NEURODEVELOPMENTAL OUTCOMES OF EXTREMELY LOW GESTATIONAL AGE NEONATES WITH LOW GRADE INTRAVENTRICULAR-PERIVENTRICULAR HEMORRHAGE

By

ALLISON H. PAYNE, MD

Submitted in partial fulfillment of the requirements for the degree of Master of Science

Thesis Advisors: Deanne Wilson-Costello, MD
Douglas Einstadter, MD

Clinical Research Scholars Program
CASE WESTERN RESERVE UNIVERSITY

January 2012
We hereby approve the thesis/dissertation of

Allison Payne, M.D.

candidate for the M.S. in Clinical Research degree *.

(signed) Deanne Wilson-Costello, M.D.
(chair of the committee)

Michele C. Walsh, M.D., M.S.

Douglas Einstadter, M.D., M.S.

October 4, 2011

*We also certify that written approval has been obtained for any proprietary material contained therein.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>5</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>6</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>7</td>
</tr>
<tr>
<td>Abstract</td>
<td>9</td>
</tr>
<tr>
<td>Background</td>
<td>10</td>
</tr>
<tr>
<td>Methods</td>
<td>12</td>
</tr>
<tr>
<td>Results</td>
<td>19</td>
</tr>
<tr>
<td>Discussion</td>
<td>29</td>
</tr>
<tr>
<td>Bibliography</td>
<td>38</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIGURE 1: Cohort Enrollment Flow Diagram ................................................................. 20

FIGURE 2: Comparison of Adjusted Odds Ratios for Categorical Neurodevelopmental Outcomes ........................................................................................................... 26
LIST OF TABLES

TABLE 1: Population Demographics................................................................. 22

TABLE 2: Unadjusted Neurodevelopmental Outcomes by IV-PVH Grade............. 24

TABLE 3: Adjusted Neurodevelopmental Outcomes by IV-PVH Grade .................. 25

TABLE 4: Independent Predictors of Adjusted Neurodevelopmental Outcomes. ....... 28
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following for their support and encouragement throughout the duration of this project and the Master’s program:

- Dr. Deanne Wilson-Costello: Mentor and Thesis Committee Chair
- Drs. Michele Walsh and Anna Maria Hibbs: Thesis Committee members
- Dr. Douglas Einstadter: CRSP Advisor and Thesis Committee member
- Drs. Susan Hintz and Betty Vohr: NICHD Neonatal Research Network Collaborators
- Carla Bann, PhD: Statistician, RTI International and the Neonatal Research Network
- Dr. Richard Martin: Principal Investigator, NIH 5-T32-HD-060537 -- “Training in Neonatal Research” (Tuition Support)
- Rainbow Fellows Research Award in Pediatrics: Tuition Support
- The NICHD Neonatal Research Network for permitting me access to their valuable databases and the NRN research nurses for their tireless collection of this data.
- All of my patients and their families who motivate me every day to strive for clinical and research excellence.
LIST OF ABBREVIATIONS

IV-PVH: Intraventricular-periventricular hemorrhage

VLBW: Very low birth weight (<1500g)

NICHD: National Institute of Child Health and Development

NRN: Neonatal Research Network

ELGA: Extremely low gestational age (<27-28 weeks gestation)

GDB: Generic Database

ANS: Antenatal steroids

CPR: Cardiopulmonary resuscitation

BPD: Bronchopulmonary dysplasia

CPAP: Continuous positive airway pressure

PDA: Patent ductus arteriosus

PVL: Periventricular leukomalacia

NEC: Necrotizing enterocolitis

GMFCS: Gross Motor Function Classification System

CP: Cerebral palsy

NDI: Neurodevelopmental impairment

SD: Standard deviation

EPIPAGE: Etude Epidémiologique sur les Petits Ages Gestationnels

ELBW: Extremely low birth weight (<1000g)

NIH: National Institutes of Health
MDI: Mental Developmental Index

MRI: Magnetic resonance imaging
Neurodevelopmental Outcomes of Extremely Low Gestational Age Neonates
With Low Grade Intraventricular-Periventricular Hemorrhage

Abstract

by

ALLISON H. PAYNE, MD

Low grade intraventricular-periventricular hemorrhage (IV-PVH) is a common neurologic morbidity among extremely low gestational age (ELGA, < 27 weeks gestation) neonates yet the outcomes associated with this morbidity are poorly understood. We sought to determine the effect of low grade IV-PVH on neurodevelopmental outcomes at 18-22 months corrected age using a three year cohort of a large multicenter database. Mixed effects regression modeling was utilized to account for potential confounders including gestational age, gender, race/ethnicity, chorioamnionitis, antenatal and/or postnatal steroids, high frequency ventilation, sepsis, and patent ductus arteriosus, as well as center differences. There were no significant differences in unadjusted or adjusted motor, cognitive, language, or composite mild or moderate-severe neurodevelopmental impairment outcomes when comparing ELGA infants with low grade IV-PVH to ELGA infants without evidence of IV-PVH on cranial ultrasound.
BACKGROUND

Intraventricular-periventricular hemorrhage (IV-PVH) is the most common form of neonatal intracranial hemorrhage among premature infants.\(^1\) While IV-PVH is one of the most common morbidities in this population, neonatologists continue to face uncertainty when attempting to relate the importance of low grade hemorrhages to parents of affected infants.

The incidence of IV-PVH among extremely premature infants varies depending on the era of evaluation and the definition of the population in the denominator. Rates of IV-PVH among very low birth weight (VLBW, < 1500 g) infants were reported as high as 40-50% in the 1980s and as low as 15% in the 1990s.\(^1\)\(^-\)\(^3\) Focusing on neonates < 29 weeks gestation, Stoll et al\(^4\) recently reported on a 2003-2007 NICHD Neonatal Research Network (NRN) cohort with an incidence for any degree of IV-PVH of approximately 32%.

