Title

Correlating Quantitative MRI-based Apparent Diffusion Coefficient Metrics with 24-month Neurodevelopmental Outcomes in Neonates from the HEAL Trial.

Journal

Radiology, 308(3)

Authors

Calabrese, Evan
Wu, Yvonne
Scheffler, Aaron
et al.

Publication Date

2023-09-01

DOI

10.1148/radiol.223262

Peer reviewed

Radiology

ORIGINAL RESEARCH • PEDIATRIC IMAGING

Correlating Quantitative MRI-based Apparent Diffusion Coefficient Metrics with 24-month Neurodevelopmental Outcomes in Neonates from the HEAL Trial

Evan Calabrese, MD, PhD • Yvonne Wu, MD, MPH • Aaron Wolfe Scheffler, PhD • Jessica L. Wisnowski, PhD • Robert C. McKinstry, MD, PhD • Amit Mathur, MBBS • Hannah C. Glass, MD, MAS • Bryan A. Comstock, PhD • Patrick J. Heagerty, PhD • Shivani Gillon, MD • Sandra E. Juul, MD, PhD • Christopher P. Hess, MD, PhD • Yi Li, MD

From the Department of Radiology, Duke University Medical Center, Durham, NC (E.C.); Department of Neurology and Weill Institute for Neuroscience (Y.W., H.C.G.), Department of Pediatrics, UCSF Benioff Children's Hospital (Y.W., H.C.G.), Department of Epidemiology and Biostatistics (A.W.S.), School of Medicine (S.G.), and Neuroradiology Section, Department of Radiology and Biomedical Imaging (C.P.H., Y.L.), University of California, San Francisco, 505 Parnassus Ave, M-391, San Francisco, CA 94143-0628; Department of Radiology, Children's Hospital of Los Angeles, University of Southern California Keck School of Medicine, Los Angeles, Calif (J.L.W.); Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, Mo (R.C.M.); Department of Pediatrics, St Louis University, St Louis, Mo (A.M.); and Departments of Statistics (B.A.C., P.J.H.) and Pediatrics (S.E.J.), University of Washington, Seattle, Wash. Received January 2, 2023; revision requested March 14; revision received July 10; accepted July 18. **Address correspondence to** Y.L. (email: *yi.li@ucsf.edu*).

Y.L. supported by the 2022 American Society of Pediatric Neuroradiology Guerbet Grant.

Conflicts of interest are listed at the end of this article.

See also the editorial by Huisman in this issue.

Radiology 2023; 308(3):e223262 • https://doi.org/10.1148/radiol.223262 • Content codes: NR MR PD

Background: Multiple qualitative scoring systems have been created to capture the imaging severity of hypoxic ischemic brain injury.

Purpose: To evaluate quantitative volumes of acute brain injury at MRI in neonates with hypoxic ischemic brain injury and correlate these findings with 24-month neurodevelopmental outcomes and qualitative brain injury scoring by radiologists.

Materials and Methods: In this secondary analysis, brain diffusion-weighted MRI data from neonates in the High-dose Erythropoietin for Asphyxia and Encephalopathy trial, which recruited participants between January 2017 and October 2019, were analyzed. Volume of acute brain injury, defined as brain with apparent diffusion coefficient (ADC) less than 800×10^{-6} mm^2/sec, was automatically computed across the whole brain and within the thalami and white matter. Outcomes of death and neurodevelopmental impairment (NDI) were recorded at 24-month follow-up. Associations between the presence and volume (in milliliters) of acute brain injury with 24-month outcomes were evaluated using multiple logistic regression. The correlation between quantitative acute brain injury volume and qualitative MRI scores was assessed using the Kendall tau-b test.

Results: A total of 416 neonates had available MRI data (mean gestational age, 39.1 weeks ± 1.4 [SD]; 235 male) and 113 (27%) showed evidence of acute brain injury at MRI. Of the 387 participants with 24-month follow-up data, 185 (48%) died or had any NDI. Volume of acute injury greater than 1 mL (odds ratio [OR], 13.9 [95% CI: 5.93, 32.45]; $P < .001$) and presence of any acute injury in the brain (OR, 4.5 [95% CI: 2.6, 7.8]; $P < .001$) were associated with increased odds of death or any NDI. Quantitative whole-brain acute injury volume was strongly associated with radiologists' qualitative scoring of diffusion-weighted images (Kendall tau-b = 0.56; $P < .001$).

Conclusion: Automated quantitative volume of brain injury is associated with death, moderate to severe NDI, and cerebral palsy in neonates with hypoxic ischemic encephalopathy and correlated well with qualitative MRI scoring of acute brain injury.

Clinical trial registration no. NCT02811263

Supplemental material is available for this article.

Hypoxic ischemic encephalopathy (HIE) affects approximately 1.5 of every 1000 live births in high-resource countries (1) and is responsible for nearly a quarter of neonatal deaths worldwide annually (2). Among survivors, HIE is a major cause of chronic neurologic disability, with long-term sequelae including developmental delay and cerebral palsy (3,4). Neonates with HIE are identified at birth by a clinical examination characterized by abnormalities in categories such consciousness, activity, tone, posture, primitive reflexes, and autonomic nervous system (5); laboratory abnormalities such as metabolic acidosis; and clinical factors predisposing them to perinatal asphyxia. Therapeutic hypothermia is the standard of care for neonates with moderate to severe HIE (3), leading to improved survival and outcomes (6,7). However, approximately 35% of neonates who receive therapeutic hypothermia will develop moderate to severe neurodevelopmental impairment (NDI) (4).

MRI plays an important role in the evaluation of neonates with HIE. Several qualitative MRI scoring systems (8–14) have been created to classify patterns and severity of hypoxic ischemic injury for purposes of outcome characterization and prognostication and as secondary end points of clinical trials. These scoring systems are subject to interrater variability (13,15) and often require time-consuming consensus scoring to achieve agreement among experts (16).

The High-dose Erythropoietin for Asphyxia and Encephalopathy (HEAL) trial was a multicenter, phase 3, randomized, double-blinded, placebo-controlled trial conducted

Abbreviations

ADC = apparent diffusion coefficient, DTI = diffusion tensor imaging, DWI = diffusion-weighted image, GMFCS = Gross Motor Function Classification System, HEAL = High-dose Erythropoietin for Asphyxia and Encephalopathy, HIE = hypoxic ischemic encephalopathy, NDI = neurodevelopmental impairment, OR = odds ratio

Summary

Quantitative diffusion MRI–based metrics of acute brain injury in neonates with hypoxic ischemic encephalopathy were associated with death, moderate to severe neurodevelopmental impairment, and cerebral palsy and correlated strongly with qualitative MRI scoring of brain injury by radiologists.

Key Results

- In this secondary analysis of the High-dose Erythropoietin for Asphyxia and Encephalopathy trial, quantitative analysis of diffusion-weighted MRI data from 416 encephalopathic neonates identified acute brain injury in 113 of 416 participants (27%).
- Global volume of acute brain injury greater than 1 mL was strongly associated with death (odds ratio [OR], 42.2; $P < .001$), moderate to severe neurodevelopmental impairment (OR, 2.00; $P = .02$), and cerebral palsy (OR, 2.74; $P = .006$) at 24 months.
- Automatically derived quantitative diffusion metrics showed a strong positive association with radiologists' scoring of brain injury at neonatal MRI (Kendall tau-b = 0.56; $P < .001$).

between January 25, 2017, and October 9, 2019, that evaluated the efficacy of erythropoietin as a neuroprotective adjunct to therapeutic hypothermia in the treatment of neonates with moderate to severe encephalopathy. The trial ultimately concluded that erythropoietin was not associated with a decrease in risk of death or NDI (17). In this secondary analysis of the multicenter cohort of neonates enrolled in the HEAL trial, we developed an automated quantitative approach to classifying acute injury based on apparent diffusion coefficient (ADC) analysis. We aimed to evaluate quantitative volumes of acute brain injury at MRI in neonates with HIE and correlate these findings with 24-month neurodevelopmental outcomes and qualitative brain injury scoring by radiologists.

