C, Ploner CJ, Leithner C, Scheel M (2022) Inter-rater agreement between humans and computer in quantitative assessment of computed tomography after cardiac arrest. *Front Neurol* 13:990208. <https://doi.org/10.3389/fneur.2022.990208>
- Vagal A, Wintermark M, Nael K, Bivard A, Parsons M, Grossman AW, Khatri P (2019) Automated CT perfusion imaging for acute ischemic stroke: pearls and pitfalls for real-world use. *Neurology* 93(20):888–898. <https://doi.org/10.1212/WNL.00000000000008481>
- Hanning U, Bernhard Sporns P, Lebedez P, Niederstadt T, Zoubi T, Schmidt R, Knecht S, Heindel W, Kemmling A (2016) Automated assessment of early hypoxic brain edema in non-enhanced CT predicts outcome in patients after cardiac arrest. *Resuscitation* 104:91–94. <https://doi.org/10.1016/j.resuscitation.2016.03.018>
- Kenda M, Scheel M, Kemmling A, Aalberts N, Guettler C, Streitberger KJ, Storm C, Ploner CJ, Leithner C (2021) Automated assessment of brain CT after cardiac arrest-an observational derivation/validation cohort study. *Crit Care Med* 49(12):e1212–e1222. <https://doi.org/10.1097/CCM.00000000000005198>
- Lang M, Nielsen N, Ullen S, Abul-Kasim K, Johnsson M, Helbok R, Leithner C, Cronberg T, Moseby-Knappe M (2022) A pilot study of methods for prediction of poor outcome by head computed tomography after cardiac arrest. *Resuscitation* 179:61–70. <https://doi.org/10.1016/j.resuscitation.2022.07.035>
- Scheel M, Storm C, Gentsch A, Nee J, Luckenbach F, Ploner CJ, Leithner C (2013) The prognostic value of gray-white-matter ratio in cardiac arrest patients treated with hypothermia. *Scand J Trauma Resusc Emerg Med* 21:23. <https://doi.org/10.1186/1757-7241-21-23>
- Moseby-Knappe M, Pellis T, Dragancea I, Friberg H, Nielsen N, Horn J, Kuiper M, Roncarati A, Siemund R, Uden J, Cronberg T (2017) Head computed tomography for prognostication of poor outcome in comatose patients after cardiac arrest and targeted temperature management. *Resuscitation* 119:89–94. <https://doi.org/10.1016/j.resuscitation.2017.06.027>
- Hinson HE (2015) Blurred lines: reduced discrimination of gray white junctions observed on CT in out-of-hospital cardiac arrest (OHCA) portends poor prognosis. *Resuscitation* 92:A1. <https://doi.org/10.1016/j.resuscitation.2015.04.029>
- Torbey MT, Selim M, Knorr J, Bigelow C, Recht L (2000) Quantitative analysis of the loss of distinction between gray and white matter in comatose patients after cardiac arrest. *Stroke* 31(9):2163–2167. <https://doi.org/10.1161/01.str.31.9.2163>



23. Kirsch K, Heymel S, Günther A, Vahl K, Schmidt T, Michalski D, Fritzenwanger M, Schulze PC, Pfeifer R (2021) Prognostication of neurologic outcome using gray-white-matter-ratio in comatose patients after cardiac arrest. *BMC Neurol* 21(1):456. <https://doi.org/10.1186/s12883-021-02480-6>
24. Streitberger KJ, Endisch C, Ploner CJ, Stevens R, Scheel M, Kenda M, Storm C, Leithner C (2019) Timing of brain computed tomography and accuracy of outcome prediction after cardiac arrest. *Resuscitation* 145:8–14. <https://doi.org/10.1016/j.resuscitation.2019.09.025>
25. Dankiewicz J, Cronberg T, Lilja G, Jakobsen JC, Levin H, Ullen S, Rylander C, Wise MP, Oddo M, Cariou A, Belohlavek J, Hovdenes J, Saxena M, Kirkegaard H, Young PJ, Pelosi P, Storm C, Taccone FS, Joannidis M, Callaway C, Eastwood GM, Morgan MPG, Nordberg P, Erlinge D, Nichol AD, Chew MS, Hollenberg J, Thomas M, Bewley J, Sweet K, Grejs AM, Christensen S, Haenggi M, Levis A, Lundin A, During J, Schmidbauer S, Keeble TR, Karamasis GV, Schrag C, Faessler E, Smid O, Otahal M, Maggiorini M, Wendel Garcia PD, Jaubert P, Cole JM, Solar M, Borgquist O, Leithner C, Abed-Maillard S, Navarra L, Annborn M, Unden J, Brunetti I, Awad A, McGuigan P, Bjorkholt Olsen R, Cassina T, Vignon P, Langeland H, Lange T, Friberg H, Nielsen N, Investigators TTM2 (2021) Hypothermia versus normothermia after out-of-hospital cardiac arrest. *N Engl J Med* 384(24):2283–2294. <https://doi.org/10.1056/NEJMoa2100591>