Grading of IV-PVH has traditionally been based on the four tiered Papile classification system.\(^5\) Grade 1 refers to hemorrhage isolated to the area of the germinal matrix. Grade 2 refers to hemorrhage extending into the ventricle without ventricular dilation. Grade 3 hemorrhages extend into the ventricle with the added complication of ventricular dilation. Grade 4 hemorrhage is recognized as having parenchymal or periventricular involvement. Debate regarding the appropriate nomenclature for IV-PVH exists but the traditional grading schema remains pervasive in both the literature and the clinical setting.\(^6\)\(^-\)\(^7\) In the context of this paper, “low grade”
will refer to the presence of Grade 1 or 2 IV-PVH and “severe” or “high grade” will refer to the presence of Grade 3 or 4 IV-PVH.

Developmental biology may suggest that hemorrhages of any grade for extremely low gestational age (ELGA, <27-28 weeks gestation) infants have the potential to destroy glial precursor cells. The germinal matrix is the source of neuronal precursors through the 20th week of gestation but is devoted to the development and proliferation of glial precursors to oligodendrocytes and astrocytes during the early third trimester.\(^1\) Interruption of oligodendrocyte development may result in abnormal myelination while interruption of astrocyte development and migration may affect the organization of the cerebral cortex. This loss of organization has been linked to reduction in cortical gray matter volume.\(^8\) Destruction of either type of precursors could therefore theoretically result in adverse neurodevelopmental outcomes.

While severe IV-PVH is well understood to correlate strongly with adverse motor and cognitive outcomes, the outcomes of low grade hemorrhages are not as fully understood despite accounting for 50-80% of all IV-PVH cases, particularly among survivors.\(^3,4,9\) A brief search of the literature will reveal that most neurodevelopmental outcome studies include or comment only on severe IV-PVH as a predictor of outcome. Despite the lack of clarity, it is widely believed that low grade IV-PVH has minimal impact and that neurologic and developmental outcomes are comparable to infants without evidence of hemorrhage.

Previous studies attempting to address the subject of low grade hemorrhage outcomes are inconsistent in their conclusions. Some studies fully support\(^{10-13}\) the
assumption of no difference in outcomes while others span a gamut from increased risk of mild delays in specific domains5, 14, 15, moderate-severe delays in specific domains16-18, or more “global” delays.19-23 Differences in the results of these studies may in some degree be attributable to study design, definition of the cohort of interest, evaluation methods, and the evolution of practice management and survival within the ELGA population.

In this study, we attempt to characterize the outcomes associated with low grade IV-PVH among ELGA infants with comparison both to a control group of ELGA infants without evidence of hemorrhage and to infants with severe IV-PVH for clinical context. To address the issues outlined above, we use a large, contemporary multicenter cohort over three years without known significant practice changes and with stable evaluation methods. We use mixed effects regression modeling to account for potential confounders including antenatal and postnatal steroids as well as center differences. We hypothesize that ELGA infants with low grade IV-PVH are at increased risk of cognitive impairment compared to those without hemorrhage but at lower risk compared to infants with severe IV-PVH.

METHODS:

Population

The cohort for this study includes ELGA infants with documented cranial ultrasound results born at less than 27 weeks gestation (≤ 26 6/7 weeks) in the 16-
center NICHD Neonatal Research Network (NRN) between January 1, 2006 and December 31, 2008 and surviving to complete follow-up assessment at 18-22 months corrected age. Additionally, subjects were excluded if there was known major congenital anomaly, meningitis, hydrocephalus requiring shunt, or porencephalic cyst evident on cranial ultrasound prior to 28 days of life. Each of these exclusion criteria has been independently associated with poor neurodevelopmental outcomes and was excluded to reduce the confounding of the outcomes in this study and allow as clear an interpretation of low grade hemorrhage as possible.

**Baseline Demographics and Characteristics**

Research nurses at each NRN institution prospectively gathered maternal, delivery, and neonatal course characteristics during the infant’s initial admission according the Generic Database (GDB) Manual of Operations.24

1. **Maternal Characteristics**

Maternal report of infant race for this study was collapsed to Black, Hispanic, White, or Other. Prolonged rupture of membranes was defined as rupture > 18 hours prior to delivery. Presence of chorioamnionitis was determined by clinical obstetric documentation. Maternal hypertension refers to obstetric documentation of any hypertension diagnosis during the pregnancy, acute or chronic. Antenatal steroid (ANS) exposure was defined as maternal receipt of any corticosteroid for the purpose of
accelerating fetal lung maturity. ANS course was considered to be complete if the mother received either two doses of betamethasone 12-24 hours apart or four doses of dexamethasone 6 hours apart and delivery occurred at least 24 hours from timing of the initial dose.