Materials and Methods

Study Design and Participants

This secondary analysis of the HEAL trial data was supported by the American Society of Pediatric Neuroradiology Guerbet Grant (awarded to Y.L.) funded by the Guerbet Corporation. The authors had sole control of the data submitted for publication, independent of the grant funding agency.

The HEAL trial (ClinicalTrials.gov identifier NCT02811263) enrolled 500 consecutive neonates with HIE at 17 sites after written informed parental consent (17–20) (Fig 1). The trial was approved by the institutional review boards of all participating sites (Table S1). All participants were born at 36 weeks gestational age or older, had moderate to severe HIE, and received therapeutic hypothermia. Full detailed inclusion and exclusion criteria for the trial are provided in Table S2. Clinical information was acquired, including maternal race or ethnicity, neonatal sex, gestational age at birth, Apgar scores at 5 and 10 minutes, and severity of HIE (moderate vs severe) based on a modified Sarnat examination. In the parent HEAL trial, treatment with erythropoietin did not affect 2-year neurodevelopmental outcome, so all infants were included in this analysis regardless of the treatment arm (17). This current study differs from previously published reports from the HEAL trial in that it describes and validates a quantitative diffusion MRI analysis pipeline and its correlation with 2-year neurodevelopmental outcome. Previously published articles from the HEAL cohort have used only expert qualitative scoring of injury.

MRI Scan Acquisition

Neonates underwent 3.0-T brain MRI at a median age of 119.5 hours (IQR, 108.2–139.1 hours), including T1-weighted, T2-weighted, and 30-direction diffusion tensor imaging (DTI) sequences, according to a protocol harmonized across 17 study sites and nine MRI platforms, as previously published (16). Diffusion MRI data were acquired from DTI in the axial plane at b of 1000 sec/mm^2, with specific parameters per platform (Table S3). The use of sedation versus feed-and-swaddle technique was not recorded. All imaging data were manually reviewed for quality by a neuroradiologist (Y.L.) with 5 years of attending experience.

Diffusion MRI Processing and Measurements

A customized automated DTI processing pipeline was implemented in Python 3.8 (Python Software Foundation) using Neuroimaging in Python Pipelines and Interfaces (Nipype, version 1.6 [21]) to interface with non-Python software. DTI volumes were converted to Neuroimaging Informatics Technology Initiative, or NIfTI, format with use of dcm2niix 1.0. Subsequently, DTI NIfTI volumes were eddy current–corrected and processed using the Eddy and DTIFIT modules from FMRIB Software Library (version 6.0.2, FMRIB) (22), which yielded several parametric maps, including the ADC. All DTI-derived data were automatically aligned to the University of North Carolina neonatal brain atlas (23) (23 regions per hemisphere, 46 regions total) by using a multistep, multiresolution diffeomorphic registration approach implemented with Advanced Normalization Tools 2.3.5 (24). DTI processing pipeline steps were automated, and computation time was less than 10 minutes per scan with use of a dedicated desktop workstation.

ADC Threshold Volume Calculation

Acute brain injury was defined as all voxels with ADC less than 800×10^{-6} mm^2/sec and corresponding diffusion MRI hyperintensity (defined as diffusion-weighted trace [b = 1000 sec/mm^2] signal intensity above the mean intensity of the whole brain parenchyma). This threshold was chosen based on prior research investigating average mean diffusivity of acute injury in gray matter in the setting of therapeutic hypothermia (25–27). Several processing steps were undertaken to avoid mislabeling voxels as acute injury. First, the threshold for ADC images was set at ADC less than 800×10^{-6} mm^2/sec, yielding a binary image. Next, all voxels with a corresponding diffusion-weighted image (DWI) intensity below the whole-brain mean (corresponding to signal contribution from regions of T2 hypointensity) were

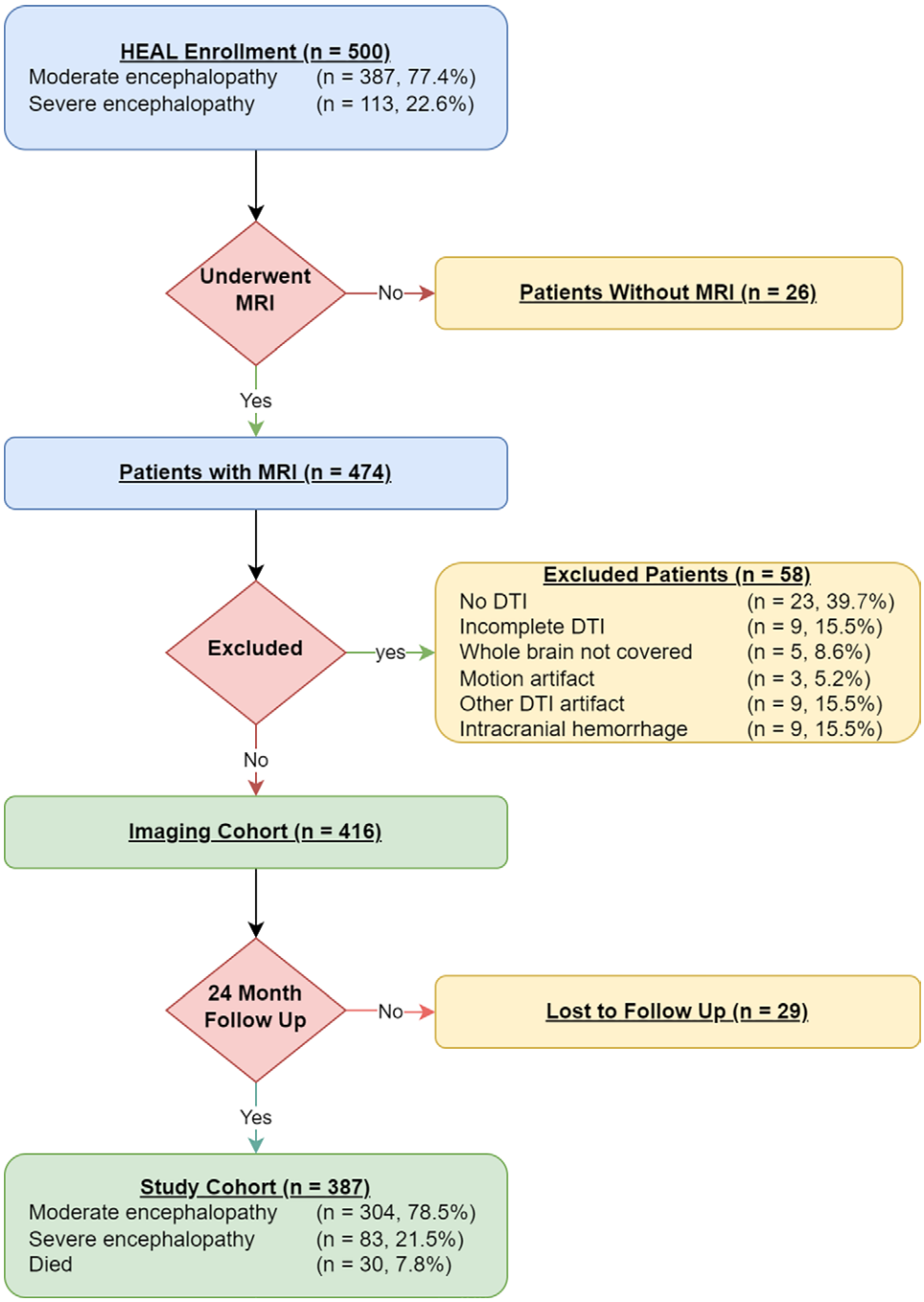

Figure 1: Flowchart of participant exclusion criteria from enrollment to 24-month follow-up. DTI = diffusion tensor imaging, HEAL = High-dose Erythropoietin for Asphyxia and Encephalopathy.

removed. Next, binary injury volumes were smoothed using morphologic erosion and dilation with a 3 × 3 × 3-voxel spherical kernel. Finally, all 26 connected components of fewer than nine voxels (corresponding to potential areas of random noise) were removed. Quantitative ADC analysis code was adapted from a prior study focused on adult patients with cardiac arrest (28) and is provided at *https://github.com/ecalabr/neonatal_anox* (commit ID: 0abebff).

