

Preschool Language Outcomes following Perinatal Hypoxic-Ischemic Encephalopathy in the Age of Therapeutic Hypothermia

Eric M. Chin^{a, b} Srishti Jayakumar^{a, b} Ezequiel Ramos^{a, b} Gwendolyn Gerner^{a, b}
Bruno P. Soares^{b, d} Elizabeth Cristofalo^{b, c} Mary Leppert^{a, b} Marilee Allen^{a, b}
Charla Parkinson^b Michael Johnston^{a, b} Frances Northington^{a, b} Vera Joanna Burton^{a, b}

^aDepartment of Neurology and Developmental Medicine, Kennedy Krieger Institute, Baltimore, MD, USA;

^bNeurosciences Intensive Care Nursery, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA; ^cDepartment of Neonatology, Frederick Memorial Hospital, Frederick, MD, USA;

^dNeuroradiology, Larner College of Medicine at The University of Vermont, Burlington, VT, USA

Keywords

Language development · Hypoxic-ischemic encephalopathy · Therapeutic hypothermia

Abstract

Early studies following perinatal hypoxic-ischemic encephalopathy (HIE) suggested expressive language deficits and academic difficulties, but there is only limited detailed study of language development in this population since the widespread adoption of therapeutic hypothermia (TH). Expressive and receptive language testing was performed as part of a larger battery with 45 children with a mean age of 26 months following perinatal HIE treated with TH. Overall cohort outcomes as well as the effects of gender, estimated household income, initial pH and base excess, and pattern of injury on neonatal brain MRI were assessed. The cohort overall demonstrated expressive language subscore, visual-reception subscore, and early learning composite scores significantly below test norms, with relative sparing of receptive language subscores. Poorer expressive language manifested as decreased vocabulary size and shorter utterances. Expressive language subscores showed a significant gender effect, and estimated socioeconomic status showed a significant effect

on both receptive and expressive language subscores. Initial blood gas markers and modified Sarnat scoring did not show a significant effect on language subscores. Binarized MRI abnormality predicted a significant effect on both receptive and expressive language subscores; the presence of specific cortical/subcortical abnormalities predicted receptive language deficits. Overall, the language development profile of children following HIE in the era of hypothermia shows a relative strength in receptive language. Gender and socioeconomic status predominantly predict expressive language deficits; abnormalities detectable on MRI predominantly predict receptive language deficits.

© 2019 S. Karger AG, Basel

Introduction

Neonatal hypoxic-ischemic encephalopathy (HIE) affects approximately 1.5 in 1,000 births and is the most common cause of perinatal brain injury in full-term neonates [1]. Even with therapeutic hypothermia (TH), now the standard of care when available for neonates with HIE, there is incomplete neuroprotection; on long-term follow-up, 17% of these children develop cerebral palsy

(CP), and 27% demonstrate IQ <70 [2]. Before the advent of TH, a number of studies demonstrated poorer speech and language skills including selective reading and spelling difficulties at school age in children with HIE even in the absence of more serious cognitive or motor difficulties [3–5]; it is therefore critical to evaluate speech and language skills in the era of TH.

There are only a few evaluations of language development in children with HIE post-TH. Both the NICHD and TOBY trial of hypothermia for perinatal HIE performed neurocognitive follow-up studies on children aged 6–7 years (using either the Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition [6] or the Wechsler Intelligence Scale for Children, 4th Edition [7]). The NICHD follow-up study [2] noted low-to-normal verbal IQ in both the hypothermia (mean standard score 85.9) and normothermia (mean standard score 86.4) groups, although a fraction of the sample population was deemed to be too low-functioning to assess verbal IQ. The TOBY trial [8] reported average verbal IQ in both the hypothermia (mean standard score 105) and normothermia (mean standard score 101) groups. While correlations between verbal IQ and scores on language assessments are generally moderate to high [9], they measure different skills and differently predict academic function [10].

In attempts to understand later neurodevelopmental outcomes following neonatal HIE, degree of injury as measured by modified Sarnat score, blood biomarkers, and patterns of brain injury seen on MRI have been used to predict outcome [11, 12]. There is evidence that such measures predict broad outcomes such as mortality, and eventual diagnoses of CP and intellectual disability [11]. Little is known about the value of these markers in predicting disorders that are more frequent but less severe, e.g., language impairment. To date, the most promising predictive markers have been MRI findings, i.e., white-matter or basal-ganglia injury have been shown to correlate with lower language scores at age 30 months [14, 18] and with both lower performance and verbal scores at the age of 4 years [14].

In our cohort of children with moderate-to-severe HIE treated with TH, we hypothesized that while gender and socioeconomic status would influence language outcomes [15–17], some early markers of hypoxic-ischemic injury severity (initial pH/base deficit, and severity of HIE on postcooling MRI) would also correlate with language outcomes. We also hypothesized that specific injury patterns seen on MRI may predict certain types of language skill deficits. Specifically, basal ganglia injury has a strong relationship with motor impairment, thought to affect speech and would therefore be expected to predominantly affect

expressive language measures [18]. Cortical injury might be expected to more likely affect language learning, resulting in receptive language deficits [13, 14, 19].

Materials and Methods

Study Design

This is a convenience sample of infants with HIE who were treated with TH and participated in a comprehensive follow-up evaluation at the age of 2 years.

Neonatal Time Point

All neonates who underwent whole-body cooling for moderate-to-severe encephalopathy between 2010 and 2014 at the Johns Hopkins Children's Hospital neonatal intensive care unit (NICU) and participated in a comprehensive neurodevelopmental battery at the age of 2 years were included in the study. Infants eligible for cooling were diagnosed with moderate-to-severe HIE on clinical exam based on the modified Sarnat criteria [20, 21] and blood gas from the umbilical cord or first hour of life with a pH <7.15 or base deficit >10 mmol/L. If the blood gas measurement was not available, a 10-min Apgar score <5 or assisted ventilation for ≥10 min after birth, an acute perinatal event, and moderate-to-severe encephalopathy were used to diagnose HIE. Additional eligibility criteria for this study included a gestational age ≥35 weeks, a birth weight ≥1,800 g, the initiation of whole-body cooling within 6 h of birth, and a parent with English as their primary language. Neonates with a contraindication to hypothermia therapy (e.g., coagulopathy), with active bleeding or congenital anomalies that could make cooling unsafe, were not eligible for the study.