II. Neonatal Characteristics

Infant gestational age was determined by best obstetric estimate using last menstrual period or ultrasonography. Apgar scores and receipt of cardiopulmonary resuscitation (CPR) were drawn from the immediate delivery period with CPR defined as requirement for chest compressions or epinephrine. Respiratory distress refers to demonstration of clinical features within the first 24 hours of life or the requirement of oxygen or positive pressure ventilation for greater than 6 hours within the first 24 hours of life. Bronchopulmonary dysplasia (BPD) was defined by the physiologic definition (at 36 weeks post menstrual age: continued need for continuous positive airway pressure (CPAP) or ventilator support or inability to wean from supplemental oxygen with FiO2> 21% or cannula flow > 0.5 liters per minute if eligible for oxygen challenge). Postnatal steroids were defined as any doses or courses of systemic corticosteroids for the express purpose of prevention or treatment of BPD. Patent ductus arteriosus (PDA) was diagnosed by characteristic chest radiograph or features of left to right shunt on clinical or echocardiographic exam. PDA-Medical refers to any attempt to close the ductus arteriosus by pharmacologic means with either indomethacin or ibuprofen. PDA-Surgical refers to those cases where the PDA was closed by surgical ligation. Prophylaxis
indicates the receipt of indomethacin within the first 24 hours of life for the purpose of either IV-PVH prevention or empiric closure of the PDA. Infants were considered to have periventricular leukomalacia (PVL) if there was evidence of cystic lesions in the periventricular area on any cranial ultrasound during the neonatal admission. Sepsis refers to positive blood culture at any time during the neonatal admission. Necrotizing enterocolitis (NEC) in this setting refers to infants with Bell’s Staging Criteria of at least Stage IIA (proven, no surgery).²⁷

**Neurodevelopmental Assessment**

Follow-Up Study assessment at 18-22 months corrected age routinely includes medical history, parent interview, and a battery of neurologic, developmental, and behavioral tests. Neuromotor examinations utilized Amiel-Tison methods.²⁸ Gross motor function was assessed with Palisano’s Gross Motor Function Classification System (GMFCS)²⁹,³⁰ by certified examiners. Experienced and certified examiners administered the Bayley Scales of Infant Development, 3rd Edition³¹ for the entirety of this cohort. The Bayley III includes measures of cognitive and language development and the tests are standardized to a mean score of 100 and standard deviation of 15.

**Outcome Definitions**

Cerebral palsy (CP) is a nonprogressive disorder of the central nervous system. For the purposes of this study, CP was defined as having abnormal tone or reflexes in at
least one extremity *and* abnormal control of movement or posture to a degree that interferes with age-appropriate activity. A diagnosis of “Moderate to Severe CP” additionally required the infant to be non-ambulatory or in need of an assistive device for ambulation. GMFCS score ≥ 2 indicates gross motor functional limitation.

Severe visual impairment, blindness, is defined as having bilateral visual acuity < 20-200. Severe hearing loss is defined as bilateral permanent hearing loss that precludes a child from understanding the examiner or communicating despite amplification.

Traditionally, neurodevelopmental impairment (NDI) is a composite outcome defined as having any of the following: moderate-severe CP with GMFCS ≥ 2, severe visual impairment, severe hearing loss, or cognitive score < 70 (-2 SD). This study additionally includes a “mild NDI” outcome with an alternative cognitive score cut-off of < 85 (-1 SD) and all other composite components the same as the traditional definition.

The primary outcome for this study was the Bayley III cognitive score. Secondary outcomes include: Bayley III language score, a Bayley III cognitive score < 70, a Bayley III cognitive score < 85, a Bayley III language score < 70, a Bayley III language score < 85, any cerebral palsy, a GMFCS score ≥ 2, severe visual impairment, severe hearing loss, and NDI with cognitive score cut-offs of either < 70 (“moderate-severe NDI”) and < 85 (“mild NDI”).
Sample Size Estimate

Sample size estimates for this study were calculated based on two-sided t-testing of pair-wise comparisons for the continuous cognitive score as the primary outcome. Assuming a standard deviation in cognitive score for the ELGA population of 15 points, and aiming to detect a 5 point difference in group means with 80% power and $\alpha=0.05$, each group would need to contain at least 143 subjects. The least common denominator groups of Grade 1 and Grade 2 IV-PVH would then need at least 143 subjects each. This would subsequently yield a Low Grade group of 286 infants which for comparisons of Low Grade vs No IV-PVH would have more than adequate power for detecting differences in the continuous cognitive outcome. Further assuming a 30% incidence in IV-PVH, 70% of IV-PVH cases being low grade, equal distribution between grade 1 and grade 2 hemorrhages, 10% meeting exclusion criteria, 75% survival of eligible infants, 15% loss to follow-up, and an estimated 6% of subjects being unable to complete follow-up testing, a projected cohort of approximately 2900 infants meeting the initial inclusion criteria of < 27 weeks gestation with cranial ultrasound data within the first 28 days of life would be required. Based on typical rates of enrollment to the NRN databases, we estimated a three year birth cohort would be required.

Statistical Analyses

Bivariate analyses were conducted to compare demographic, neonatal, and maternal characteristics and unadjusted neurodevelopmental outcomes by IV-PVH
grade. Chi-squared tests were used for categorical variables and t-tests for continuous variables.

Two levels of three-group analyses were conducted with No IV-PVH infants within the birth cohort serving as controls and Grade 3/4 infants serving as comparators for clinical frame of reference. First, a comparison of No IV-PVH vs Low Grade (Grade 1/2) vs Severe (Grade 3/4) IV-PVH was performed. Second, a comparison within low grade IV-PVH was carried out: No IV-PVH vs Grade 1 vs Grade 2. Multivariate mixed effects regression modeling was performed to adjust for potential confounding factors. Linear regression modeling was used for continuous outcomes and logistic regression modeling for categorical outcomes. To preserve the largest possible sample, missing values for predictor variables, but not outcomes, were imputed as not having the exposure. Model covariates were determined a priori to address suspected confounders and modifiers including: IV-PVH grade (as a 3-level covariate), gestational age, gender, race/ethnicity, chorioamnionitis (clinical), sepsis (positive blood culture, early or late), antenatal steroids (partial or complete course), postnatal steroids, high frequency ventilation, and patent ductus arteriosus. NRN Center was included in all models as a random effect to control for site-to-site differences in clinical management and local cranial ultrasound readings. As the model was not intended to be predictive and each covariate was considered important to regard, the model was applied en bloc for each outcome without paring down procedures such as stepwise regression.
Since this model includes several covariates, a correlation matrix to assess potential multicollinearity was created. Multicollinearity refers to the statistical situation wherein at least two predictor variables are strongly correlated. If multicollinearity exists, the resulting coefficient estimates may be quite sensitive to small changes in the model or data. The correlation matrix for the specified model employed in this study indicated no significant correlations between any pairing of covariates, including with IV-PVH, with all $R^2$ values $< |0.3|$. 