Volume of injury in milliliters was assessed across the whole brain and within the thalami and white matter. These regional areas of interest were chosen a priori based on the high frequency of thalamic injury in the basal ganglia/thalamic pattern of injury and the high frequency of white matter injury in the watershed pattern of injury (29–31). Volumes of acute injury were calculated for all participants with adequate imaging, including those with no detectable (0 mL) acute injury.

Qualitative MRI Scoring

All MRI scans were previously scored for the presence and extent of injury with use of a validated scoring system by readers blinded to clinical information (9). The scoring system incorporates separate scores for abnormality on each of three

sequences (diffusion-, T1-, and T2-weighted imaging) and captures regional scores for five subregions per hemisphere, including white matter and thalami. For each sequence and subregion, possible injury scores include 0 (no signal abnormality), 1 (signal abnormality in <25% of the region), 2 (signal abnormality in 25%–50% of the region), or 3 (widespread injury involving >50% of the region). The total possible score was 18 for thalamic and for white matter subregions. The total possible global injury score was 138. All studies were initially independently scored by two of three readers (J.L.W., with 19 years of experience in neonatal MRI research scoring; A.M., with 15 years of experience in neonatal MRI research scoring; and R.C.M., with 24 years of experience in neuroradiology), who were assigned studies randomly, and all final scores were agreed on by consensus. The global and regional DWI scores were considered the reference standard for comparison with quantitative ADC measurements.

Neurodevelopmental Outcome Assessment

Participants were evaluated at 24 months of age with use of the Bayley Scales of Infant Development III, a standardized neurologic examination (32), and the Gross Motor Function Classification System (GMFCS) (18). NDI was defined as any of the following: GMFCS level 1 or higher, GMFCS level 0 or 0.5 and cerebral palsy at neurologic examination, or Bayley III cognitive score less than 90 (0.67 SD below the mean). As in the parent HEAL trial, the primary outcome was death or any NDI at 2 years of age (22–36 months). Secondary outcomes were *(a)* death; *(b)* cerebral palsy, as determined by validated neurologic examination; and *(c)* moderate to severe NDI, defined as GMFCS level 1 or higher and cerebral palsy, GMFCS level 2 or higher, quadriplegic cerebral palsy, or a Bayley III cognitive score less than 85 (1 SD below the mean).

Development of Quantitative Diffusion Metrics

We evaluated the relationship between the primary and secondary end points and several quantitative diffusion metrics on the basis of previously described patterns of injury: volume of acute injury in the whole brain, thalami, and white matter, as well as binary presence of acute injury (greater than 0 mL) in the whole brain, thalami, and white matter. Due to the skewed distribution of injury volume (Fig S1), descriptive statistics of volume of acute injury report the 75th and 90th percentiles instead of IQR. Volume of acute injury across the whole brain was further evaluated using receiver operating characteristic analysis, and a proposed threshold for outcome prediction was defined as the point on the receiver operating characteristic curve that was easily recognized and yielded a minimum euclidean distance with the upper left corner of the receiver operating characteristic graph (Fig S2). Based on this analysis, we adopted presence of acute injury greater than 1 mL in the whole brain as an additional binary quantitative metric for evaluation.

Statistical Analysis

Statistical analysis was performed using Stata, version 16 (StataCorp) by two authors (Y.L. and A.W.S., with 5 and 10 years of experience in statistical analysis, respectively). Clinical characteristics and administration of erythropoietin were compared between participants with and without acute injury with use of Wilcoxon rank-sum and χ^2 tests. Quantitative volume of acute injury in the whole brain as well as in thalami and white matter was correlated with the corresponding qualitative MRI scores by using the Kendall tau-b test (33). For Kendall tau-b, absolute values less than 0.2 correspond to weak associations, values less than 0.3 correspond to moderate associations, and values of 0.3 and above correspond to strong associations (34). The proportion of participants who died or had any NDI or moderate to severe NDI were compared between those with and without acute brain injury with use of χ^2 tests. ADC-derived quantitative imaging metrics of the entire sample and stratified by those without NDI and those who died or had any NDI were evaluated using χ^2 and Wilcoxon rank-sum tests. A multiple logistic regression was used to model the association of quantitative ADC-based metrics and DWI scores with the primary and secondary outcomes controlling for gestational age, sex, and the administration of erythropoietin. The site of enrollment was included in the model as a site-specific random intercept, and estimation was carried out by means of maximum likelihood methods. In instances where maximum likelihood methods experienced quasi-separation leading to unstable estimates of associations for predictors of primary interest, exact multiple logistic regression was used to produce less biased point estimates and CIs (35). Statistically significant difference was defined as $P < .05$ using two-sided tests.

Results

Participant Characteristics

Of the 500 neonates initially enrolled, 474 (95%) underwent brain MRI. A total of 58 neonates were excluded because DTI was either not available ($n = 23$), substantially degraded by artifact ($n = 12$) (Fig S3), incomplete ($n = 9$), or lacked whole-brain coverage ($n = 5$) or because of substantial intracranial hemorrhage causing mass effect ($n = 9$) (Fig S4). The remaining 416 neonates constituted the imaging cohort (Fig 1).

A total of 416 neonates were included in the imaging cohort (mean gestational age, 39.1 weeks ± 1.4 [SD]; 235 male, 181 female), and the mean postnatal age at MRI was 119.5 hours (IQR, 108.2–139.1 hours). Evidence of acute brain injury was observed at MRI in 113 of 416 participants (27%) (Table 1). Within the imaging cohort, 332 of 416 participants (80%) had moderate HIE at clinical examination and 84 (20%) had severe HIE. A total of 216 of 416 participants (52%) were randomized to receive erythropoietin.