26. World Medical A (2013) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 310(20):2191–2194. <https://doi.org/10.1001/jama.2013.281053>
27. Lilja G, Ullen S, Dankiewicz J, Friberg H, Levin H, Nordstrom EB, Heimburg K, Jakobsen JC, Ahlqvist M, Bass F, Belohlavek J, Olsen RB, Cariou A, Eastwood G, Fanebust HR, Grejs AM, Grimmer L, Hammond NE, Hovdenes J, Hrecko J, Iten M, Johansen H, Keeble TR, Kirkegaard H, Lascarrou JB, Leithner C, Lesona ME, Levis A, Mion M, Moseby-Knappe M, Navarra L, Nordberg P, Pelosi P, Quayle R, Rylander C, Sandberg H, Saxena M, Schrag C, Siranec M, Tiziano C, Vignon P, Wendel-Garcia PD, Wise MP, Wright K, Nielsen N, Cronberg T (2023) Effects of hypothermia vs normothermia on societal participation and cognitive function at 6 months in survivors after out-of-hospital cardiac arrest: a predefined analysis of the TTM2 randomized clinical trial. *JAMA Neurol* 80(10):1070–1079. <https://doi.org/10.1001/jamaneurol.2023.2536>
28. Dankiewicz J, Cronberg T, Lilja G, Jakobsen JC, Belohlavek J, Callaway C, Cariou A, Eastwood G, Erlinge D, Hovdenes J, Joannidis M, Kirkegaard H, Kuiper M, Levin H, Morgan MPG, Nichol AD, Nordberg P, Oddo M, Pelosi P, Rylander C, Saxena M, Storm C, Taccone F, Ullen S, Wise MP, Young P, Friberg H, Nielsen N (2019) Targeted hypothermia versus targeted normothermia after out-of-hospital cardiac arrest (TTM2): a normalisation clinical trial-rationale and design. *Am Heart J* 217:23–31. <https://doi.org/10.1016/j.ahj.2019.06.012>
29. Turella S, Dankiewicz J, Friberg H, Jakobsen JC, Leithner C, Levin H, Lilja G, Moseby-Knappe M, Nielsen N, Rossetti AO, Sandroni C, Zubler F, Cronberg T, Westhall E (2024) The predictive value of highly malignant EEG patterns after cardiac arrest: evaluation of the ERC-ESICM recommendations. *Intensive Care Med* 50(1):90–102. <https://doi.org/10.1007/s00134-023-07280-9>
30. Genske U, Jahnke P (2022) Human observer net: a platform tool for human observer studies of image data. *Radiology* 303(3):524–530. <https://doi.org/10.1148/radiol.211832>
31. Rorden C, Bonilha L, Fridriksson J, Bender B, Karnath HO (2012) Age-specific CT and MRI templates for spatial normalization. *Neuroimage* 61(4):957–965. <https://doi.org/10.1016/j.neuroimage.2012.03.020>
32. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) Fsl. *Neuroimage* 62(2):782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
33. Lilja G, Nielsen N, Ullen S, Blennow Nordstrom E, Dankiewicz J, Friberg H, Heimburg K, Jakobsen JC, Levin H, Callaway C, Cariou A, Eastwood GM, Helbok R, Hovdenes J, Kirkegaard H, Leithner C, Morgan MPG, Nordberg P, Oddo M, Pelosi P, Rylander C, Saxena M, Taccone FS, Siranec M, Wise MP, Young PJ, Cronberg T (2020) Protocol for outcome reporting and follow-up in the targeted hypothermia versus targeted normothermia after out-of-hospital cardiac arrest trial (TTM2). *Resuscitation* 150:104–112. <https://doi.org/10.1016/j.resuscitation.2020.03.004>
34. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HC, Bossuyt PM (2016) STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 6(11):e012799. <https://doi.org/10.1136/bmjopen-2016-012799>