Variables gathered from the NICU period used for analysis in this study included initial blood gas (either cord or within the first hour of life, as available), pH, and base deficit. Lactate was not consistently obtained in this cohort (Table 1). Household income was not directly assessed on admission, but socioeconomic status was estimated using well-established [22, 23] proxies: (1) a 2015 median income of the census tract containing the family's home address at the time of delivery and (2) family insurance status (private vs. medical assistance/public).

Categorization of encephalopathy was determined by review of neurologic status as described in the admission history and physical exam. A neonatologist and a pediatric neurologist, both with additional training in developmental disabilities, independently reviewed the admission records to determine moderate or severe encephalopathy. Level of encephalopathy was determined using modified Sarnat criteria [20, 21]. If infants had features from both categories, the level of encephalopathy was determined with the most criteria present. Agreement between the record reviewers was $\kappa = 0.78$. Disagreements were reviewed and determined by consensus.

MRI was obtained as a part of usual NICU care in the first 2 weeks of life in 44/45 subjects (average age 8.7 ± 3.0 days). NICHD MRI grade (a grading scheme ranging between 0 [or no detectable HIE-related injury] to 3 ["global devastation"]) with a distinction between cortical/subcortical and basal ganglia/thalamic injury was based on T1- and T2-weighted sequences. See Table 1 for a brief description of NICHD MRI grades (described in detail in [24]) assessed by a board-certified pediatric neuroradiologist blinded to outcomes.

Table 1. Participants' characteristics

	<i>n</i>	%	Mean ± SD	Range
Female sex	21/45	47.0		
Private insurance	34/45	76.0		
Maternal education of 12 years (completed high school)	18/45		15.1±2.1	11–17
Annual household income (USD 1,000)				
Parent-reported ^a	18/45		126.1±83.7	31.2–305.9
Estimated from median by census tract	45/45		75.5±33.4	16.9–167.6
<i>NICU biomarkers</i>				
Modified Sarnat score ^b				
2 (moderate encephalopathy)	32/45	71.0		
3 (severe encephalopathy)	13/45	29.0		
NICHD MRI grade ^c				
0 (normal)	25/44	57.0		
1A (minimal cerebral lesions)	11/44	25.0		
1B (more extensive cerebral lesions without other involvement)	2/44	4.5		
2A (basal ganglia, thalamic, internal capsule lesions only)	3/44	6.8		
2B (basal ganglia, thalamic, internal capsule, and cerebral lesions)	3/44	6.8		
3 (cerebral devastation)	0/44	0.0		
pH	45/45		6.96±0.13	6.62 to 7.26
Base excess/deficit	43/45		−15.9±6.2	−33.8 to −7.0
Serum lactate (mmol/L)	22/45		3.9±4.1	0.7 to 16.7
<i>Other diagnoses at age 2 years</i>				
Cerebral palsy	3/45	6.7		

Measures obtained at NICU admission include proxies of socioeconomic status (parental public vs. private insurance and home census tract-based estimate of household income) and injury markers.

^a Obtained from the 2-year research visit.

^b Documented degree of encephalopathy on exam within the first 6 h of life.

^c Both the degree and gross distribution of brain injury on conventional MRI and included the categories: normal MRI (0), cortical/subcortical involvement only (1A) and more extensive (1B), basal ganglia/thalamic/deep white matter involvement only (2A), or involvement of both areas (2B), and (3). Parent-reported household tax bracket obtained from the 2-year research visit.

Preschool Age Time Point

As the standard of care (since 2010), all children treated for neonatal HIE in our NICU are referred for neurodevelopmental follow-up at discharge. Families were eligible to participate in a research evaluation between the ages of 18 and 30 months. Children who were involved in the foster-care system at the time of neurodevelopmental follow-up were not included in the research study.

As part of a larger battery, the research evaluation included the Mullen Scales of Early Learning (MSEL; Pearson Education, Inc., London, UK) and the MacArthur-Bates Communicative Development Inventories (MB-CDI; Brookes Publishing; Baltimore, MD, USA). MSEL is a developmental battery assessing gross motor (GM), fine motor (FM), visual reception (VR), expressive language (EL), and receptive language (RL), which also provides a summary early learning composite (ELC) incorporating information from the VR, EL, RL, and FM subscores. Testing norms are established for children from birth to 68 months of age. Subscores are reported as T-scores (population mean of 50 [SD 10] and floor of instrument 20), and the ELC is reported as a standard score (population mean of 100 [SD 15]). The MB-CDI is a detailed, standardized parent-report instrument including subscores describing the over-

all volume of word production (MB-PROD) and average length of the longest 3 sentences (MB-M3L). No other MB-CDI subscores were included in this analysis as they reflect skills (e.g., correct contextual matching of later-developed word endings) that are not expected to be reliably present at this age. Norms for age and gender are reported as percentiles and are available for the age range 8–30 months. In normative samples, a subscore performance below the 10th percentile was deemed significant for language disability.

Forty-five children in this convenience sample had completed at least MSEL-RL, MSEL-EL, and MSEL-VR, and were included in the analysis. Of these, 44/45 completed all MSEL subscores, and completed MB-CDI reports were obtained for 33/45 within the normed window. We were particularly interested in MB-PROD and MB-M3L as measures of speech output and complexity, respectively; for these measures, normed percentile values were available for 33/45 and 32/45 subjects, respectively.

In addition to primary study outcome measurements, other information collected at the preschool-age time point included whether the individual had been clinically diagnosed with CP. Additional demographic information (e.g., household tax bracket) was also elicited but with an incomplete response (Table 1).