Outcome data at 24 months postnatal age were available for 387 of 416 participants (93%) and constituted the study cohort (Fig 1). Of participants in the study cohort, 185 (48%) experienced the primary outcome of death or any NDI (Table 2). Among those with the primary outcome, 30 of 387 participants (7.8%) died and 155 of 387 (40%) had any NDI. Among those with any NDI, 111 of 387 (29%)

Table 1: Clinical Characteristics of Participants Stratified by Presence of Acute Brain Injury at Neonatal MRI

Clinical Characteristic	All Participants (n = 416)	Participants without Acute Injury (n = 303)	Participants with Acute Injury (n = 113)	P Value
Gestational age (wk)*	39.1 ± 1.4	39.1 ± 1.4	39.2 ± 1.4	.32
Sex				.97
M	235 (56)	171 (56)	64 (57)	
F	181 (44)	132 (44)	49 (43)	
Birthweight (g)*	3378 ± 594	3381 ± 578	3372 ± 640	.45
Apgar score at 5 minutes†	3 (2–5)	4 (2–5)	3 (1–4)	<.001
Apgar score at 10 minutes†	5 (4–7)	5 (4–7)	4 (2–5.5)	<.001
Encephalopathy				<.001
Moderate	332 (80)	262 (86)	70 (62)	
Severe	84 (20)	41 (14)	43 (38)	
Received erythropoietin	216 (52)	155 (51)	61 (54)	.61
Age at MRI (h)†	119.5 (108.2–139.1)	120.0 (108.6–136.5)	119.0 (108.3–141.5)	.55
Maternal race or ethnicity				.90
Asian	27 (6.5)	20 (6.6)	7 (6.2)	
Black	54 (13)	42 (14)	12 (11)	
White	300 (72)	215 (71)	85 (75)	
Multiple	13 (3.1)	10 (3.3)	3 (2.7)	
Unknown	22 (5.3)	16 (5.3)	6 (5.3)	

Note.—Unless stated otherwise, data are numbers of participants, with percentages in parentheses. For statistical comparisons, continuous variables were compared using a two-tailed t test for parametric data and Wilcoxon rank-sum test for nonparametric data, and proportional variables were compared using a χ^2 test.

* Data are means ± SDs.

† Data are medians, with IQRs in parentheses.

Table 2: Neurodevelopmental Follow-up at 24 Months Stratified by Presence of Acute Brain Injury at Neonatal MRI

Parameter	All Participants (n = 387)	Participants without Acute Injury (n = 280)	Participants with Acute Injury (n = 107)	P Value
Death or any NDI	185 (48)	113 (40)	72 (67)	<.001
Death	30 (7.8)	3 (1.1)	27 (25)	<.001
Any NDI	155 (40)	110 (39)	45 (42)	.009
Moderate to severe NDI*	111 (29)	70 (25)	41 (38)	<.001
Cerebral palsy	48 (12)	24 (8.6)	24 (22)	<.001
GMFCS score†	0 (0, 0.5)	0 (0, 0)	0 (1, 4)	<.001
Bayley Scales of Infant Development cognitive score‡	89 ± 17	91 ± 15	84 ± 20	.001

Note.—Unless otherwise specified, data are numbers of participants, with percentages in parentheses. For statistical comparisons, continuous variables were compared using a two-tailed t test for parametric data and Wilcoxon rank-sum test for nonparametric data, and proportional variables were compared using a χ^2 test. GMFCS = Gross Motor Function Classification System, NDI = neurodevelopmental impairment.

* One participant with NDI did not have sufficient follow-up information to assess severity of NDI.

† Data are medians, with 75th and 90th percentiles in parentheses.

‡ Data are means ± SDs.

had moderate to severe NDI and 48 of 387 (12%) had cerebral palsy, which were secondary outcomes of clinical interest and were not mutually exclusive.

Volume of Acute Brain Injury

Figure 2 shows two representative examples of corresponding DWI and ADC images with and without acute brain injury segmentation overlays, one from a participant with basal ganglia–thalamic injury pattern and one from a participant with watershed injury pattern. Injury volumes were calculated across the entire imaging cohort, including 303 of 416 participants (73%) with no detectable acute injury (resulting in medians of 0). The median volume of acute injury in the whole brain was 0 mL (75th, 90th percentiles: 0, 6.2 mL). The proportion of participants with volume of acute brain injury greater than 1 mL was 66 of 416 (16%). The median

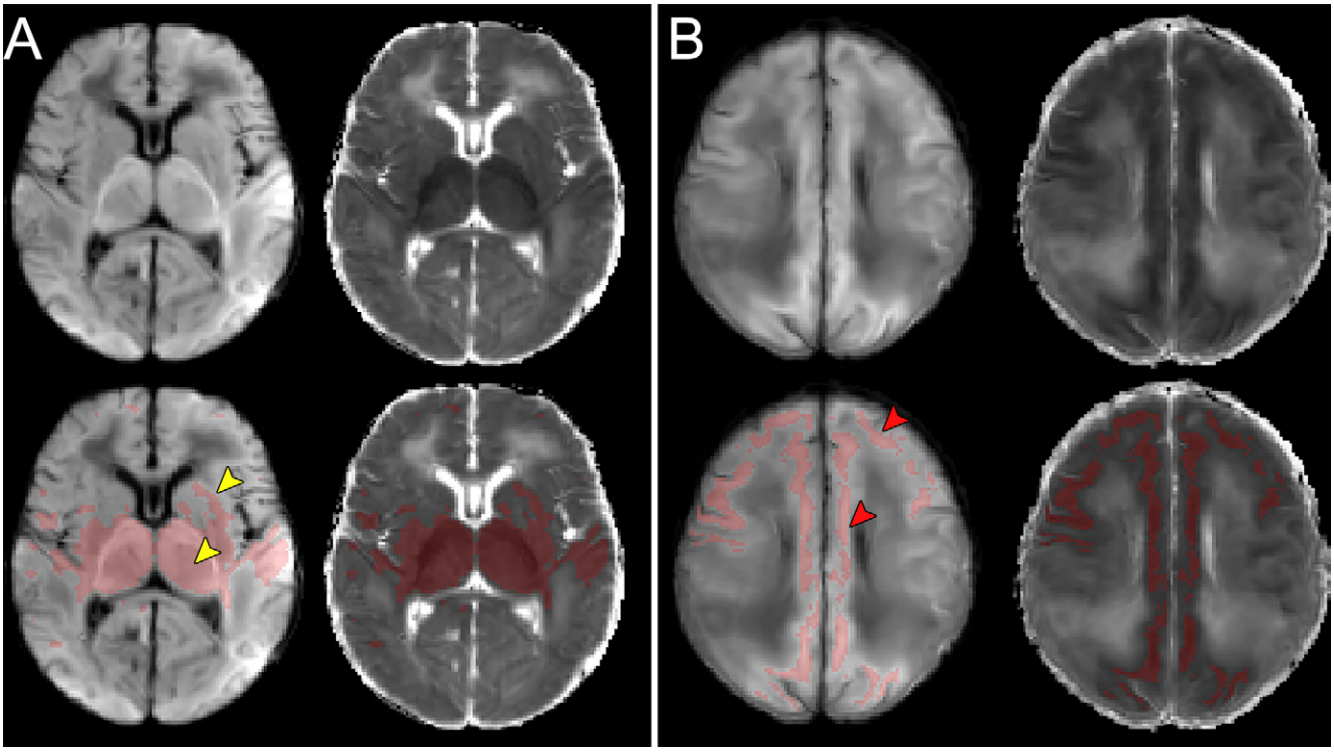

Figure 2: Representative examples of MRI-based diffusion-weighted images (left) and corresponding apparent diffusion coefficient (ADC) maps (right) without (top row) and with (bottom row) automated acute injury segmentation color overlays. **(A)** Axial diffusion-weighted and ADC images at the level of the thalamus demonstrate a basal ganglia–thalamic pattern of injury in a participant with basal ganglia–thalamic predominant pattern of acute injury (arrowheads). This male participant was born via emergent Cesarian section at 40+6 weeks after uterine rupture with Apgar scores of 1, 3, and 4 at 1, 5, and 10 minutes of life, respectively. The participant was initially treated for pulmonary hypertension, systemic hypotension, and disseminated intravascular coagulation. Subsequent electroencephalography consistently showed low voltage burst suppression, and the participant died on day of life 5 after transition to comfort care. **(B)** Axial diffusion-weighted and ADC images at the level of the centrum semiovale show a watershed pattern of injury in a participant with cortical and subcortical predominate watershed pattern of acute injury (arrowheads). This male participant was born at 40+0 weeks gestation via unassisted vaginal delivery and was noted to have a tight nuchal umbilical cord. Apgar scores were 2, 3, and 3 at 1, 5, and 10 minutes of life, respectively. The participant was treated for systemic hypotension and subclinical seizures with good response and was subsequently discharged home off all antiseizure medication.

volume of acute injury was 0 mL (75th, 90th percentiles: 0, 0.2 mL) in the thalami and 0 mL (75th, 90th percentiles: 0, 1.8 mL) in the white matter. A frequency map of acute injury for the study cohort stratified by 2-year neurodevelopmental outcome is shown in Figure 3.

