Prognostic Accuracy of Heart Rate Variability Analysis in Neonatal Encephalopathy: A Systematic Review

Vânia Oliveira, Rui Martins, Natasha Liow, Justinas Teiserskas, Wilhelm von Rosenberg, Tricia Adjei, Vijayakumar Shivamurthappa, Peter J. Lally, Danilo Mandic, Sudhin Thayyil

Centre for Perinatal Neuroscience, Imperial College London, London, UK; Royal London Hospital, Barts Health NHS Trust, London, UK; Neonatal Unit, Imperial Healthcare NHS Trust, London, UK; Electrical and Electronic Engineering, Imperial College London, London, UK

Keywords
Heart rate variability · Neonatal encephalopathy · Hypoxic ischemic encephalopathy

Abstract
Background: Heart rate variability analysis offers real-time quantification of autonomic disturbance after perinatal asphyxia, and may therefore aid in disease stratification and prognostication after neonatal encephalopathy (NE). Objective: To systematically review the existing literature on the accuracy of early heart rate variability (HRV) to predict brain injury and adverse neurodevelopmental outcomes after NE. Design/Methods: We systematically searched the literature published between May 1947 and May 2018. We included all prospective and retrospective studies reporting HRV metrics, within the first 7 days of life in babies with NE, and its association with adverse outcomes (defined as evidence of brain injury on magnetic resonance imaging and/or abnormal neurodevelopment at ≥1 year of age). We extracted raw data wherever possible to calculate the prognostic indices with confidence intervals. Results: We retrieved 379 citations, 5 of which met the criteria. One further study was excluded as it analysed an already-included cohort. The 4 studies provided data on 205 babies, 80 (39%) of whom had adverse outcomes. Prognostic accuracy was reported for 12 different HRV metrics and the area under the curve (AUC) varied between 0.79 and 0.94. The best performing metric reported in the included studies was the relative power of high-frequency band, with an AUC of 0.94. Conclusions: HRV metrics are a promising bedside tool for early prediction of brain injury and neurodevelopmental outcome in babies with NE. Due to the small number of studies available, their heterogeneity and methodological limitations, further research is needed to refine this tool so that it can be used in clinical practice.

Introduction
Neonatal encephalopathy (NE) results primarily from a presumed lack of oxygen and blood flow to the fetal brain around the time of birth and accounts for 1 million...
deaths worldwide every year [1]. Although rescue hypothermic neuroprotection improves survival and neurodevelopmental outcome after NE, the early identification of babies at risk of brain injury and adverse outcomes is challenging due to the evolving clinical picture. An abnormal voltage or pattern of amplitude-integrated electroencephalography (aEEG) has been used as a bedside tool for quantifying brain injury and as an inclusion criterion for some of the cooling trials [2, 3]. More recent data suggests that aEEG does not offer an added advantage over clinical examination [4] and its prognostic accuracy in cooled infants is poor [5].

As autonomic disturbance is the hallmark of perinatal hypoxic injury, heart rate variability (HRV) analysis may offer a promising solution. HRV analysis quantifies variations in heartbeat intervals, offering a non-invasive measure of autonomic function, as it reflects the actions of the sympathetic and parasympathetic nervous systems. The simplest measure of HRV is determined by the standard deviation of the difference between consecutive RR intervals (SDNN), but many other measures of HRV continue to be developed [6, 7]. Currently, HRV analyses can include not only linear measures in time and frequency domains but also several non-linear metrics derived from complexity analysis [8].

Loss of HRV occurs in traumatic brain injury and during seizures [9, 10], and is associated with brain injury and adverse neurodevelopment in very-low-birth-weight babies [11, 12]. In babies with birth asphyxia, the early identification of brain injury risk via HRV analysis could benefit patients by facilitating early disease stratification and the implementation of the most appropriate treatment. Nonetheless, it is unclear what the prognostic value of these electrocardiographic (ECG) changes are in identifying babies at risk of brain injury and whether they are useful in accurately predicting neurodevelopmental outcomes.