Table 2. Whole-cohort testing characteristics

	Estimate	95% CI	<i>p</i> value
Mullen RL T-score	49.0	45.0–52.5	0.57
Mullen EL T-score	44.5	41.0–48.0	0.0023
Mullen VR T-score	42.5	38.5–46.5	0.0015
Mullen FM T-score	42.0	38.5–45.0	0.000062
Mullen ELC standard score	90.5	85.0–96.0	0.0016
MB-PROD percentile	27.5	20.0–40.0	0.00087
MB-M3L percentile	35.0	25.0–47.5	0.016
	Estimated group difference	95% CI	<i>p</i> value
Mullen RLt – VRt	+6.0	1.0–11.0	0.018
Mullen RLt – ELCt	+5.0	0.67–10.0	0.037
Mullen RLt – ELt	+4.0	0–9.0	0.083

Bolded entries indicate $p < 0.05$ for significant group difference below testing norms. T-scores are normed against typical validation groups with a mean score of 50 (SD 10). Standard scores are normed against typical validation groups with a mean score of 100 (SD 15). ELt, expressive language T-score; RLt, receptive language T-score; VRt, visual reception T-score; FM, fine motor; ELC, Early Learning Composite; MB, MacArthur-Bates Communicative Development Inventories; PROD, overall volume of word production; M3L, mean length of 3 longest utterances.

Statistical Analysis

Overall cohort outcomes were assessed; statistics were performed using R 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria), and figures were constructed using MATLAB R2018b (MathWorks, Natick, MA, USA). Distributions of scores for each subtest were described and compared to typical average scores for age using a nonparametric Wilcoxon signed-rank test so as not to assume normally distributed data. Developmental dissociation was assessed between subscores within the tested cohort using the Wilcoxon rank-sum test.

To better define motor contributions to language scores, post hoc analysis was performed on (1) subjects with a clinical diagnosis of CP, and (2) all subjects. As only 3 subjects in this convenience sample were diagnosed with CP, language scores were described narratively. To identify motor and cognitive contributions more quantitatively, a 2-way type II ANOVA model was constructed using MSEL-VR (as a proxy for nonverbal intelligence) and MSEL-FM (a measure of fine-motor performance) as predictors of MSEL-EL and MSEL-RL.

To investigate the value of potential prognostic factors available in the NICU setting to later language ability, we examined the covariance of selected expressive and receptive language subscores (MSEL-EL, MSEL-RL, and MB-PROD) with gender, geocode-based estimated income, initial pH and base excess, modified Sarnat scores, and binarized normal/abnormal MRI. Data from additional variables of interest (reported household income, level of maternal education, and lactate) were only available from a few subjects. Cohort-wide values were summarized (Table 1) but these variables were not included in further analysis.

We examined covariance of selected subscores (MSEL-EL, MSEL-RL, MSEL-VR, MB-PROD, MB-M3L) with respect to 2×2 covariates of the presence versus absence of cortical/subcortical lesions and basal ganglia/thalamic lesions, respectively. Determi-

nation of anatomical location was based on NICHD MRI criteria: grades 1A and 1B were considered to have cortical/subcortical involvement only; grade 2A was considered to have basal ganglia/thalamic involvement only; and grades 2B and 3 were considered to have both types of involvement.

Given the sample size, analysis was first performed in a univariate manner on subjects with data including the variable of interest. Again, nonparametric statistics were used; a Wilcoxon rank-sum test was used for binary covariates, and a significance test on Kendall's tau correlation versus a null hypothesis of no effect was used for continuous covariates. The 2×2 model based on factors of the presence/absence of (1) cortical/subcortical (MRI-CORT) and (2) basal ganglia/thalamic lesions (MRI-BGT) was analyzed using a two-way type II ANOVA structure; significance was determined based on an F test. To better understand the relative contributions of variables studied, we also constructed a 4-way type II ANOVA model including all variables that showed significant associations on univariate analyses (MRI-CORT, MRI-BGT, gender, insurance status, initial pH, and initial base excess).

Results

Sample Characteristics

127 infants were treated with TH during the study period; 14 of these had died at the time of follow-up. A subset of 45 of children participated in a comprehensive research evaluation at around 2 years of age. Participants' characteristics are summarized in Table 1.

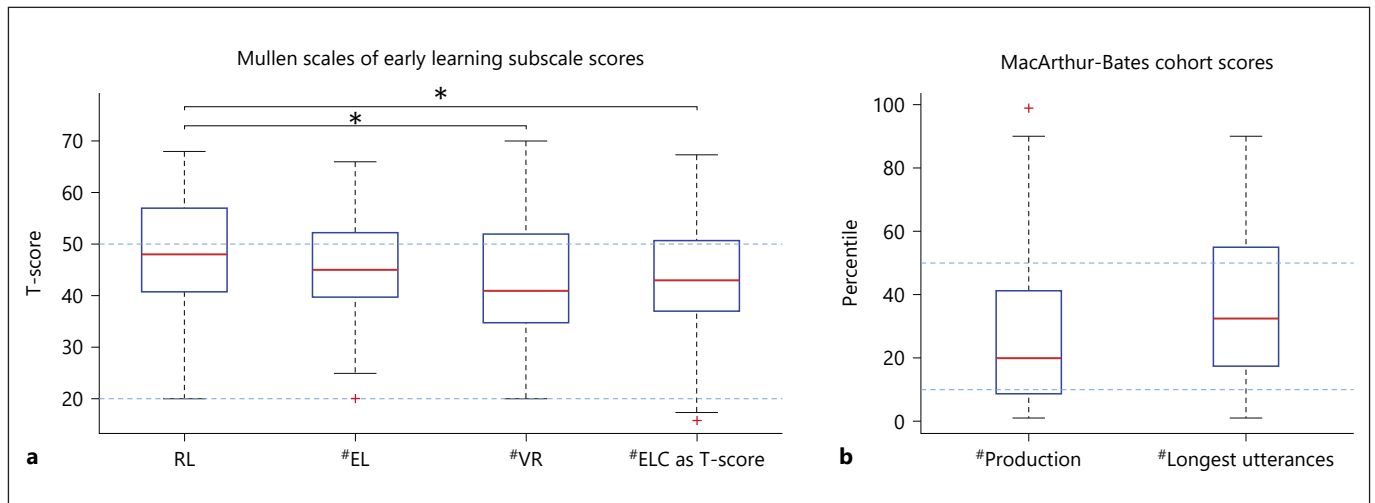


Fig. 1. Whole-cohort testing characteristics. **a** Mullen Scales. **b** MacArthur-Bates. # Significant group difference below testing norms ($p < 0.05$). * Significant discrepancy in performance between domains ($p < 0.05$). T-scores are normed against typical validation groups with a mean score of 50 (SD 10). RL, receptive language; EL, expressive language; VR, visual reception; ELC, early learning composite.