RESULTS

2514 infants less than 27 weeks gestational age met initial inclusion criteria. (Figure 1) 202 infants were ineligible for the cohort by exclusion criteria. Of the 2312 eligible infants, 627 died before 18-22 month follow-up. Of the surviving infants, 178 infants were lost to follow-up and 35 infants had incomplete neurodevelopmental testing and were excluded to allow modeling without imputation of the outcome of interest. Therefore, 87% of eligible survivors completed follow-up. The final cohort consists of 1472 ELGA infants surviving to and completing follow-up at 18-22 months corrected age. Of the final cohort, 451 (30.6%) infants had some degree of IV-PVH: 140 with Grade 1, 130 with Grade 2, and 181 with Grades 3 or 4. Thus, 60% of IV-PVH among survivors could be characterized as “Low Grade”.
Baseline Characteristics

Demographics, maternal and neonatal course characteristics are shown in Table 1. There was no difference in mean gestational age for infants with low grade or no IV-PVH. Infants with high grade hemorrhage were significantly younger than those with low grade or no IV-PVH. Within low grade IV-PVH, infants with Grade 2 were slightly younger only when compared to those with no IV-PVH (24.9 vs. 25.1 weeks, p=0.02).
Infants with any grade IV-PVH were more likely to be male; however there was no
association between gender and severity of IV-PVH. Race/ethnicity were distributed
similarly between no IV-PVH, low grade, and high grade groups. However, within low
grade comparisons, the Grade 1 group had a larger proportion of black infants
compared to Grade 2 (45% vs. 32%, p=0.03) while Grade 2 has a larger proportion of
white infants compared to Grade 1 (41% vs. 29%, p=0.03). Infants with any grade IV-
PVH were less likely to be delivered by Cesarean section but there was no association
between mode of delivery and IV-PVH grade. Maternal hypertension during pregnancy
was less common with any grade of IV-PVH when comparing no IV-PVH to low grade or
high grade. Within the low grade comparisons, there were no significant differences,
but there was similar non-significant decreased incidence of exposure to maternal
hypertension for those with Grade 1 or Grade 2 hemorrhage. All groups were similar in
birth weight and maternal age, education, and marital status. Infants with Grade 3/4 IV-
PVH were more likely to have been exposed to clinical chorioamnionitis but less likely to
be exposed to prolonged rupture of membranes > 18 hours than infants with no IV-PVH
or low grade IV-PVH. Infants with low grade or no IV-PVH received full or partial courses
of antenatal steroids with similar frequency. However, infants with high grade
hemorrhages were less likely to have received full or partial antenatal steroids when
compared to no IV-PVH or low grade IV-PVH groups.

Neonatal course characteristics were very similar for infants with either no IV-
PVH or low grades of IV-PVH with two exceptions. First, infants with low grade
hemorrhage were more likely to have been exposed to high frequency ventilation. This
# Table 1. Population Demographics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No IV-PVH n = 1021</th>
<th>Low Grade n = 270</th>
<th>High Grade n = 181</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infant Characteristics</strong></td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>Gestational Age (wks, mean(sd))</td>
<td>25.1 (0.9)</td>
<td>25.0 (1)</td>
<td>24.7 (1)</td>
<td>b, c</td>
</tr>
<tr>
<td>Birth Weight (g, mean(sd))</td>
<td>769 (154)</td>
<td>769 (151)</td>
<td>749 (154)</td>
<td></td>
</tr>
<tr>
<td>Gender (Male)</td>
<td>47</td>
<td>61</td>
<td>57</td>
<td>a, b</td>
</tr>
<tr>
<td>Race (Black)</td>
<td>39</td>
<td>39</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td><strong>Maternal Characteristics</strong></td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>Age (yrs, mean(sd))</td>
<td>28 (6)</td>
<td>27 (7)</td>
<td>27 (6)</td>
<td></td>
</tr>
<tr>
<td>≥ High School Education</td>
<td>83</td>
<td>81</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>47</td>
<td>47</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>23</td>
<td>16</td>
<td>14</td>
<td>a, b</td>
</tr>
<tr>
<td>Prolonged ROM</td>
<td>28</td>
<td>28</td>
<td>19</td>
<td>b, c</td>
</tr>
<tr>
<td>Chorioamnionitis</td>
<td>17</td>
<td>21</td>
<td>29</td>
<td>b, c</td>
</tr>
<tr>
<td>Cesarean Section</td>
<td>68</td>
<td>58</td>
<td>57</td>
<td>a, b</td>
</tr>
<tr>
<td>Antenatal Steroids, Any</td>
<td>91</td>
<td>89</td>
<td>78</td>
<td>b, c</td>
</tr>
<tr>
<td>Antenatal Steroids, Full</td>
<td>61</td>
<td>57</td>
<td>44</td>
<td>b, c</td>
</tr>
<tr>
<td><strong>Neonatal Course</strong></td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>5-minute Apgar &lt; 6</td>
<td>21</td>
<td>26</td>
<td>31</td>
<td>b</td>
</tr>
<tr>
<td>CPR</td>
<td>9</td>
<td>8</td>
<td>14</td>
<td>b, c</td>
</tr>
<tr>
<td>Surfactant</td>
<td>88</td>
<td>90</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>High Frequency Ventilator</td>
<td>37</td>
<td>46</td>
<td>61</td>
<td>a, b, c</td>
</tr>
<tr>
<td>Pneumothorax</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Postnatal Steroids</td>
<td>14</td>
<td>14</td>
<td>22</td>
<td>b, c</td>
</tr>
<tr>
<td>CLD, physiologic</td>
<td>47</td>
<td>53</td>
<td>60</td>
<td>b</td>
</tr>
<tr>
<td>PDA - Medical</td>
<td>38</td>
<td>41</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>PDA - Surgical</td>
<td>17</td>
<td>20</td>
<td>29</td>
<td>b</td>
</tr>
<tr>
<td>IV-PVH Prophylaxis</td>
<td>44</td>
<td>38</td>
<td>35</td>
<td>b</td>
</tr>
<tr>
<td>PVL</td>
<td>2</td>
<td>5</td>
<td>14</td>
<td>a, b, c</td>
</tr>
<tr>
<td>Sepsis</td>
<td>39</td>
<td>44</td>
<td>49</td>
<td>b</td>
</tr>
<tr>
<td>NEC</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