Whole-Brain and Regional ADC Analysis

The distribution of acute injury volume in the whole brain was different between outcome groups, with a median of 0 mL (75th, 90th percentiles: 2.1, 68.4 mL) for participants who died or showed any NDI versus 0 mL (75th, 90th percentiles: 0, 0.03 mL) for those without NDI ($P < .001$) (Table 3). The proportion of participants with a brain injury volume greater than 1 mL was higher in the death or any NDI outcome group (54 of 185 participants [29%]) compared with the no-NDI outcome group (eight of 202 [4.0%]; $P < .001$) (Table 3). Furthermore, differences in the distribution of acute injury volume in the thalami were observed between participants who died or had any NDI (median volume, 0 mL [75th, 90th percentiles: 0, 2.4 mL]) and those with no NDI (median volume, 0 mL [75th, 90th percentiles: 0, 0 mL]; $P < .001$) (Table 3). The distribution of acute injury volume in the white matter was also different between outcome groups (death or any NDI median volume, 0 mL [75th, 90th percentiles: 0.6, 23.0 mL] vs no-NDI median volume, 0 mL [75th, 90th percentiles: 0, 0 mL]; $P < .001$) (Table 3). No differences were observed in global and regional injury volumes between participants with and without available outcome data at 2 years (Table S4).

Association of Acute Brain Injury Volume and Qualitative Score with Neurodevelopmental Outcomes

Presence and volume of acute injury in the whole brain and regionally in the thalami and white matter were each associated with several neurodevelopmental outcome measures (Table 4). An acute injury volume greater than 1 mL was associated with higher odds of death (odds ratio [OR], 42.2 [95% CI: 15.3, 116.2]; $P < .001$), moderate to severe NDI (OR, 2.00 [95% CI: 1.11, 3.61]; $P = .02$), and cerebral palsy (OR, 2.74 [95% CI: 1.3, 5.6]; $P = .006$). The presence of any acute injury in the whole brain, thalami, and white matter was also associated with death, moderate to severe NDI, and cerebral palsy (Table 4).

Expert reader DWI scores for the whole brain, thalami, and white matter were also associated with death, moderate to severe NDI, and cerebral palsy (Table 4).

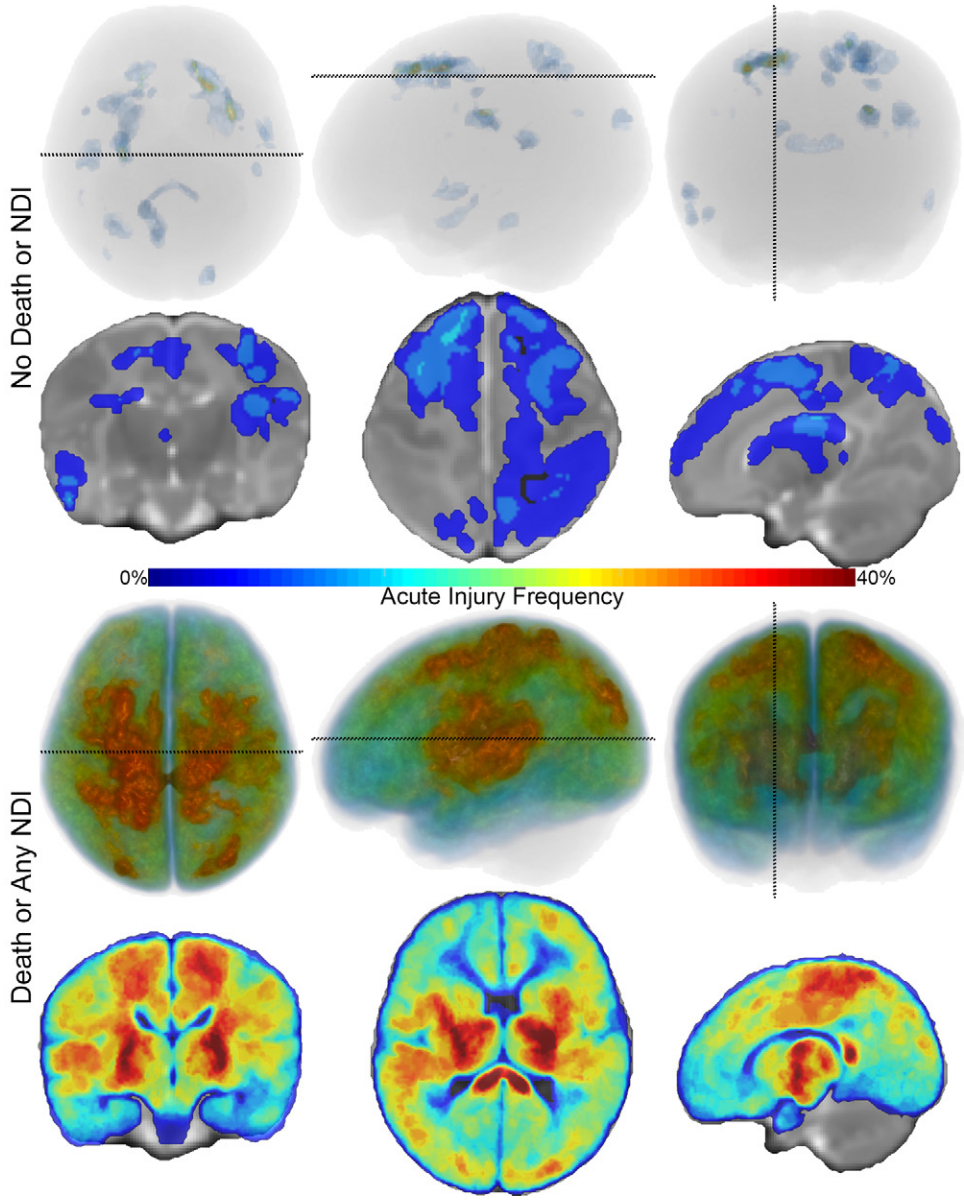

Figure 3: Acute injury frequency map for the two primary outcome groups displayed as colorized heat maps. Frequency maps for participants who did not die or have neurodevelopmental impairment (NDI) at follow-up are shown in the top two rows. The top row shows superior, lateral, and anterior projection views of volume rendered injury maps, with cool colors indicating relatively low injury frequency. The second row shows coronal, axial, and sagittal sections of the population-averaged neonatal brain atlas with colorized heat map overlays using the same color scale. The location of the displayed sections is indicated on the volume rendered projections with a dotted line. Brain regions where there was no acute injury in this outcome group are displayed as clear to show the underlying brain atlas anatomy (gray scale). Rows 3 and 4 are identical to rows 1 and 2 except that they represent frequency of acute injury in participants who died or had NDI at follow-up. Warm colors indicate a relatively higher frequency of acute injury and correspond to typical locations of acute brain injury described in the setting of hypoxic ischemic encephalopathy.