Table 3. Cognitive and motor determinants of language function

	MSEL expressive language T-score	MSEL receptive language T-score
MSEL-VR	27.2%***	23.1%***
MSEL-FM	10.0%*	2.5%
MSEL-VR \times MSEL-FM interaction effect	1.7%	5.4%
Total explained variance	38.9%	31.0%

Variance of each outcome measure by each predictor explained within a 2-factor ANOVA based on proxies of nonverbal cognitive skills and motor function. The third row represents the interaction term between the 2 factors as a predictor. Bolded entries indicate $p < 0.05$ by F test. * $p < 0.05$; *** $p < 0.001$. MSEL, Mullen Scales of Early Learning; VR, visual reception T-score; FM, fine motor T-score.

Whole-Cohort Language and Cognitive Outcomes

Whole-cohort testing characteristics are summarized in Table 2. MSEL-RL scores were not significantly different from those of the test norms (T-score estimate 49.0, confidence interval [CI] 45–52.5, $p = 0.57$), but MSEL-EL (44.5, CI 41–48, $p = 0.0023$), MSEL-VR (42.5, CI 38.5–46.5, $p = 0.0015$), and MSEL-ELC (SS = 90.5, CI 85–96, $p = 0.0016$) were all significantly lower than test norms (Fig. 1a, indicated by *). 4/44 subjects (9.1%) had MSEL-ELC standard scores < 70 ; 3/45 (6.7%) and 4/45 (8.9%) had MSEL-RL and MSEL-EL T-scores < 30 , respectively (2 SD below the population mean in all cases). MB-PROD and MB-M3L measures were both significantly lower than test norms (Fig. 1b; estimate 27.5 percentile, CI 20–40 percentiles, $p = 0.00087$, and estimate 35 percentile, CI

25–47.5 percentiles, $p = 0.016$, respectively). 24% of subjects (MB-PROD) and 16% of subjects (MB-M3L), respectively, had scores < 10 th percentile for age and gender, representing the classification “high-risk.” MSEL-RL scores were significantly higher than MSEL-VR scores ($p = 0.018$) and the MSEL-ELC ($p = 0.037$) but nonsignificantly higher than MSEL-EL scores ($p = 0.083$; Fig. 1a).

Covariate Analysis

All 3 children with CP had basal ganglia/thalamic involvement: 1 had MRI grade 2A and the other 2 had MRI grade 2B. Those with grade 2B on MRI had variable motor functioning: 1 had Gross Motor Functional Classification System (GMFCS) I and the other had GMFCS V, and very poor language outcomes (MSEL-RL and MSEL-EL