ROM: rupture of membranes, CPR: cardiopulmonary resuscitation (chest compressions or epinephrine), CLD: chronic lung disease, PDA: patent ductus arteriosus, IV-PVH: intraventricular-periventricular hemorrhage prophylaxis (indomethacin or ibuprofen), PVL: periventricular leukomalacia, NEC: necrotizing enterocolitis (≥ Bell Stage IIA).

a: p < 0.05 for Grade 1/2 vs No IV-PVH, b: p < 0.05 for Grade 3/4 vs No IV-PVH, c: p < 0.05 for Grade 3/4 vs Grade 1/2
significant difference is largely related to a difference in exposure between the Grade 2 and no IV-PVH groups. Second, infants with low grade IV-PVH were more likely to have periventricular leukomalacia than infants without any IV-PVH (5% vs. 2%, p<0.01). This difference is seemingly driven by an increased incidence in the Grade 1 group compared to the no IV-PVH group (6% vs 2%, p<0.01). Of note, infants with either no IV-PVH or low grade hemorrhage had similar rates of postnatal steroid exposure, physiologic BPD, medical or surgical PDA intervention, exposure to IV-PVH prophylaxis, sepsis, and NEC.

Not surprisingly, infants with high grade IV-PVH were more likely to experience other neonatal morbidities including poor 5 minute Apgar score, requirement for CPR in the delivery room, BPD, high frequency ventilation, postnatal steroids, surgical ligation of a PDA, periventricular leukomalacia, and sepsis. Infants with high grade bleeds were also less likely to have received IV-PVH prophylaxis compared to infants without hemorrhage.

Unadjusted Neurodevelopmental Outcomes

Low grade IV-PVH was not associated with any significant increases in poor neurodevelopmental outcome compared to ELGA infants without evidence of IV-PVH (Table 2). However, continuous language score (83 vs 86, p=0.06) and the categorical outcome of Language < 85 (53% vs 45%, p=0.06) demonstrate a non-significant increase in risk of poor outcome for the low grade group.
Within the low grades, the continuous language score was significantly decreased for infants with Grade 1 (82±15) but not for infants with Grade 2 (85±16) when compared to control infants without evidence of IV-PVH (86±17, p=0.03 and p=0.56 respectively).

Infants with high grade IV-PVH demonstrated significantly higher rates for every marker of poor neurodevelopmental outcome when compared to infants without IV-PVH. In addition, infants with high grade IV-PVH had a significantly increase higher rate of poor outcome when compared to infants with low grade hemorrhage for every outcome except Language < 85.
Table 3. Adjusted Neurodevelopmental Outcomes by IV-PVH Grade

<table>
<thead>
<tr>
<th></th>
<th>Low Grade vs No IV-PVH</th>
<th>High Grade vs No IV-PVH</th>
<th>High vs Low Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any CP</td>
<td>1.00 (0.61, 1.64)</td>
<td>3.44 (2.24, 5.28)*</td>
<td>3.43 (1.96, 5.98)*</td>
</tr>
<tr>
<td>GMFCS &gt; 2</td>
<td>0.66 (0.32, 1.39)</td>
<td>2.52 (1.43, 4.45)*</td>
<td>3.80 (1.68, 8.64)*</td>
</tr>
<tr>
<td>Cognitive Score</td>
<td>-0.45 (SE = 0.92)</td>
<td>-4.46 (SE = 1.10)*</td>
<td>-4.01 (SE = 1.28)*</td>
</tr>
<tr>
<td>Cognitive &lt; 70</td>
<td>0.94 (0.55, 1.61)</td>
<td>1.38 (0.79, 2.38)</td>
<td>1.47 (0.75, 2.88)</td>
</tr>
<tr>
<td>Cognitive &lt; 85</td>
<td>1.03 (0.74, 1.42)</td>
<td>1.82 (1.26, 2.64)*</td>
<td>1.77 (1.15, 2.74)*</td>
</tr>
<tr>
<td>Language Score</td>
<td>-0.20 (SE=1.09)</td>
<td>-3.52 (SE = 1.33)*</td>
<td>-3.32 (SE = 1.53)*</td>
</tr>
<tr>
<td>Language &lt; 70</td>
<td>0.76 (0.52, 1.13)</td>
<td>1.57 (1.04, 2.37)*</td>
<td>2.06 (1.25, 3.40)*</td>
</tr>
<tr>
<td>Language &lt; 85</td>
<td>1.06 (0.79, 1.43)</td>
<td>1.45 (1.00, 2.09)*</td>
<td>1.37 (0.90, 2.08)*</td>
</tr>
<tr>
<td>NDI (Cog &lt; 70)</td>
<td>0.82 (0.51, 1.31)</td>
<td>1.68 (1.07, 2.65)*</td>
<td>2.05 (1.15, 3.64)*</td>
</tr>
<tr>
<td>NDI (Cog &lt; 85)</td>
<td>0.99 (0.72, 1.36)</td>
<td>1.78 (1.24, 2.57)*</td>
<td>1.80 (1.17, 2.76)*</td>
</tr>
</tbody>
</table>

*: p<0.05 Adjusted OR (95% CI) for categorical outcomes by logistic regression, score adjustment for continuous outcomes by linear regression. CP: cerebral palsy, GMFCS: Gross Motor Function Classification Staging, SE: standard error, NDI: neurodevelopmental impairment.