Our aim was to systematically review the available evidence about the prognostic accuracy of HRV analysis in predicting brain injury and adverse neurodevelopmental outcomes after NE.

**Methods**

**Search Strategy**

Two reviewers (V.O. and R.M.) independently searched the literature according to the strategy outlined in a predefined protocol, using Ovid SP (selecting Embase Classic+Embase, Global Health, HMIC Health Management Information Consortium, Ovid MEDLINE(R) ALL, Maternity & Infant Care Database), Web of Science, Scopus, and the Cochrane Register of Diagnostic Test Accuracy Studies, between May 1947 and May 2018. The databases were searched using the following combination of keywords: (heart rate variability) and (birth asphyxia OR neonatal brain injury OR hypoxic ischemic encephalopathy OR hypoxic ischaemic encephalopathy OR neonatal encephalopathy). No language restrictions were applied. An example of a full search strategy is provided in Appendix 1. We also searched the reference lists of the retrieved articles. Our report follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [13].

**Study Selection**

We included all prospective and retrospective cohort or case-control studies comparing the HRV of babies with NE, with either brain injury on neonatal magnetic resonance imaging (MRI) or an adverse neurodevelopmental outcome. All studies on term and late preterm babies (≥35 weeks gestational age) with NE, who were ≤1 week old at the start of ECG measurements, were eligible.

The index test was HRV analysis, and included any metrics (time and frequency domain, or complexity analysis), regardless of the method of ECG data collection (single or multiple leads; real-time or archived data) or post-processing software.

The reference test was defined as either a structured neurodevelopmental examination at ≥12 months of age, or, as a surrogate biomarker, abnormalities on brain MRI performed within 6 weeks of birth that are known to be associated with adverse neurological outcomes [14].

**Data Extraction and Analysis**

Two reviewers (V.O. and R.M.) independently screened the titles and abstracts of the retrieved articles, obtained full texts and online supplements of all relevant studies, and extracted the data.
Prognostic Accuracy of HRV Analysis in NE

In total, these 4 studies included 228 patients, 205 of whom had NE and ECG recordings/HRV calculated: 46 (22.4%) with mild NE on admission, 88 (42.9%) with moderate NE, and 71 (34.6%) with severe NE. In total, 39% (80/205) of the babies had an adverse outcome, including a mortality rate of 12% (25/205).

One study reported data from babies admitted to the NICU before cooling therapy was standard care and 3 studies included babies who received cooling treatment. All 4 studies were observational and retrospective.

Characteristics of Index and Reference Tests

Of the 4 included studies, 2 used a structured neurodevelopmental examination (1 with the Bayley Scales of Infant Development III [BSID-III] and 1 with Griffiths) as a gold standard indicator of adverse neurodevelopmental outcome; 2 used neonatal MRI within 2 weeks after birth (Table 1).

The ECG trace was obtained from bedside multiparameter monitors \( n = 100 \), a combined EEG monitor (NicotOneTM, Vyasis Healthcare, San Diego, CA, USA) \( n = 61 \), and a HeRO monitor \( n = 67 \). ECG recordings were performed from “as soon as possible after birth” up to day 7, although the earliest ECG data published refers to 24 h of postnatal life. Three studies used in-house post-processing analysis pipelines, while 1 study used HeRO monitor scores (MPSC™, Charlottesville, VA, USA).

All studies reported HRV data at 24 h from birth (or day 1). Three studies also reported data at different time points, i.e., 48 h, 96 h, and 7 days postnatal age. Conse-