There were eight neonates with volume of acute injury greater than 1 mL who had no NDI. Further evaluation of the MRI data revealed that these participants had injury patterns different from hypoxic ischemic injury, including perinatal arterial ischemic stroke (n = 6), focal right splenial reduced diffusion (n = 1), and bilateral occipital cortical and white matter reduced diffusion suggestive of neonatal hypoglycemia (n = 1).

Correlation between Quantitative and Qualitative Injury Metrics

Quantitative whole-brain acute injury volume was strongly associated with expert qualitative scoring of global DWI signal abnormality (Kendall tau-b = 0.56; probability of rank agreement = 0.78; P < .001). Similarly, acute injury volume in the white matter correlated strongly with the qualitative white matter DWI score (Kendall tau-b = 0.62; probability of rank

Table 3: Automated Quantitative Neonatal MRI–based Apparent Diffusion Coefficient Metrics of Brain Injury Stratified by Neurodevelopmental Outcome at 24-month Follow-up

Metric	All Participants (n = 387)	No NDI (n = 202)	Death or Any NDI (n = 185)	P Value
Acute injury volume (mL)*				
Whole brain	0 (0, 6.2)	0 (0, 0.03)	0 (2.1, 68.4)	<.001
Thalamus	0 (0, 0.2)	0 (0, 0)	0 (0, 2.4)	<.001
White matter	0 (0, 1.8)	0 (0, 0)	0 (0.6, 23.0)	<.001
Presence of acute injury				
Whole brain	107 (28)	35 (17)	72 (39)	<.001
Thalamus	45 (12)	1 (0.5)	44 (24)	<.001
White matter	84 (22)	20 (10)	64 (35)	<.001
Volume of global acute injury >1 mL	62 (16)	8 (4.0)	54 (29)	<.001

Note.—Unless otherwise specified, data are numbers of participants, with percentages in parentheses. For statistical comparisons, continuous variables were compared using Wilcoxon rank-sum test, and proportional variables were compared using a χ^2 test. NDI = neurodevelopmental impairment.

* Data are medians, with 75th and 90th percentiles in parentheses.

agreement = 0.81; P < .001), and acute injury volume in the thalamus correlated strongly with the thalamic DWI score (Kendall tau-b = 0.67; probability of rank agreement = 0.84; P < .001).

Discussion

Neonatal hypoxic ischemic injury is a major cause of neonatal death and chronic neurologic disability and an area of active research into new therapeutic adjuncts to therapeutic hypothermia. At present, standard-of-care neonatal MRI can be used to identify children at highest risk for major disability but is limited in its ability to discriminate between degrees of neurodevelopmental impairment (20). Over the years, multiple qualitative MRI scoring systems (8–14) have been created to classify patterns and severity of hypoxic ischemic injury in an effort to predict outcome and to serve as a secondary end point for clinical trials. These scoring systems are all subject to interrater variability (13) and, for purposes of clinical trials, require time-consuming consensus scoring to achieve agreement among experts. This highlights the important need to create an objective, automated, quantitative method for capturing injury severity in neonates with hypoxic ischemic encephalopathy. The specific benefits of automated image quantification include objectivity of results, reproducibility across sites and users, and flexibility to adjust injury thresholds and incorporate results into statistical or predictive models.

In this secondary analysis of the HEAL randomized controlled trial, we performed automated, quantitative, volumetric ADC analysis of neonatal brain MRI scans in 416 participants with HIE. We identified acute injury in 113 of 416 participants (27%). A global volume of acute injury greater than 1 mL was strongly associated with death (OR, 42.2 [95% CI: 15.3, 116.2]; P < .001), moderate to severe NDI (OR, 2.00 [95% CI: 1.11, 3.61]; P = .02), and cerebral palsy (OR, 2.74 [95% CI: 1.3, 5.6]; P = .006) at 24 months. An acute injury volume of 1 mL is clinically relevant and visually detectable (examples in Fig S5) and corresponds to approximately 10 × 10 × 10 abnormal pixels in a 1-mm^3 isotropic image.

The distribution of volume of acute injury was different between participants with an outcome of death or any NDI (0 mL [75th, 95th percentiles: 2.1, 68.4 mL]) versus those with an outcome of no NDI (0 mL [75th, 95th percentiles: 0, 0.03 mL]; P < .001). Furthermore, a higher proportion of participants with acute brain injury than without experienced outcomes of death or any NDI (67% with acute injury vs 40% without acute injury; P < .001), death (25% with acute injury vs 1.1% without acute injury; P < .001), moderate to severe NDI (38% with acute injury vs 25% without acute injury; P < .001), and cerebral palsy (22% with acute injury vs 8.6% without acute injury; P < .001) at 24-month follow-up.

Our automated quantitative acute injury assessment showed a strong positive correlation with radiologists' qualitative scoring of injury on DWIs (Kendall tau-b = 0.56; P < .001), which we used as a reference standard. Because our study aimed to determine the association between automated acute injury scores and HIE outcomes, a direct comparison of the predictive value of our automated quantitative method versus the predictive value of the qualitative injury scores is beyond the scope of this study. We did, however, find that the automated acute injury volume in the thalamus was very strongly associated with death or any NDI (OR, 61.3). In contrast, the qualitative MRI score in the thalamus was also associated with death or any NDI, but with a more modest OR of 2.2. This study contributes to the current body of literature in neonatal HIE by describing a new method for automated quantification of acute brain injury that is closely associated with expert human consensus injury scoring and is highly correlated with patient outcomes. Further studies that focus on predictive modeling are needed to determine the specific quantification method that will more accurately predict outcomes after HIE.

Several smaller prior studies have explored the association of quantitative ADC values with outcome in neonatal HIE. Unlike our study, these prior studies explored the magnitude of ADC reduction in association with outcome, not volume of acute injury as defined by thresholded ADC values. Vermeulen et al (36) found that in neonates imaged 0–4 days after birth, those with poor motor outcome (defined as death or 2-year Bayley Scales of Infant Development motor score less than 70) had lower ADC values in multiple brain regions that corresponded to the basal ganglia/thalamic pattern of injury. Our study builds on the findings of this prior study, with several advancements: We used an automated quantitative technique to evaluate volume of acute injury instead of magnitude of ADC reduction in a manually drawn region of interest, thus providing a more comprehensive and objective measure of injury

Table 4: Multiple Logistic Regression Evaluating Association between Neonatal MRI–based Apparent Diffusion Coefficient Metrics and Neurodevelopmental Outcome at 24-month Follow-up