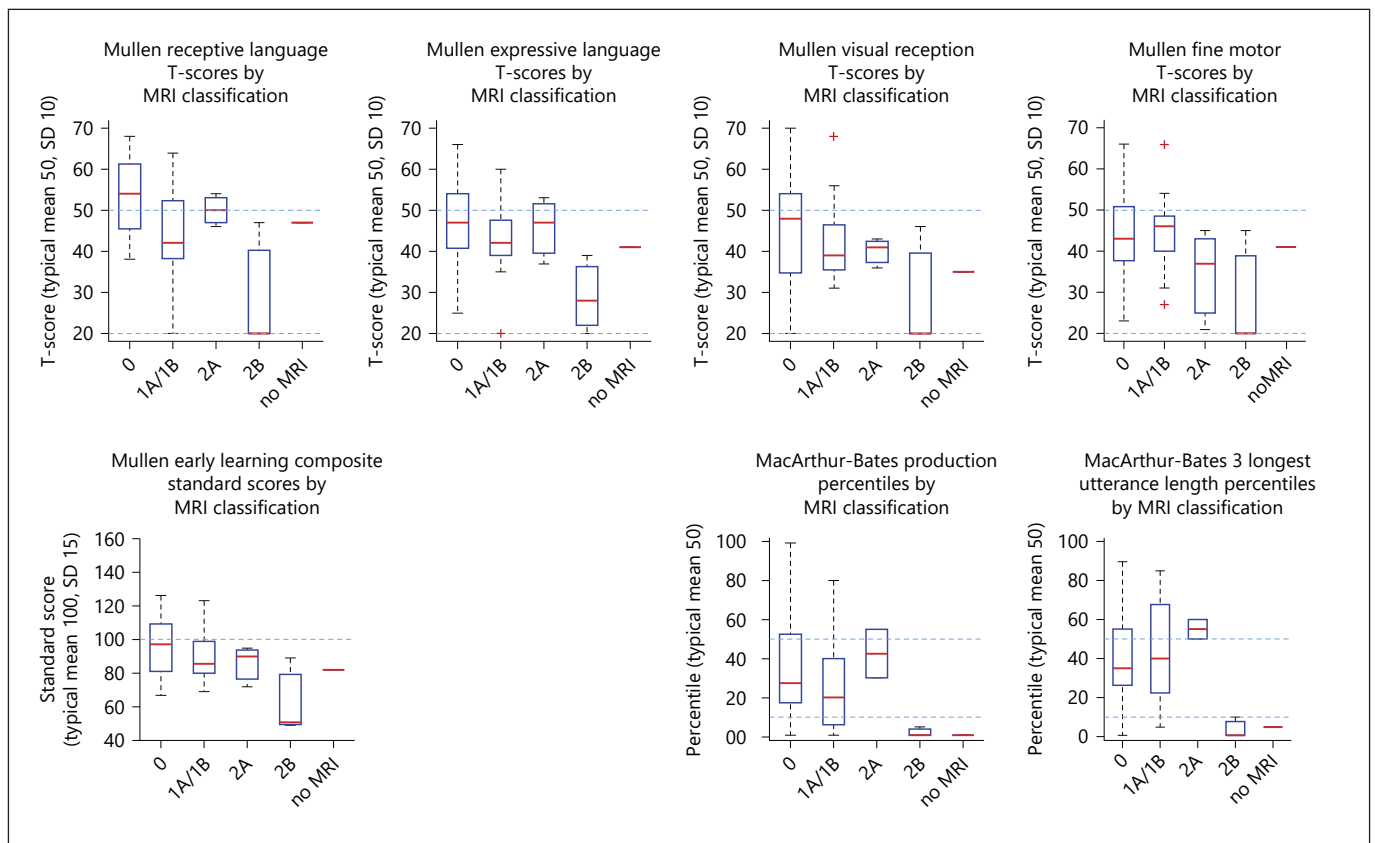


Fig. 2. Assessment outcomes by modified NICHD MRI grade. NICHD MRI grade reflects both the degree and gross distribution of brain injury on conventional MRI. Grades include categories for normal MRI (0), cortical/subcortical involvement only (1A [minimal cerebral lesions] and 1B [more extensive cerebral lesions]), basal ganglia/thalamic/deep white-matter involvement only (2A

[basal ganglia, thalamic, internal capsule lesions only]), or involvement of both areas (2B [basal ganglia, thalamic, internal capsule, and cerebral lesions] and 3 [cerebral devastation]). T-scores are normed against typical validation groups with a mean score of 50 (upper dotted line) and SD 10; a T-score of 20 represents the floor of the instrument for individual subscores (lower dotted line).

T-scores <25). The child with grade 2A on MRI had GM-FCS II and borderline language outcome (MSEL-RL T-score 46 and MSEL-EL T-score 37).

Variance in MSEL-VR (a proxy for nonverbal cognitive skills) predicted a significant minority of variance in both MSEL-EL (27.2% of variance; $p = 0.00014$) and MSEL-RL (23.1% of variance; $p = 0.00074$). Variance in MSEL-FM contributed significantly only to variance in MSEL-EL (10.0% of variance; $p = 0.014$; Table 3).

Univariate evaluation of effects of covariates is summarized in Table 4. There were gender effects in MSEL-EL subscores (estimated T-score 7.0 points more in females than males; $p = 0.024$) but not in MSEL-RL subscores. Estimated household income predicted MSEL-EL and MSEL-RL deficits. pH and base excess did not predict MSEL-EL, MSEL-RL, or MB-PROD. Modified Sarnat scores did not predict deficits. Binarized normal versus abnormal MRI

(NICHD score > 0) showed significant effects on both the expressive language (estimated difference in T-scores 7.0 points; $p = 0.038$) and receptive language (estimated difference in T-scores 10.1 points; $p = 0.0061$) subscores.

In the 2-factor anatomical injury model (Fig. 2; Table 5), only the presence of cortical/subcortical lesions showed a significant effect on receptive language subscores ($p = 0.0023$), but neither showed a significant effect on the expressive language or visual reception subscore. There was no significant interaction effect between the 2 factors. This model explained a total of 28.4% of variance in MSEL-RL, versus 17.8% in MSEL-EL and 10.6% in MSEL-VR. The same model also explained a total of 10.9% of variance in MB-PROD and 18.9% in MB-M3L; in the latter, it was predominantly the interaction term (the presence of both cortical/subcortical and basal ganglia/thalamic lesions) that predicted deficits ($p = 0.036$).

Table 4. Effects of NICU covariates

	Estimated group difference	95% CI	<i>p</i> value
Mullen EL FM	+7.0	1.0–14.0	0.024
MB-PROD percentile FM	+15.0	–9.5 to 35.0	0.063
Mullen RL FM	+4.0	–3.0 to 11.0	0.32
Mullen EL	+4.0	–3.0 to 12.0	0.23
ModSarnat 2 – ModSarnat 3			
MB-PROD percentile	+0.0	–19.0 to 25.0	0.85
ModSarnat 2 – ModSarnat 3			
Mullen RL	+3.0	–5.0 to 11.0	0.42
ModSarnat 2 – ModSarnat 3			
Mullen EL NI – Abnl MRI	+7.0	0.0–14.0	0.038
Mullen RL NI – Abnl MRI	+10.1	3.0–19.0	0.0061
Mullen VR NI – Abnl MRI	+7.0	–2.0 to 15.0	0.10
Mullen FM NI – Abnl MRI	+1.0	–6.0 to 9.0	0.89
MB-PROD percentile NI – Abnl MRI	+15.0	0.0 to 34.0	0.070
MB-M3L percentile NI – Abnl MRI	+11.6	–10.0 to 30.0	0.28
	Estimated Kendall's rank correlation coefficient (τ)		<i>p</i> value
Mullen EL vs. estimated income	0.305		0.0035
MB-PROD percentile vs. estimated income	0.180		0.15
Mullen RL vs. estimated income	0.250		0.017
Mullen EL vs. pH	–0.045		0.67
MB-PROD percentile vs. pH	0.10		0.43
Mullen RL vs. pH	–0.0010		0.99
Mullen EL vs. BE	0.11		0.33
MB-PROD percentile vs. BE	–0.027		0.84
Mullen RL vs. BE	0.091		0.40

Bolded entries indicate $p < 0.05$ for significant between-group differences. T-scores are normed against typical validation groups with a mean score of 50 (SD 10). EL, expressive language T-score; RL, receptive language T-score; VR, visual reception T-score; FM, fine motor T-score; MB, MacArthur-Bates Communicative Development Inventories; M3L, mean of 3 longest utterances percentile; PROD, overall volume of word production; ModSarnat, Modified Sarnat score; NI, normal; Abnl, abnormal; BE, base excess.