Table 1. Characteristics of the included studies

<table>
<thead>
<tr>
<th>First author</th>
<th>Study design</th>
<th>Characteristics of the study population</th>
<th>NE stage, n (%)</th>
<th>TH</th>
<th>n/N</th>
<th>Index test</th>
<th>Age at index test</th>
<th>Gold standard</th>
<th>Age at gold standard</th>
<th>Prevalence of adverse outcome, n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massaro [16]</td>
<td>retrospective cohort</td>
<td>NE, GA 36+ weeks, NICHD cooling criteria</td>
<td>moderate 11 (55) severe 9 (45)</td>
<td>Yes</td>
<td>20/20</td>
<td>rHF rLF</td>
<td>5–41 h</td>
<td>death or BSID-II &lt;2SD</td>
<td>15 months</td>
<td>10/20 (50)</td>
</tr>
<tr>
<td>Metzler [20]</td>
<td>retrospective cohort</td>
<td>NE, GA 35+ weeks, NICHD cooling criteria</td>
<td>moderate 55 (74) severe 19 (26)</td>
<td>Yes</td>
<td>74/80</td>
<td>RMSS RMSS rLF</td>
<td>24–27 h</td>
<td>death of predominant basal ganglia or global injury on MRI</td>
<td>10 days</td>
<td>24/74 (32)</td>
</tr>
<tr>
<td>Goulding [17]</td>
<td>retrospective cohort</td>
<td>NE, GA 37+ weeks, and 2 of the following: pH &lt;7.1, 5-min Apgar ≤6, lactate &gt;7 mmol, and abnormal neurology/seizures</td>
<td>mild 22 (36) moderate 9 (20) severe 13 (30)</td>
<td>No</td>
<td>44/61</td>
<td>SDNN TINN VLF LF, HF</td>
<td>12–48 h</td>
<td>death, cerebral palsy, or Griffiths developmental quotient &lt;87</td>
<td>24 months</td>
<td>20/44 (45)</td>
</tr>
<tr>
<td>Vergales [18]</td>
<td>retrospective cohort</td>
<td>NE, GA 34+ weeks, NICHD cooling criteria</td>
<td>normal 5 (8) mild 19 (28) moderate 13 (19) severe 30 (45)</td>
<td>Yes</td>
<td>67/67</td>
<td>HRC index HRV</td>
<td>1–7 days</td>
<td>death or “moderate” or “severe” injury on MRI</td>
<td>7 days</td>
<td>26/87 (29)</td>
</tr>
</tbody>
</table>

TH, therapeutic hypothermia; n/N, babies included in review/total reported in the study; Index test, HRV metrics for which prognostic accuracy metrics were reported; NE, neonatal encephalopathy; GA, gestational age; NICHD, National Institute of Child Health and Human Development; rLF, relative power of high-frequency band; rLF, relative power of low-frequency band; RMSS, root mean square at short time scale (15–50 beats); SDNN, standard deviation of the RR intervals; TINN, triangular interpolation of the NN intervals histogram; VLF, power of very-low-frequency band; LF/HF, low-/high-frequency power ratio; HRC index, heart rate characteristics index (proprietary); HRV, heart rate variability (proprietary); BSID, Bayley Scales of Infant Development.
sequently, pooling of data could only be attempted for the HRV metrics provided at 24 h of postnatal age (although different metrics were reported at this time point).

The relative power of low frequencies of the ECG spectrogram (rLF) was reported in 2 studies (both using 0.05–0.25 Hz as a threshold for LF). Otherwise, all other HRV metrics were reported only once across all studies. The selection method of reported HRV metrics was not pre-specified.

Prognostic Accuracy

Twelve different HRV metrics were reported by the included studies. The AUC for prognosticating an adverse outcome from these metrics ranged from 0.79 to 0.94. As presented in Table 2, the best performing feature reported was the relative power of the high-frequency band (rHF) at 24–27 h with an AUC of 0.94, a sensitivity of 0.84, and a specificity of 0.87 [16]. This was based on a sample size of only 20 babies. High frequency was defined score (proprietary); AUC, area under the receiver operating characteristics curve; PPV, positive predictive value; NPV, negative predictive value; OR, odds ratio; n.r., data not reported.