	Death or Any NDI (n = 185)		Death (n = 30)		Moderate to Severe NDI (n = 111)		Cerebral Palsy (n = 48)	
Metric	OR	*P* Value	OR	*P* Value	OR	*P* Value	OR	*P* Value
Acute injury volume (mL)								
Whole brain	1.45 (1.07, 1.97)	.02	1.0 (1.0, 1.0)	<.001	1.0 (1.0, 1.0)	.38	1.0 (1.0, 1.0)	.39
Thalamus*	61.3 (10.1, 369.4)	<.001	3.9 (2.7, 6.2)	<.001	0.84 (0.64, 1.10)	.20	1.0 (0.74, 1.39)	.95
White matter	1.45 (1.07, 2.0)	.02	1.06 (1.04, 1.08)	<.001	1.0 (0.98, 1.0)	.42	1.0 (0.99, 1.02)	.40
Volume of global acute injury >1 mL	13.9 (5.93, 32.45)	<.001	42.2 (15.3, 116.2)	<.001	2.00 (1.11, 3.61)	.02	2.74 (1.3, 5.6)	.006
Presence of acute injury								
Whole brain	4.5 (2.6, 7.8)	<.001	32.3 (9.3, 111.5)	<.001	1.45 (1.36, 3.79)	.002	3.4 (1.8, 6.5)	<.001
Thalamus	89.8 (11.7, 692.2)	<.001	48.3 (17.9, 147.4)	<.001	1.88 (0.96, 3.68)	.07	3.67 (1.67, 8.04)	.001
White matter	6.3 (3.4, 11.7)	<.001	34.5 (11.6, 102.7)	<.001	2.33 (1.37, 3.98)	.002	3.9 (2.0, 7.7)	<.001
DWI score								
Whole brain	1.18 (1.12, 1.23)	<.001	1.22 (1.1, 1.3)	<.001	1.02 (1.00, 1.05)	.04	1.06 (1.03, 1.09)	<.001
Thalamus	2.2 (1.7, 2.8)	<.001	2.9 (2.2, 3.7)	<.001	1.15 (1.02, 1.31)	.03	1.45 (1.24, 1.69)	<.001
White matter	1.5 (1.3, 1.7)	<.001	2.1 (1.7, 2.7)	<.001	1.18 (1.05, 1.32)	.03	1.36 (1.17, 1.56)	<.001

Note.—Data in parentheses are 95% CIs. All regressions were adjusted for sex, gestational age, and erythropoietin administration, with site of enrollment as a random effect. DWI = diffusion-weighted image, NDI = neurodevelopmental impairment, OR = odds ratio.

* Exact multiple logistic regression was used due to issues of quasi-separation of the predictor among outcome groups.

severity. Furthermore, our primary and secondary neurodevelopmental outcomes capture a wider, more nuanced spectrum of developmental abnormality, with analysis of cognitive impairment and differing degrees of NDI, all of which were not assessed in this prior study. In a similar study, Hunt et al (37) explored the lowest ADC values in the posterior limb of the internal capsule in neonates with HIE and found the lowest ADC values to be different between neonates who survived versus those who died and to be different across neuromotor outcome groups, with lower ADC values in those with severe neuromotor impairment. Our study again improves on this prior study by providing an automated calculation of acute injury volume and exploring its association with a wider range of clinically important neurodevelopmental outcomes. Mulkey et al (25) used manually delineated regions of interest to manually calculate a whole-brain volume of acute injury and found that higher volume of acute injury correlated with lower 12-month neurodevelopmental scores. Our study advances this by using an automated volumetric method for acute injury volume calculation and correlating with longer-term, 24-month neurodevelopmental outcomes.

All participants in our study were prospectively enrolled in the HEAL trial based on a clinical picture of moderate to severe encephalopathy and suspected HIE. Nonetheless, some of these participants may have experienced neonatal encephalopathy secondary to other causes, some of which (such as perinatal stroke) are known to have relatively better outcomes (38). In our evaluation of patients with greater than 1 mL of acute injury with relatively favorable outcomes, we found these patients had patterns of injury that were atypical for hypoxic ischemic injury. Thus, the greater than 1 mL threshold of acute injury volume that is associated with death and NDI may not be as generalizable to infants who have causes of neonatal encephalopathy other than HIE.

Our study has several limitations. First, nondiffusion MRI signal abnormalities (eg, T1- and T2-weighted) that could represent subacute or chronic injury were not considered. Analysis of signal abnormality at multisequence neonatal MRI (ie, including T1-, T2-, and susceptibility-weighted imaging) will be an important future area of research. Second, since approximately half of participants in this cohort received erythropoietin, the results may be confounded by this intervention, although the parent HEAL trial found no therapeutic benefit to erythropoietin (17) and we found no difference in the presence of acute injury between treatment groups. Third, as our study was performed primarily for validation of our automated method, we did not explore the association of volumes of acute injury in all regions with outcome, but intentionally only analyzed two regions defined a priori based on prior research. This does, however, limit our ability to detect other regions of the brain that may influence developmental outcome. Fourth, ADC can be heterogeneous across multiple sites. To mitigate this potential variability, we calculated ADC values from DWI by using third-party software (FMRIB Software Library) instead of relying on vendor ADC maps. Finally, participants with substantial acute intracranial hemorrhage and incomplete DTI data were excluded. While the number of excluded participants was relatively low

(58 of 474 [12%] overall and nine of 474 [1.9%] for intracranial hemorrhage), our results may not be relevant to cohorts with substantial intracranial hemorrhage.

Automated quantitative diffusion MRI–based metrics of acute brain injury in neonates with hypoxic ischemic encephalopathy were strongly associated with subjective grading of brain injury at MRI by expert radiologists. Additionally, the presence of any acute brain injury as well as volume of acute injury greater than 1 mL were associated with death, moderate to severe neurodevelopmental impairment, and cerebral palsy. Future research in this field should explore quantitative multiparametric analysis of different sequences in association with neurodevelopmental outcome.

Author contributions: Guarantors of integrity of entire study, **E.C., Y.W., A.M., Y.L.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **E.C., A.M., S.G., Y.L.**; clinical studies, **E.C., Y.W., J.L.W., A.M., H.C.G., S.G., S.E.J., C.P.H., Y.L.**; experimental studies, **R.C.M., Y.L.**; statistical analysis, **A.W.S., J.L.W., B.A.C., P.J.H., Y.L.**; and manuscript editing, **E.C., Y.W., A.W.S., J.L.W., R.C.M., A.M., H.C.G., P.J.H., S.G., S.E.J., C.P.H., Y.L.**

Data sharing: Data analyzed during the study were provided by a third party. Requests for data should be directed to the corresponding author.

Disclosures of conflicts of interest: E.C. Research grant from the American Society of Neuroradiology; royalties from Elsevier. **Y.W.** Research grant from the National Institutes of Health; payment for expert testimony; scientific advisory board member for the Cerebral Palsy Alliance Research Foundation. **A.W.S.** Training grant from the American Society of Pediatric Neuroradiology. **J.L.W.** No relevant relationships. **R.C.M.** Grant from the National Institutes of Health; honoraria for lectures from Siemens Healthineers; support for travel from Siemens Healthineers and Philips Healthcare; medical advisory board member for Turing Medical (formerly Nous Imaging); stock in Turing Medical. **A.M.** No relevant relationships. **H.C.G.** Grants to institution from the National Institutes of Health; payment for expert opinion; participation on the data safety monitoring board for iAQUIRE (National Institutes of Health); board of directors member for the Pediatric Academic Societies; spouse holds shares in Elemeno Health. **B.A.C.** Grant from the National Institute of Neurological Disorders and Stroke. **P.J.H.** Grant from the National Institutes of Health. **S.G.** No relevant relationships. **S.E.J.** Grants to institution from the National Institutes of Health; royalties from Elsevier; patent planned, issued, or pending with the University of Washington; participation on the data safety monitoring board for ALBINO; director of the Intellectual and Developmental Disabilities Research Center at the University of Washington and the Institute on Human Development and Disabilities. **C.P.H.** Consulting fees from GE Healthcare; medical advisory board membership for Kheiron Medical Technologies; participation on a data safety monitoring board for the Focused Ultrasound Foundation and uniQure Biopharma. **Y.L.** No relevant relationships.