Table 5. Outcomes by pattern of injury on MRI

	MSEL-EL	MSEL-RL	MSEL-VR	MB-PROD	MB-M3L
Presence of cortical/subcortical injury	8.1%	18.5%**	3.3%	4.3%	1.3%
Presence of basal ganglia/thalamic injury	5.7%	6.5%	6.3%	1.6%	3.4%
(Cortical/subcortical injury) \times (basal ganglia/thalamic injury) interaction effect	4.0%	3.5%	1.1%	4.9%	14.1%*
Total explained variance	17.8%	28.4%	10.6%	10.9%	18.9%

Variance of each outcome measure by each predictor is explained within a 2-factor ANOVA based on NICHD MRI scoring, i.e., grade 1A/1B (cortical/subcortical injury only), grade 2A (basal ganglia/thalamic injury only), and grades 2B and 3 (both types of injury). The third row represents the interaction term between the 2 factors as a predictor. Bolded entries indicate $p < 0.05$ by F test. * $p < 0.05$; ** $p < 0.01$. MSEL, Mullen Scales of Early Learning; MB, MacArthur-Bates Communicative Development Inventories; EL, expressive language T-score; RL, receptive language T-score; VR, visual reception T-score; PROD, overall volume of word production; M3L, mean length of 3 longest utterances.

Table 6. 4-factor ANOVA including all significant covariates

	MSEL-EL	MSEL-RL	MSEL-VR	MB-PROD	MB-M3L
MRICort	4.0%	13.7%**	0.9%	0.4%	0.1%
MRIBgt	1.0%	2.1%	2.6%	0.0%	0.1%
Gender	10.5%*	1.1%	4.4%	9.6%	0.7%
TractEstIncome	16.4%**	12.5%*	19.1%*	1.0%	5.4%
MRIBgt:Gender	1.1%	0.9%	0.0%	0.0%	0.0%
MRIBgt:MRICort	0.3%	0.0%	0.1%	2.4%	9.9%
Gender:MRICort	0.6%	2.3%	0.6%	8.5%	0.5%
MRIBgt:TractEstIncome	0.7%	0.9%	0.6%	0.2%	3.4%
Gender:TractEstIncome	0.0%	0.0%	4.0%	0.5%	0.5%
MRICort:TractEstIncome	1.1%	2.4%	0.0%	3.0%	8.0%
MRIBgt:Gender:MRICort	0.0%	0.0%	0.0%	0.0%	0.0%
MRIBgt:Gender:TractEstIncome	0.0%	0.0%	0.0%	0.0%	0.0%
MRIBgt:MRICort:TractEstIncome	1.0%	0.5%	1.2%	0.0%	0.0%
Gender:MRICort:TractEstIncome	0.7%	4.9%	0.2%	0.6%	2.9%
MRIBgt:Gender:MRICort:TractEstIncome	0.0%	0.0%	0.0%	0.0%	0.0%
Total explained variance	37.4%	41.4%	33.8%	26.3%	31.5%

Variance of each outcome measure by each predictor explained within the 4-factor model. Rows with multiple predictors separated by colons represent interaction term predictors. Bolded entries indicate $p < 0.05$ by F test. * $p < 0.05$; ** $p < 0.01$. MSEL, Mullen Scales of Early Learning; EL, expressive language T-score; RL, receptive language T-score; VR, visual reception T-score; MB, MacArthur-Bates Communicative Development Inventories; PROD, overall volume of word production; M3L, mean length of 3 longest utterances; MRIBgt, the presence of basal ganglia/thalamic injury on MRI; MRICort, the presence of cortical/subcortical injury on MRI; TractEstIncome, household income as estimated based on the home census tract.

In the 4-factor ANOVA model (Table 6), total explained variance ranged between 26.3 (MB-PROD) and 41.4% (MSEL-RL). Estimated household income was the most consistent predictor with a significant impact on MSEL-EL ($p = 0.0068$), MSEL-RL ($p = 0.014$), and MSEL-VR ($p = 0.0047$). Gender also contributed significantly to MSEL-EL ($p = 0.027$), and MRI-CORT to MSEL-RL ($p = 0.010$).

Discussion

Mean performance in language measures fell within the normal range, and most individuals performed within the normal range for cognitive and language measures. While receptive language scores on MSEL were not significantly different from existing normative samples, the expressive language and visual reception subscales were significantly lower, as were the composite scores. These findings are consistent with previous studies, i.e., that as a group, children with perinatal HIE treated with TH perform in the average to low-to-average range in some language measures. This study demonstrates that patterns of strengths and weakness previously demonstrated at the age of 4–7 years are evident

already at 2 years of age [2, 13]. Receptive language at 2 years was most preserved when compared to expressive language and a nonverbal visual reception task. While expressive language measures fell in the average range for the group, mean performance was significantly lower than the published normative samples in the performance-based measure, MSEL, indicating a larger number of children than would be predicted with low-to-average or below-average expressive and nonverbal skills. Additionally, parent report of expressive one-word vocabulary on the MB-CDI was consistent with the children's performance on MSEL, with 24% of the study group classified as high-risk. While language performance at the age of 2 years is variable and children with isolated expressive language delays have a good prognosis, early language disability can be an indicator for neurodevelopmental problems and should be monitored in a high-risk population such as children with a history of perinatal HIE. There is evidence of good concurrent and predictive validity of parent-report measures such as the MB-CDI in this age range [25]. Furthermore, there is growing evidence that while “late talkers” (defined as children between the ages of 18–35 months who acquire language at a slower rate than their typically developing peers) may catch up in terms of vo-

cabulary, language-based learning difficulties re-emerge as reading and writing difficulties at school age [26–29]. Pre-TH outcomes have found these very areas of difficulty in children with perinatal HIE [3, 5].