Adjusted Neurodevelopmental Outcomes

After adjusting for the specified model covariates via linear regression, continuous cognitive and language scores were not significantly different for infants with low grade IV-PVH compared to those without IV-PVH (Table 3, Figure 2). For infants with severe hemorrhages, cognitive and language scores were significantly lower compared to infants without IV-PVH.

After adjusting for specified model covariates, the odds ratios for each categorical neurodevelopmental outcome among the low grade group were not significant. This indicates that the 18-22 month outcomes for infants with Grade 1 or 2 hemorrhages are not significantly different than for infants without evidence of IV-PVH.
Figure 2. Comparison of Adjusted Odds Ratios for Categorical Neurodevelopmental Outcomes. Odds ratios with 95% confidence intervals are represented by the horizontal colored bars. Vertical green line is a reference line equal to OR=1 (null). Confidence intervals crossing the reference line are not significant (p>0.05).

The non-significant increased risk for language impairment among the low grade group no longer exists after adjustment for potential confounding.

Among comparisons of infants with high grade hemorrhage to infants without IV-PVH, only the adjusted odds ratio for “Cognitive < 70” was not significantly different
between groups. The adjusted odds of poor motor outcomes, language impairment, composite NDI, or mild cognitive impairment were significantly increased for this group.

Independent predictors of the outcomes of interest after adjusting for all other specified covariates are summarized in Table 4. None of the model covariates was an independent predictor of every outcome after adjusting for all other covariates. Low grade hemorrhage and chorioamnionitis did not independently predict any impairment outcome. PDA as a predictor of mild language delay approached significance (p=0.06) but was otherwise not a significant predictor of other outcomes. For gestational age, younger gestational ages were at increased risk of poor outcome. In all cases of gender, “male” was the predictor of poor outcome. Where antenatal steroids were a predictor, this was a protective effect against the outcome. Presence of severe hemorrhage, postnatal steroids, high frequency ventilation, and sepsis were negative predictors of outcome. Where race was a predictor of outcome, black infants were at increased risk of outcome in each instance. Hispanic infants were at increased risk for decreased cognitive score, language score, and mild or moderate language impairment. “Other race” was associated with an increased risk of decreased language score (p=0.1) and mild language impairment (p=0.06), but neither of these associations reached statistical significance.
Table 4. Independent Predictors of Adjusted Neurodevelopmental Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Low Grade IV-PVH</th>
<th>High Grade IV-PVH</th>
<th>Gestational Age</th>
<th>Gender</th>
<th>Race/Ethnicity</th>
<th>Antenatal Steroids</th>
<th>Chorioamnionitis</th>
<th>Postnatal Steroids</th>
<th>High Frequency Ventilation</th>
<th>Sepsis</th>
<th>Patent Ductus Arteriosus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any CP</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFCS &gt; 2</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Score</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive &lt; 70</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive &lt; 85</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Score</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>#</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language &lt; 70</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language &lt; 85</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDI (Cog &lt; 70)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDI (Cog &lt; 85)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

✓: p<0.05 after adjusting for all other model factors; #: (0.05 < p ≤ 0.1) after adjusting for all other model factors. CP: cerebral palsy, GMFCS: Gross Motor Function Classification Staging, NDI: neurodevelopmental impairment.

To allay concerns about postnatal steroids and high frequency ventilation potentially existing in the causal pathway between IV-PVH and outcome, regression was also run using models without either of these covariates as well as with models removing only one of these factors at a time. There were no significant differences in any outcome using these strategies.
DISCUSSION

This analysis of a large, multi-center cohort of ELGA infants surviving to follow-up at 18-22 months corrected age reaffirms the traditionally held belief that neurodevelopmental outcomes at this stage do not differ significantly between infants without IV-PVH and infants with Grade 1 or 2 hemorrhages in a contemporary cohort, particularly when adjusting for potential confounders. The inclusion of comparisons between high grade and no IV-PVH provide internal validity to our results.

The results of our study are in contrast to a recent cluster of studies reporting on outcomes of low grade hemorrhage for extremely preterm infants born in the 1990s. Sherlock et al (2005)\textsuperscript{18}, in a regional Australian cohort of 8 year olds born in the early 1990s, found no difference in rates of CP for those with a history of Grade 1 IV-PVH but an increase in risk of CP for those with Grade 2 IV-PVH. There were no significant differences in cognitive or educational outcomes for either degree of low grade hemorrhage. From the French multi-center EPIPAGE cohort (born 1997), Ancel (2006)\textsuperscript{16} and Beaino (2011)\textsuperscript{32} found increased risk of cerebral palsy with low grade hemorrhage (particularly Grade 2) but no difference in the 5-year cognitive outcomes for those with low grade hemorrhage after adjusting for potential confounders. While these three studies do report on low grade IV-PVH outcomes within a larger framework, they were not specifically designed to evaluate low grade outcomes. Finally, Patra et al (2006)\textsuperscript{20} specifically investigated the outcomes of extremely low birth weight (ELBW, <1000g) infants with and without low grade IV-PVH born between 1992 and 2000 in a single-
center retrospective cohort study and found significant and dramatic increases in the rates of neurodevelopmental impairment, major neurologic abnormality, and cognitive/language impairment at 20 months corrected age, all with adjusted odds ratios approximately two-fold.