References

1. Kurinczuk JJ, White-Koning M, Badawi N. Epidemiology of neonatal encephalopathy and hypoxic-ischaemic encephalopathy. Early Hum Dev 2010;86(6):329–338.
2. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet 2015;385(9966):430–440.
3. Tagin MA, Woolcott CG, Vincer MJ, Whyte RK, Stinson DA. Hypothermia for neonatal hypoxic ischemic encephalopathy: an updated systematic review and meta-analysis. Arch Pediatr Adolesc Med 2012;166(6):558–566.
4. Shankaran S, Pappas A, McDonald SA, et al. Childhood outcomes after hypothermia for neonatal encephalopathy. N Engl J Med 2012;366(22):2085–2092. [Published correction appears in N Engl J Med 2012;367(11):1073.]
5. Sarnat HB, Sarnat MS. Neonatal encephalopathy following fetal distress. A clinical and electroencephalographic study. Arch Neurol 1976; 33(10):696–705.
6. Jacobs SE, Berg M, Hunt R, Tarnow-Mordi WO, Inder TE, Davis PG. Cooling for newborns with hypoxic ischaemic encephalopathy. Cochrane Libr 2013;2013(1):CD003311.
7. Edwards AD, Brocklehurst P, Gunn AJ, et al. Neurological outcomes at 18 months of age after moderate hypothermia for perinatal hypoxic ischaemic encephalopathy: synthesis and meta-analysis of trial data. BMJ 2010;340:c363.
8. Barkovich AJ, Hajnal BL, Vigneron D, et al. Prediction of neuromotor outcome in perinatal asphyxia: evaluation of MR scoring systems. AJNR Am J Neuroradiol 1998;19(1):143–149.
9. Trivedi SB, Vesoulis ZA, Rao R, et al. A validated clinical MRI injury scoring system in neonatal hypoxic-ischemic encephalopathy. Pediatr Radiol 2017;47(11):1491–1499.
10. Rutherford M, Ramenghi LA, Edwards AD, et al. Assessment of brain tissue injury after moderate hypothermia in neonates with hypoxic-ischaemic encephalopathy: a nested substudy of a randomised controlled trial. Lancet Neurol 2010;9(1):39–45.
11. Shankaran S, Barnes PD, Hintz SR, et al. Brain injury following trial of hypothermia for neonatal hypoxic-ischaemic encephalopathy. Arch Dis Child Fetal Neonatal Ed 2012;97(6):F398–F404. [Published correction appears in Arch Dis Child Fetal Neonatal Ed 2014;99(3):301.]
12. Weeke LC, Groenendaal F, Mudigonda K, et al. A novel magnetic resonance imaging score predicts neurodevelopmental outcome after perinatal asphyxia and therapeutic hypothermia. J Pediatr 2018;192:33–40.e2.
13. Goergen SK, Ang H, Wong F, et al. Early MRI in term infants with perinatal hypoxic-ischaemic brain injury: interobserver agreement and MRI predictors of outcome at 2 years. Clin Radiol 2014;69(1):72–81.
14. Cheong JLY, Coleman L, Hunt RW, et al. Prognostic utility of magnetic resonance imaging in neonatal hypoxic-ischemic encephalopathy: substudy of a randomized trial. Arch Pediatr Adolesc Med 2012;166(7): 634–640.
15. Ozturk A, Sasson AD, Farrell JAD, et al. Regional differences in diffusion tensor imaging measurements: assessment of intrarater and interrater variability. AJNR Am J Neuroradiol 2008;29(6):1124–1127.
16. Wisnowski JL, Bluml S, Panigrahy A, et al. Integrating neuroimaging biomarkers into the multicentre, high-dose erythropoietin for asphyxia and encephalopathy (HEAL) trial: rationale, protocol and harmonisation. BMJ Open 2021;11(4):e043852.
17. Wu YW, Comstock BA, Gonzalez FF, et al. Trial of erythropoietin for hypoxic-ischemic encephalopathy in newborns. N Engl J Med 2022;387(2):148–159.
18. Juul SE, Comstock BA, Heagerty PJ, et al. High-dose Erythropoietin for Asphyxia and Encephalopathy (HEAL): a randomized controlled trial – background, aims, and study protocol. Neonatology 2018;113(4):331–338.
19. Glass HC, Wusthoff CJ, Comstock BA, et al. Risk of seizures in neonates with hypoxic-ischemic encephalopathy receiving hypothermia plus erythropoietin or placebo. Pediatr Res 2023;94(1):252–259.
20. Wu YW, Monsell SE, Glass HC, et al. How well does neonatal neuroimaging correlate with neurodevelopmental outcomes in infants with hypoxic-ischemic encephalopathy? Pediatr Res 2023. 10.1038/s41390-023-02510-8. Published online March 1, 2023.
21. Gorgolewski K, Burns C, Madison C, et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 2011;5:13.
22. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. Neuroimage 2012;62(2):782–790.
23. Shi F, Yap PT, Wu G, et al. Infant brain atlases from neonates to 1- and 2-year-olds. PLoS One 2011;6(4):e18746.
24. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 2011;54(3):2033–2044.
25. Mulkey SB, Ramakrishnaiah RH, McKinstry RC, et al. Erythropoietin and brain magnetic resonance imaging findings in hypoxic-ischemic encephalopathy: volume of acute brain injury and 1-year neurodevelopmental outcome. J Pediatr 2017;186:196–199.
26. Bednarek N, Mathur A, Inder T, Wilkinson J, Neil J, Shimony J. Impact of therapeutic hypothermia on MRI diffusion changes in neonatal encephalopathy. Neurology 2012;78(18):1420–1427.
27. Rutherford M, Counsell S, Allsop J, et al. Diffusion-weighted magnetic resonance imaging in term perinatal brain injury: a comparison with site of lesion and time from birth. Pediatrics 2004;114(4):1004–1014.
28. Calabrese E, Gandhi S, Shih J, et al. Parieto-occipital injury on diffusion MRI correlates with poor neurologic outcome following cardiac arrest. AJNR Am J Neuroradiol 2023;44(3):254–260.

29. Barkovich AJ. MR and CT evaluation of profound neonatal and infantile asphyxia. AJNR Am J Neuroradiol 1992;13(3):959–972; discussion 973–975.
30. Baenziger O, Martin E, Steinlin M, et al. Early pattern recognition in severe perinatal asphyxia: a prospective MRI study. Neuroradiology 1993;35(6):437–442.
31. Rutherford MA, Pennock JM, Dubowitz LMS. Cranial ultrasound and magnetic resonance imaging in hypoxic-ischemic encephalopathy: a comparison with outcome. Dev Med Child Neurol 1994;36(9):813–825.
32. Kuban KCK, Allred EN, O'Shea M, et al. An algorithm for identifying and classifying cerebral palsy in young children. J Pediatr 2008;153(4):466–472.
33. Kerridge DF. The interpretation of rank correlations. J R Stat Soc Ser C Appl Stat 1975;24(2):257–258.
34. Botsch R. Significance and Measures of Association. In: Scopes and Methods of Political Science. https://polisci.usca.edu/apls301/Text/. Published 2011. Accessed November 2022.
35. Mansournia MA, Geroldinger A, Greenland S, Heinze G. Separation in logistic regression: causes, consequences, and control. Am J Epidemiol 2018;187(4):864–870.
36. Vermeulen RJ, van Schie PEM, Hendrikx L, et al. Diffusion-weighted and conventional MR imaging in neonatal hypoxic ischemia: two-year follow-up study. Radiology 2008;249(2):631–639.
37. Hunt RW, Neil JJ, Coleman LT, Kean MJ, Inder TE. Apparent diffusion coefficient in the posterior limb of the internal capsule predicts outcome after perinatal asphyxia. Pediatrics 2004;114(4):999–1003.
38. Jan MMS, Camfield PR. Outcome of neonatal stroke in full-term infants without gnificant birth asphyxia. Eur J Pediatr 1998;157(10):846–848.