Variation in general cognitive factors, and, in the case of expressive language, motor skills, appears to explain a portion of the variation in preschool language skills. However, large portions of variation in language functioning also remain independent of these factors, highlighting the need for serial developmental monitoring across domains in this high-risk population.

As expected, there were intervening gender and environmental effects on language outcomes following perinatal HIE. Girls had better performance than boys on the expressive portions of MSEL, but there were no receptive differences. As the MB-CDI normative percentiles are based on gender and age, gender was already accounted for in these measures. Proxies of socioeconomic status predicted differences in both expressive and receptive portions of the MSEL. This finding is not unexpected, given the large body of literature demonstrating socioeconomic effects on language production; however, it reinforces the idea that additional exposures, especially in low-income households, are important for language learning. This is particularly important in a vulnerable population, such as children with a history of HIE, and it provides some evidence for early intervention.

Consistent with previous studies, while neither modified Sarnat score nor pH/BE independently correlated with the 2-year language outcomes, injury as demonstrated on early MRI did relate to outcome. Specifically, children with normal early MRI had significantly higher expressive and receptive language scores than those with any MRI abnormalities. As expected, children with the most extensive injury patterns had lower scores for all measures. The presence of cortical/subcortical lesions showed effects on the receptive language subscore, but neither cortical/subcortical nor basal ganglia/thalamic types of injury showed significant effects on the expressive language or visual reception subscores.

Limitations

This study was a convenience sample and was not fully powered for multivariate analysis. We did not intentionally select for participants based on injury characteristics or socioeconomic factors, but inclusion only of individuals maintaining research follow-up may have biased our sample.

As with any preschool outcome, caution should be used when interpreting the relationship between perfor-

mance in any specific domain and later functional outcomes. This is especially true for language, which is rapidly developing at this age.

High-grade injury (according to encephalopathy score or MRI) was not equally represented in this sample, and our findings apply most directly to individuals without devastating injury on MRI (NICHD grade 3). That said, the percentage of normal (NICHD grade 0) MRI scores in our cohort was similar to that in prior studies (e.g., 57 vs. 52% in [24]), and the data collected reflect the developmental status of individuals who sustained a broad spectrum of hypoxic-ischemic brain injury patterns.

Implications

This cohort of children with perinatal HIE treated with TH had relatively preserved receptive language skills at the age of 2 years. Expressive language and visual reception were in the normal range but significantly lower than in normative samples. The MB-CDI has been validated in many populations, and it similarly appears to be appropriate to identify expressive language deficits in preschool children with a history of HIE. While blood gas variables and modified Sarnat score did not predict outcome, only a single subject in our study with a normal MRI demonstrated below-normal expressive language functioning (and none demonstrated below-normal receptive language functioning), allowing for some reassurance about language outcomes at the age of 2 years for children with a normal MRI. Furthermore, one of the major determinants of expressive and receptive language was estimated socioeconomic status, suggesting that much remaining risk following HIE treated with TH may be modifiable through environmental enrichment and early intervention. While portions of variance in language outcomes appear to be driven by variance in the motor and cognitive domains, significant unexplained variance remains. This highlights the need for close monitoring of all developmental domains in this high-risk population.

Acknowledgement

We acknowledge the members of the Johns Hopkins Neurosciences Intensive Care Nursery whose collaboration provided a platform for this research. We thank the participants of the 11th Hershey Conference on Developmental Brain Injury for their early feedback on our data. We are particularly grateful to the participants and their families, without whom this research would not have been possible.

Statement of Ethics

This study was performed following IRB approval. Parents/guardians of subjects provided written informed consent for all study procedures.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Sources

Vera Joanna Burton was supported by the NINDS/NIH K12-NS001696 during data collection for this project. Vera Joanna Burton, Gwyn Gerner, and Frances Northington were all supported in part by NIH/NICHD 1 R01 HD086058-01A1 during the preparation of the paper.

Author Contributions

Eric Chin participated in the analysis and writing of the manuscript. Srishti Jayakumar participated in the analysis and reviewed the manuscript. Ezequiel Ramos participated in the analysis and reviewed the manuscript. Gwendolyn Gerner participated in the conceptualization and data collection for the project and the writing of the manuscript. Bruno Soares reviewed and scored the MRIs and reviewed the manuscript. Elizabeth Cristofalo participated in the conceptualization and data collection for the project and reviewed the manuscript. Mary Leppert participated in the conceptualization and data collection for the project and reviewed the manuscript. Marilee Allen participated in the conceptualization and analysis and writing of the manuscript. Charla Parkinson participated in the conceptualization and data collection for the project and reviewed the manuscript. Michael Johnston participated in the conceptualization of the project and reviewed the manuscript. Frances Northington participated in the conceptualization and data collection for the project and reviewed the manuscript. Vera Joanna Burton participated in the conceptualization and data collection for the project and the analysis and writing of the manuscript.