The results of Patra et al\textsuperscript{20} are intriguing as they reflect higher risk of major disability than has generally been reported in relation to low grade hemorrhages but potential pitfalls in the study exist as previously outlined by Inder\textsuperscript{33}. The differences in these study results highlight that comparisons between prior studies to our current study, and to each other, must be carefully considered. As mentioned previously, differences in the results of these studies may be attributable to definition of the cohort of interest, study design (e.g. single vs. multicenter studies), evaluation methods, and the evolution of practice management.

Previous studies reporting on low grade IV-PVH have used various cohort definitions including narrow and broad ranges of either birth weight cohorts (from $<1000g$ to $<2000g$) or gestational age cohorts ($<27$ weeks to $<32$ weeks).\textsuperscript{5, 10-23, 32} Results for broadly defined cohorts may be biased towards the null hypothesis as the incidence of IV-PVH in older and larger preterm infants is lower and the impact of IV-PVH on outcomes may be less than for extremely premature infants. “ELGA” in this study was defined to include those infants born at $<27$ weeks gestational age for two reasons. Logistically, the gestational age ceiling of the ongoing NRN follow-up study is
currently 26 6/7 weeks. Clinically, this gestational age population is also the most at-risk for IV-PVH.\textsuperscript{4, 34-36}

Our study uses a multi-center cohort. A design of this type is beneficial in allowing relatively rapid accumulation of a large sample size. The period of time required to establish a large cohort is not only a matter of efficiency but it is the risk of any broad longitudinal study to result in a heterogeneous cohort in terms of demographics, clinical exposures, treatment, or evaluation. This is particularly true in neonatology which is a continuously evolving field where significant changes in practice generally occur every few years. In order to minimize such issues, cohorts would ideally be gathered over short periods of time but this is often not feasible as large numbers of infants are often required to achieve adequate power. Also, while single-center studies can provide very useful information, the ability to generalize those results beyond that center may be limited by site-specific characteristics, such as population demographics or clinical management strategies.

The benefits of multi-center cohorts do not come without potential consequences. Center-to-center differences in population, clinical practice (such as the use of prophylactic indomethacin or initial ventilation strategies), and diagnoses may exist and must be considered. Vohr et al (2004)\textsuperscript{37} demonstrated significant NRN center differences in neurodevelopmental outcomes even after adjusting for demographics and antenatal interventions and concluded that center-specific practices strongly influence outcome.
To control for clustering of infants in centers and potential center differences in our study, “center” was included as a random effect in mixed effects regression modeling. While classic regression modeling assumes independence of each observation, center differences may lead to a lack of independence for subjects, or clustering. Mixed effects modeling assumes two sources of variation – within cluster (fixed effects) and between cluster (random effects). In our model, IV-PVH grade, gestational age, gender, race/ethnicity, antenatal steroids, chorioamnionitis, postnatal steroids, high frequency ventilation, sepsis, and patent ductus arteriosus would be considered “fixed effects”. Although the homogeneity of the populations across centers is not explicitly stated in the EPIPAGE\textsuperscript{16, 32} nor the Sherlock et al\textsuperscript{18} studies, neither controlled for center differences within their multi-center studies.

The historical and clinical practice context must also be considered when comparing studies across time. As one example, steroids have taken on a very important role in neonatology over the past 20-30 years. The use of antenatal steroids promotes acceleration of fetal lung maturity and has been shown to reduce respiratory distress, duration of mechanical ventilation, and IV-PVH. Since the 1994 publication of an NIH Consensus statement\textsuperscript{38} recommended the practice, a single course of antenatal steroids in the setting of imminent preterm delivery has been rapidly adopted and is generally considered standard of care.\textsuperscript{39} Postnatal steroids can be used to facilitate extubation from mechanical ventilation and may decrease the incidence of chronic lung disease. However, in the late 1990s, the liberal use of postnatal steroids came under scrutiny due to concerning links to increased rates of cerebral palsy. These concerns
have been validated in meta-analyses\textsuperscript{40-42} and the empiric/widespread use has fallen out of favor.\textsuperscript{43} Following the decline in postnatal steroid use, the rate of cerebral palsy has declined.\textsuperscript{44} The eight year cohort of Patra et al\textsuperscript{20} (1992-2000) encompassed the era of these important changes and the lack of reporting or adjusting for steroid use raises the possibility that the significant differences in low grade IV-PVH outcomes may have been related to steroids. Changes in evaluation methods may further complicate comparisons of outcomes over time. The Bayley II\textsuperscript{45}, which was in use for infants born within the NRN between 1993 and 2005, reports a Mental Developmental Index (MDI) which is an inseparable composite measure of both cognitive and language domains. However, the Bayley III separates cognitive and language scores into two reportable domains. Direct comparison or simple translation of a Bayley III cognitive score into a Bayley II MDI equivalent (or vice-versa), is problematic and not currently possible although conversion methods are being sought.\textsuperscript{46} The NRN began uniformly using the Bayley III for cognitive and language assessment for infants born on or after January 1, 2006, thus the entire cohort in this study was assessed using the Bayley III. Thus, comparison of our study using an entirely Bayley III evaluated cohort to a study using an entirely Bayley II evaluated cohort should not focus on the numeric score attained by each group but the overall trend of how the low grade group behaves relative to the control group.