### Table 2. Prognostic accuracy of the best-performing HRV metrics originally reported by the included studies

<table>
<thead>
<tr>
<th>First author</th>
<th>Metric</th>
<th>TP</th>
<th>FN</th>
<th>FP</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>AUC (95% CI)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massaro et al. [16], 2014</td>
<td>rHF</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0.80 (0.44–0.98)</td>
<td>0.90 (0.56–1.00)</td>
<td>0.89 (0.52–1.00)</td>
<td>0.82 (0.48–0.98)</td>
<td>0.85 (0.69–1.00)</td>
<td>36 (2.72–476.30)</td>
</tr>
<tr>
<td>Metzler et al. [20], 2017</td>
<td>RMSS</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>38</td>
<td>0.62 (0.41–0.81)</td>
<td>0.86 (0.73–0.95)</td>
<td>0.71 (0.48–0.89)</td>
<td>0.81 (0.67–0.91)</td>
<td>0.74 (0.63–0.86)</td>
<td>10.56 (3.20–34.82)</td>
</tr>
<tr>
<td>Metzler et al. [20], 2017</td>
<td>RMSL</td>
<td>16</td>
<td>8</td>
<td>11</td>
<td>33</td>
<td>0.67 (0.45–0.84)</td>
<td>0.75 (0.60–0.87)</td>
<td>0.59 (0.39–0.78)</td>
<td>0.81 (0.65–0.91)</td>
<td>0.71 (0.59–0.82)</td>
<td>6 (2.02–17.83)</td>
</tr>
<tr>
<td>Metzler et al. [20], 2017</td>
<td>rLF</td>
<td>17</td>
<td>7</td>
<td>14</td>
<td>30</td>
<td>0.71 (0.49–0.87)</td>
<td>0.68 (0.52–0.81)</td>
<td>0.55 (0.36–0.73)</td>
<td>0.81 (0.65–0.92)</td>
<td>0.70 (0.58–0.81)</td>
<td>5.2 (1.75–15.4)</td>
</tr>
<tr>
<td>Goulding et al. [17], 2015</td>
<td>SDNN</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>11</td>
<td>0.90 (0.56–1.00)</td>
<td>0.70 (0.35–0.93)</td>
<td>0.75 (0.43–0.95)</td>
<td>0.88 (0.47–1.00)</td>
<td>0.80 (0.62–0.98)</td>
<td>21 (1.77–248.11)</td>
</tr>
<tr>
<td>Goulding et al. [17], 2015</td>
<td>TINN</td>
<td>8</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>0.80 (0.44–0.98)</td>
<td>0.50 (0.27–0.79)</td>
<td>0.53 (0.34–0.72)</td>
<td>0.83 (0.52–0.98)</td>
<td>0.69 (0.52–0.87)</td>
<td>5.71 (0.92–35.48)</td>
</tr>
<tr>
<td>Goulding et al. [17], 2015</td>
<td>VLF</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>0.80 (0.44–0.98)</td>
<td>0.65 (0.38–0.86)</td>
<td>0.50 (0.29–0.72)</td>
<td>0.85 (0.55–0.98)</td>
<td>0.72 (0.55–0.90)</td>
<td>7.33 (1.16–46.24)</td>
</tr>
<tr>
<td>Goulding et al. [17], 2015</td>
<td>LF</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>0.90 (0.56–1.00)</td>
<td>0.59 (0.33–0.82)</td>
<td>0.56 (0.30–0.80)</td>
<td>0.81 (0.59–1.00)</td>
<td>0.74 (0.59–0.90)</td>
<td>12.86 (1.31–125.78)</td>
</tr>
<tr>
<td>Cumulative</td>
<td></td>
<td>100</td>
<td>32</td>
<td>63</td>
<td>157</td>
<td>0.74 (0.67–0.80)</td>
<td>0.79 (0.73–0.84)</td>
<td>0.66 (0.59–0.73)</td>
<td>0.79 (0.73–0.84)</td>
<td>0.73 (0.69–0.77)</td>
<td>7.78 (4.75–12.76)</td>
</tr>
</tbody>
</table>

Prognostic metrics were calculated from the available raw data or information provided by the authors. Confidence intervals appear in parentheses. n.r., not reported; TP, true positive; FN, false negative; TN, true negative; AUC, area under the receiver operating characteristics curve; PPV, positive predictive value; NPV, negative predictive value; OR, odds ratio; RMSS, root mean square at short time-scale (15–50 beats); RMSL, root mean square at long time-scale (100–150 beats); rLF, relative power of low-frequency band; rHF, relative power of high-frequency band; SDNN, standard deviation of the RR intervals; TINN, triangular interpolation of the NN interval histogram; VLF, power of very-low-frequency band; LF, power of low-frequency band; HRV, heart rate variability score (proprietary); HRC index, heart rate characteristics index (proprietary).
in this study as ranging between 0.30 and 0.80 Hz, although this threshold was different from what was reported in other studies.