References

- 1 Kurinczuk JJ, White-Koning M, Badawi N. Epidemiology of neonatal encephalopathy and hypoxic-ischaemic encephalopathy. *Early Hum Dev*. 2010 Jun;86(6):329–38.
- 2 Natarajan G, Pappas A, Shankaran S. Outcomes in childhood following therapeutic hypothermia for neonatal hypoxic-ischemic encephalopathy (HIE). *Semin Perinatol*. 2016 Dec;40(8):549–55.
- 3 D'Souza SW, McCartney E, Nolan M, Taylor IG. Hearing, speech, and language in survivors of severe perinatal asphyxia. *Arch Dis Child*. 1981 Apr;56(4):245–52.
- 4 Robertson C, Finer N. Term infants with hypoxic-ischemic encephalopathy: outcome at 3.5 years. *Dev Med Child Neurol*. 1985 Aug;27(4):473–84.
- 5 Marlow N, Rose AS, Rands CE, Draper ES. Neuropsychological and educational problems at school age associated with neonatal encephalopathy. *Arch Dis Child Fetal Neonatal Ed*. 2005 Sep;90(5):F380–7.
- 6 Wechsler D, editor (Tideman E, transl.). WPPSI-III: Wechsler Preschool and Primary Scale of Intelligence. 3rd ed. San Antonio (TX): Harcourt Assessment; 2005. pp. 1–164.
- 7 Wechsler D. WISC-IV: Wechsler Intelligence Scale for Children. 4th ed. Integrated: Technical and Interpretive Manual. San Antonio (TX): Harcourt Assessment; 2004.
- 8 Azzopardi D, Strohm B, Marlow N, Brocklehurst P, Deierl A, Eddama O, et al.; TOBY Study Group. Effects of hypothermia for perinatal asphyxia on childhood outcomes. *N Engl J Med*. 2014 Jul;371(2):140–9.
- 9 Pearson Education [Internet]. Clinical Evaluation of Language Fundamentals: Technical report: Correlations between the CELF-4 and WISC-4 Integrated [cited 2008]. Available from: https://images.pearsonclinical.com/images/resource/techrpts/CELF4_WISC4_TechReport.pdf
- 10 Stage SA, Abbott RD, Jenkins JR, Berninger VW. Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: failure to validate discrepancy method. *J Learn Disabil*. 2003 Jan-Feb;36(1):24–33.
- 11 American Academy of Pediatrics. Neonatal Encephalopathy and Neurologic Outcome. 2nd Edition. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy. *Pediatrics*. 2014 May;133(5):e1482–8.
- 12 Allanson ER, Waqar T, White C, Tunçalp Ö, Dickinson JE. Umbilical lactate as a measure of acidosis and predictor of neonatal risk: a systematic review. *BJOG*. 2017 Mar;124(4):584–94.
- 13 Shapiro KA, Kim H, Mandelli ML, Rogers EE, Gano D, Ferriero DM, et al. Early changes in brain structure correlate with language outcomes in children with neonatal encephalopathy. *Neuroimage Clin*. 2017 Jun;15:572–80.
- 14 Steinman KJ, Gorno-Tempini ML, Glidden DV, Kramer JH, Miller SP, Barkovich AJ, et al. Neonatal watershed brain injury on magnetic resonance imaging correlates with verbal IQ at 4 years. *Pediatrics*. 2009 Mar;123(3):1025–30.
- 15 Washington JA, Craig HK. Performances of At-Risk, African American Preschoolers on the Peabody Picture Vocabulary Test-III. *Lang Speech Hear Serv Sch*. 1999 Jan;30(1):75–82.
- 16 Gutierrez-Clellen VF. Language diversity: Implications for assessment. In: Cole K, Dale P, Thal D, editors. *Advances in Assessment of Communication and Language*. Baltimore: Brookes; 1996. pp. 29–56.
- 17 Whitehurst GJ. Language processes in context: Language learning in children reared in poverty. In: Adamson LB, Ronski MA, editors. *Communication and Language Acquisition: Discoveries from Atypical Development*. Baltimore: Brookes; 1997. pp. 233–65.
- 18 Martinez-Biarge M, Diez-Sebastian J, Rutherford MA, Cowan FM. Outcomes after central grey matter injury in term perinatal hypoxic-ischaemic encephalopathy. *Early Hum Dev*. 2010 Nov;86(11):675–82.
- 19 Martinez-Biarge M, Bregant T, Wusthoff CJ, Chew AT, Diez-Sebastian J, Rutherford MA, et al. White matter and cortical injury in hypoxic-ischemic encephalopathy: antecedent factors and 2-year outcome. *J Pediatr*. 2012 Nov;161(5):799–807.
- 20 Sarnat HB, Sarnat MS. Neonatal encephalopathy following fetal distress. A clinical and electroencephalographic study. *Arch Neurol*. 1976 Oct;33(10):696–705.
- 21 Shankaran S, Laptook AR, Ehrenkranz RA, Tyson JE, McDonald SA, Donovan EF, et al.; National Institute of Child Health and Human Development Neonatal Research Network. Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy. *N Engl J Med*. 2005 Oct;353(15):1574–84.

- 22 Thomas AJ, Eberly LE, Davey Smith G, Neaton JD; Multiple Risk Factor Intervention Trial (MRFIT) Research Group. ZIP-code-based versus tract-based income measures as long-term risk-adjusted mortality predictors. *Am J Epidemiol*. 2006 Sep;164(6):586–90.
- 23 Diez-Roux AV, Kiefe CI, Jacobs DR Jr, Haan M, Jackson SA, Nieto FJ, et al. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Ann Epidemiol*. 2001 Aug;11(6):395–405.
- 24 Shankaran S, McDonald SA, Laptook AR, Hintz SR, Barnes PD, Das A, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Neonatal Magnetic Resonance Imaging Pattern of Brain Injury as a Biomarker of Childhood Outcomes following a Trial of Hypothermia for Neonatal Hypoxic-Ischemic Encephalopathy. *J Pediatr*. 2015 Nov;167(5):987–93.e3.
- 25 Feldman HM, Dale PS, Campbell TF, Colborn DK, Kurs-Lasky M, Rockette HE, et al. Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Dev*. 2005 Jul-Aug;76(4):856–68.
- 26 Snowling MJ, Adams JW, Bishop DV, Stothard SE. Educational attainments of school leavers with a preschool history of speech-language impairments. *Int J Lang Commun Disord*. 2001 Apr-Jun;36(2):173–83.
- 27 Rescorla L. Language and reading outcomes to age 9 in late-talking toddlers. *J Speech Lang Hear Res*. 2002 Apr;45(2):360–71.
- 28 Rice ML, Taylor CL, Zubrick SR. Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *J Speech Lang Hear Res*. 2008 Apr;51(2):394–407.
- 29 Preston JL, Frost SJ, Mencl WE, Fulbright RK, Landi N, Grigorenko E, et al. Early and late talkers: school-age language, literacy and neurological differences. *Brain*. 2010 Aug;133(Pt 8):2185–95.