The Bayley III introduces the possibility of a routine language measure in this follow-up population. Increased risk of language delay is well documented for premature infants and has been described as persisting into and beyond adolescence. This study is the first low grade IV-PVH study to report specific Bayley III language
domain scores at this age. In the unadjusted results, we did not find a significant
difference in continuous language score or mild language impairment/delay with score <
85 (-1 SD) among infants with low grade hemorrhage compared to infants without IV-
PVH. The failure to detect a significant association between IV-PVH and language delay
may be due to inadequate study power for this outcome. Alternatively, detecting a
specific language delay among infants at a developmental age of 18-22 months may not
be achievable with current measures. Further investigation of neuroimaging studies and
language ability in school age premature outcome studies is warranted.

The limitations of our study involve the interrater reliability of cranial ultrasound
detection of IV-PVH, particularly at the lowest grades, and the power of the tests of
categorical outcomes and tests between levels of low grade hemorrhage. Cranial
ultrasound technique and consequently interpretation is highly operator dependent
such that systematic differences between radiologists at a given site as well as between
radiologists between sites may exist. Hintz et al (2007)\textsuperscript{47} report 40% (moderate)
agreement for “low grade” hemorrhage between two NRN centralized readers while the
agreement for grade 1 or grade 2 hemorrhage specifically was even worse at 26% and
20% respectively. The sensitivity of local reads compared to centralized readers for
Grade 1/2 hemorrhage of 48-68%. The NRN continues to use local ultrasound reads as
there is no substantial improvement in agreement when using central readers in
addition to the added cost such a practice would entail.
In regards to power, our study was powered on a primary outcome of continuous cognitive score. While the comparison of “low grade” hemorrhage to either no IV-PVH or severe hemorrhage is highly powered, we were unable to reach the goal of sample size of 143 subjects per group for comparisons of no IV-PVH vs. Grade 1 vs. Grade 2 despite a three year cohort of infants. It is difficult to imagine that an additional 3 infants with Grade 1 or 13 infants with Grade 2 hemorrhage would significantly change the results. Simply expanding the cohort to include infants born in 2005 was not feasible as those infants were evaluated with the Bayley II. To achieve adequate power for our categorical outcomes, larger cohorts, typically 5-7 year NRN Follow-Up Study birth cohorts, would be necessary. However, the confidence intervals of the adjusted odds ratios for these categorical outcomes do provide us with some valuable information. If we assume that it is unlikely that low grade hemorrhage would lead to improved outcomes (synonymous with OR < 1), then we can look at the upper limit of the 95% confidence interval to get an idea of the “worst case scenario” increase in poor outcomes since increased sample size should lead to narrowing of the confidence interval.

While we have shown that neurodevelopmental outcomes at 18-22 months are not negatively impacted by Grade 1/2 hemorrhage, it cannot be assumed that the same could be said at later ages. The overall stability in diagnoses between toddler-hood and early school age is typically poor. Of the components of “neurodevelopmental impairment”, significant motor delay is the most likely to remain stable. In contrast, cognitive diagnoses are not stable and the direction of change varies between reports.
Hack et al (2005)\textsuperscript{48} reporting on differences in outcomes between 20 months and 8 years saw a general improvement in cognitive scores for infants with either mild (<-1 SD, > -2 SD) or moderate-severe (< -2 SD) delay at 20 months of age. However, Roberts et al (2010)\textsuperscript{50} found only 46% of composite disability remained stable with equal proportions of increase and decrease in disability status between 2 and 8 years and changes in status were attributed to shifts in cognitive classifications. Marlow et al (2005)\textsuperscript{49} showed 86% of infants at 30 months with severe disability remained in the moderate-severe disability range. However, only 50% of disability diagnoses remained stable with approximately 1/3\textsuperscript{rd} of severely disabled infants improving their status over time while 2/3\textsuperscript{rd}s of infants without disability at 30 months were then classified as having mild-severe disability at age 6. Again, it is likely these shifts are related to cognitive score shifts although this not explicitly reported.

In addition, cognitive test scores represent only one piece of the puzzle in assessing the late outcomes of school performance, academic achievement, and behavior. Nearly 2/3rds of ELBW children require special education support in school and are more likely than term peers to have subject specific learning problems.\textsuperscript{51} High prevalence/low severity disabilities such as learning disabilities, attention/deficit hyperactivity disorders, specific neuropsychological deficits, and behavioral problems may gradually emerge and may contribute to the trend of worsening outcomes over time for ELBW and VLBW children.\textsuperscript{52} Theoretically, the inability to keep up with increasing cognitive and performance demands as a child grows older may be related to subtle cognitive deficits. While Sherlock et al (2005)\textsuperscript{18} reported no significant
differences in regards to cognitive testing or major disabilities for ELBW infants with low grade hemorrhage at age 8, they did not investigate low severity disabilities. It is not clear what contribution low grade hemorrhage may have to these more subtle cognitive disabilities but the ability to quantify the risk in concert with neuroimaging could be important to understanding the pathophysiology of this common morbidity.

Cranial ultrasound has been used routinely in neonatology as a screening and diagnostic method since the late 1970s. While magnetic resonance imaging (MRI) continues to emerge as an important research tool in understanding brain development and pathophysiology, discussions with parents regarding IV-PVH diagnosis and possible outcomes often happen in the setting of clinical instability where MRI is not feasible. Cranial ultrasound continues to have a strong clinical presence due to its portability, speed, and lower cost. As such, further investigating the predictive abilities of this modality, particularly over time as more subtle disabilities may emerge with increased cognitive and performance demands, will be useful to clinical neonatologists as they continue to advise parents on the potential outcomes of their extremely premature infants.
BIBLIOGRAPHY