In the work of Goulding et al. [17], the power of the low-frequency band (LF) predicted death or adverse neurodevelopment with an AUC of 0.79 (95% confidence interval [CI] 0.57–0.94), although the difference in LF between severity groups was no longer statistically significant when adjusted for phenobarbitone use.

The study by Vergales et al. [18] reported an AUC of 0.83 (CI not reported) for HRV score (proprietary), on day 1, to predict death or abnormal MRI, although this was based on a device originally designed to predict sepsis onset in preterm infants [19]. The authors also highlighted the difference of HRV between babies with moderate or severe encephalopathy and those with mild or no encephalopathy at 24 h (15.8 vs. 25.4 respectively, \( p = 0.01 \) for heart rate characteristics [HRC] index and 2.48 vs. 1.42, \( p < 0.05 \) for HRV score [both proprietary metrics]). Equally, survivors had significantly higher variability than babies who died (HRV score 24.3 vs. 10, \( p < 0.05 \); HRC index 2.85 vs. 3.64, \( p < 0.001 \)).

In another study by Metzler et al. [20], the root mean square at short time-scale (RMSS) predicted basal ganglia injury on MRI with an AUC of 0.79 (CI not reported). In this study, dichotomisation of the outcome excluded watershed/mild basal ganglia injury. RMSS and root mean square at long time-scale (RMSL) were significantly associated with MRI brain injury scores, after adjusting for confounders.

Three studies also reported the prognostic accuracy of HRV metrics beyond 24 h: AUC 0.80 (0.62–1.00) for the power of very low frequency (VLF) at 48 h [17]; 0.82 for LF at 93–96 h; 0.81 for HF at 102–105 h [16]; and 0.75 on days 4–7 of life [18] for HRC index (CI not reported). Where >1 time point was reported, 24 h was the time point of higher accuracy but none of the studies reported data for the first 24 h of postnatal age.

As several prognostic indices and CIs were not reported in many studies, we extracted the raw data, where possible, to calculate these measures (Table 3; Fig. 2). Although meta-analysis was not appropriate due to the heterogeneity of HRV metrics, to explore what could become the potential performance of HRV to predict adverse outcomes we calculated the cumulative true/false positives/negatives and the associated prognostic values.

There was a low risk of bias across the studies (Fig. 3). Whereas all studies reported blinding to clinical information, only 1 reported that index test assessors were blinded to the results of the comparator test and vice versa. Missing data, artefacts in the ECG, and suboptimal data

---

**Fig. 2.** Schematic summary of the accuracy of the calculated prognostic metrics. Names of first authors are used to denote individual studies.
collection were reported by all authors since these were often obtained from archived data. Two of the 4 selected studies were from the same research group [16, 20]. It is not clear if there was any data overlap or if any participant data may have been included in both reports.

Additionally, given the good prognostic performance in all studies, publication bias cannot be excluded. The studies assessed the predictive value of different features of HRV, using different time points and gold standards for the assessment of outcomes. There was significant under-reporting, and pre-selection of specific HRV parameters was unclear across all studies.

**Excluded Studies**

Eight studies were excluded, 2 of which [21, 22] measured the severity of brain injury with methods that did not meet the criteria for this review (Table 4). These studies reported significant differences in mean HRV between severity groups, as well as between survivors and non-survivors. In another study [23], several parameters were higher in the unfavourable-outcome group than in the favourable-outcome group at 24 h, but not at 48 and 96 h. The 24-h result was contradictory to the findings of all other studies, but no estimates or hypothesis testing were reported, and the authors’ primary aim was not to describe prognostic accuracy. They informed us that they are currently examining a larger dataset with that aim. One excluded study [24], whose cohort overlapped with another study included in this review, reported an AUC of 0.66 which referred to the average of 60 ECG features examined. As individual accuracy was not reported, it is possible that this value is not representative of the best accuracy of individual parameters.