35. Geocadin RG, Callaway CW, Fink EL, Golan E, Greer DM, Ko NU, Lang E, Licht DJ, Marino BS, McNair ND, Peberdy MA, Perman SM, Sims DB, Soar J, Sandroni C (2019) Standards for studies of neurological prognostication in comatose survivors of cardiac arrest: a scientific statement from the American Heart Association. *Circulation* 140(9):e517–e542. <https://doi.org/10.1161/cir.0000000000000702>
36. Nahm FS (2022) Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol* 75(1):25–36
37. Fleiss JL (1971) Measuring nominal scale agreement among many raters. *Psychol Bull* 76:378–382
38. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
39. Hartmann DP (1977) Considerations in the choice of interobserver reliability estimates. *J Appl Behav Anal* 10(1):103–116. <https://doi.org/10.1901/jaba.1977.10-103>
40. Callaway CW, Donnino MW, Fink EL, Geocadin RG, Golan E, Kern KB, Leary M, Meurer WJ, Peberdy MA, Thompson TM, Zimmerman JL (2015) Part 8: post-cardiac arrest care: 2015 American Heart Association Guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation* 132(18 Suppl 2):S465–482. <https://doi.org/10.1161/CIR.0000000000000262>
41. Youn CS, Park KN, Kim SH, Lee BK, Cronberg T, Oh SH, Jeung KW, Cho IS, Choi SP (2022) External validation of the 2020 ERC/ESICM prognostication strategy algorithm after cardiac arrest. *Crit Care* 26(1):95. <https://doi.org/10.1186/s13054-022-03954-w>
42. Lopez Soto C, Dragoi L, Heyn CC, Kramer A, Pinto R, Adhikari NKJ, Scales DC (2020) Imaging for neuroprognostication after cardiac arrest: systematic review and meta-analysis. *Neurocrit Care* 32(1):206–216. <https://doi.org/10.1007/s12028-019-00842-0>
43. Lee BK, Jeung KW, Song KH, Jung YH, Choi WJ, Kim SH, Youn CS, Cho IS, Lee DH (2015) 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* 96:46–52. <https://doi.org/10.1016/j.resuscitation.2015.07.027>
44. Lee BK, Kim WY, Shin J, Oh JS, Wee JH, Cha KC, Park Y, Choi JH, Jeung KW (2016) Prognostic value of gray matter to white matter ratio in hypoxic and non-hypoxic cardiac arrest with non-cardiac etiology. *Am J Emerg Med* 34(8):1583–1588. <https://doi.org/10.1016/j.ajem.2016.05.063>
45. Metter RB, Rittenberger JC, Guyette FX, Callaway CW (2011) Association between a quantitative CT scan measure of brain edema and outcome after cardiac arrest. *Resuscitation* 82(9):1180–1185. <https://doi.org/10.1016/j.resuscitation.2011.04.001>
46. Wang GN, Chen XF, Lv JR, Sun NN, Xu XQ, Zhang JS (2018) The prognostic value of gray-white matter ratio on brain computed tomography in adult comatose cardiac arrest survivors. *J Chin Med Assoc* 81(7):599–604. <https://doi.org/10.1016/j.jcma.2018.03.003>
47. Son SH, Lee IH, Park JS, Yoo IS, Kim SW, Lee JW, Ryu S, You Y, Min JH, Cho YC, Jeong WJ, Oh SK, Cho SU, Ahn HJ, Kang C, Lee DH, Lee BK, Youn CS (2020) Does combining biomarkers and brain images provide improved prognostic predictive performance for out-of-hospital cardiac arrest survivors before target temperature management? *J Clin Med*. <https://doi.org/10.3390/jcm9030744>
48. Molinski NS, Kenda M, Leithner C, Nee J, Storm C, Scheel M, Meddeb A (2024) Deep learning-enabled detection of hypoxic-ischemic encephalopathy after cardiac arrest in CT scans: a comparative study of 2D and 3D approaches. *Front Neurosci* 18:1245791. <https://doi.org/10.3389/fnins.2024.1245791>
49. In YN, Lee IH, Park JS, Kim DM, You Y, Min JH, Jeong W, Ahn HJ, Kang C, Lee BK (2022) Delayed head CT in out-of-hospital cardiac arrest survivors: does this improve predictive performance of neurological outcome? *Resuscitation* 172:1–8. <https://doi.org/10.1016/j.resuscitation.2022.01.003>