**Discussion**

We systematically reviewed the literature on the accuracy of HRV analysis to predict brain injury on MRI or adverse neurodevelopmental outcome in babies with NE. Although we found a very small number of studies and
Limiting designs, the results are consistent throughout the studies to support that HRV metrics, acquired within days after birth, hold a good predictive value of later neurodevelopmental outcome. At its best performance, calculated HRV metrics had a sensitivity of 0.80, a specificity of 0.90, and an AUC of 0.85 at 24 h of age [16], which represents a sensitivity comparable to aEEG [5] but a higher specificity.

Unfortunately, none of the studies analysed HRV within 6 h of birth, which is the time point by which critical treatment decisions need to be made with regard to therapeutic hypothermia. This may have been due to the logistic difficulties in obtaining consent for research soon after birth. Nevertheless, future prospective studies should attempt to recruit babies immediately after birth so that the value of HRV to support early clinical decision can be examined.

It is important to mention that these studies used a dichotomised approach as has been traditional in NE outcomes research. This means that babies with mild NE are grouped with those with normal neurodevelopmental outcomes to define “favourable outcome” and babies with moderate and severe NE are grouped together as having “adverse outcome”. This approach has been com-

**Table 4. Excluded studies**

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Country</th>
<th>n</th>
<th>Reason for exclusion</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massaro [28]</td>
<td>USA</td>
<td>51</td>
<td>Correlation between HRV and temperature but not between HRV and severity of NE</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Goulding [21]</td>
<td>Ireland</td>
<td>118</td>
<td>Severity of NE classified per EEG, not MRI or NDO</td>
<td>HRV higher in TH group than in pre-TH group. Significant difference between mild and moderate NE in pre-TH group. Significant difference between moderate and severe NE in TH group.</td>
</tr>
<tr>
<td>Aliefendioglu [22]</td>
<td>Turkey</td>
<td>22</td>
<td>Group differences based on Sarnat grading, not MRI or NDO</td>
<td>Significant difference between control group and moderate/severe NE group. Significant difference between cases who died and cases who survived.</td>
</tr>
<tr>
<td>Lasky [29]</td>
<td>USA</td>
<td>2</td>
<td>Describes changes in Q-T segment with cooling</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Javorka [30]</td>
<td>Slovakia</td>
<td>3</td>
<td>Correlation between HRV and temperature but not between HRV and severity of NE</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Schneebaum [31]</td>
<td>USA</td>
<td>40</td>
<td>Examines lateralization of sympathetic and parasympathetic function; no prognostic metrics</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Matic [32]</td>
<td>The Netherlands</td>
<td>19</td>
<td>Classification with machine learning; no prognostic metrics (book chapt.)</td>
<td>Linear discriminant analysis 80%</td>
</tr>
<tr>
<td>Vesoulis [23]</td>
<td>USA</td>
<td>16</td>
<td>Describes effect of time and temperature on HRV; no prognostic metrics</td>
<td>SDNN, TINN, LF, HF, and HF/LF higher in unfavourable outcome group than in favourable outcome group at 24 h, but not at 48 and 96 h (no estimates or hypothesis testing)</td>
</tr>
<tr>
<td>Temko [24]</td>
<td>Ireland</td>
<td>38</td>
<td>Same cohort as Goulding et al. [17]</td>
<td>AUC of average HRV metrics = 0.66</td>
</tr>
</tbody>
</table>

HRV, heart rate variability score; TH, therapeutic hypothermia; NE, neonatal encephalopathy; NDO, neurodevelopmental outcome; SDNN, standard deviation of the RR intervals; TINN, triangular interpolation of the NN interval histogram; LF, power of the low-frequency band; HF, power of the high-frequency band; HF/LF, high frequency-to-low frequency ratio.
mon practice in many studies to date (because babies with mild NE were thought to be “intact survivors”), but emerging evidence suggests that babies who suffer mild NE may not have intact neurodevelopment but instead present with more subtle delays later into childhood [25, 26]. This systematic review, along with our recent survey highlighting clinician uncertainty about treatment choices in babies with mild encephalopathy [27], raises the importance of developing an accurate and early bedside tool to identify babies at risk of adverse neurodevelopment. In fact, it underlines the point that severity stratification tools in NE must shift from the traditional dichotomous approach to one that more accurately represents the full spectrum of NE.

As neurodevelopmental outcomes and severity of brain injury were dichotomised in all of the studies, the data does not inform us whether HRV analysis is sensitive to predict milder forms of adverse outcome. Equally, the fact that 35.7% of the included infants had severe HIE may have increased the sensitivity of HRV to predict adverse outcomes. It is therefore important to ascertain if the prognostic value of HRV is not driven by the inclusion of severe NE but is also able to distinguish between mild, moderate, and severe NE.

Once refined, HRV may become not only an accurate disease stratification tool but may also allow monitoring of the therapeutic effect of established and/or experimental therapies. Despite the number of ongoing trials exploring neuroprotective therapies, understanding which babies may benefit from each therapy is a question we will still struggle to answer if we lack accurate bedside tools for early disease stratification.

**Limitations**

The studies included in our review were affected by the heterogeneity of the index and the reference test, and the selective reporting of outcome measures; this impeded a quantitative meta-analysis. Nonetheless, we tried to collect further information from the authors and extracted the data that allowed for generating further prognostic metrics. The retrospective nature of most of the studies and the fact that some were nested cohorts with small sample sizes imply that adjusting for important confounders was not possible. Other reports have specifically addressed the confounding effect of temperature on HRV [23, 28], producing contradictory results. In view of the evidence that HRV may be predictive of neurodevelopmental outcome despite confounding effects, future studies of HRV in babies with NE should examine important factors such as temperature, seizures, sedation, and analgesia in order to optimize the accuracy and clinical translation of this tool.

**Conclusion**

Despite the small number of available studies and their heterogeneity, the studies included in our review suggest that HRV monitoring offers good prognostic accuracy to predict brain injury or adverse neurodevelopment in NE. Given its association with neurodevelopmental outcome, HRV analysis presents a promising bedside tool for early severity stratification in these babies. However, further refinement of the technology and standardization of risk thresholds are required before clinical use. Future studies which examine this tool must adopt more robust designs that address multiple confounders and more comprehensively describe prognostic accuracy over time, i.e., from birth and throughout the progression of encephalopathy.

**Acknowledgements**

We would like to thank Dr. Robert Goulding for providing the raw data of the work published by his group.

**Appendix 1**

**Example of Full Search Strategy**

Using Ovid SP


Embase Classic+Embase 1947 to 1 May 2018,

Global Health 1973 to 2018 Week 18,

HMIC Health Management Information Consortium 1979 to May 2018,

Ovid MEDLINE(R) ALL 1946 to 1 May 2018,

Maternity & Infant Care Database (MIDIRS) 1971 to May 2018

**Searches**

1# heart rate variability.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, bt, id, cc, nm, kf, px, rx, an, ui, sy]

2# hypoxic ischaemic encephalopathy.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, bt, id, cc, nm, kf, px, rx, an, ui, sy]

3# hypoxic ischemic encephalopathy.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, bt, id, cc, nm, kf, px, rx, an, ui, sy]

4# birth asphyxia.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, bt, id, cc, nm, kf, px, rx, an, ui, sy]

5# neonatal encephalopathy.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, bt, id, cc, nm, kf, px, rx, an, ui, sy]

6# 2 or 3 or 4 or 5

7# 1 and 6.

66

Neonatology 2019;115:59–67
DOI: 10.1159/000493002
Disclosure Statement

We have no conflicts of interest to disclose.

References


Prognostic Accuracy of HRV Analysis in NE

Disclosures

No conflicts of interest.

Funding

V.O. is funded by the National Institute for Health Research (Doctoral Fellowship), P.L. is funded by the National Institute for Health Research (Clinical Trials Fellowship).