Author/s:
Spittle, AJ;Walsh, JM;Potter, C;Mcinnes, E;Olsen, JE;Lee, KJ;Anderson, PJ;Doyle, LW;Cheong, JLY

Title:
Neurobehaviour at term-equivalent age and neurodevelopmental outcomes at 2 years in infants born moderate-to-late preterm

Date:
2017-02-01

Citation:
Spittle, A. J., Walsh, J. M., Potter, C., Mcinnes, E., Olsen, J. E., Lee, K. J., Anderson, P. J., Doyle, L. W. & Cheong, J. L. Y. (2017). Neurobehaviour at term-equivalent age and neurodevelopmental outcomes at 2 years in infants born moderate-to-late preterm. Developmental Medicine and Child Neurology, 59 (2), pp.207-215. https://doi.org/10.1111/dmcn.13297.

Received Date : 29-Feb-2016
Revised Date : 08-Sep-2016
Accepted Date : 10-Sep-2016
Article type : Original Article

(original article; 4 tables, 2 figures)

Neurobehaviour at term-equivalent age and neurodevelopmental outcomes at two years in infants born moderate-to-late preterm

ALICIA J SPITTLE[1,2,3]
JENNIFER M WALSH[2,3,4]
CODY POTTER[2]
EMMA MCINNES[2,3]
JOY E OLSEN[2,3,5]
KATHERINE J LEE[2,6]
PETER J ANDERSON[2,6]
LEX W DOYLE[2,3,5,6]
JEANIE L Y CHEONG[2,3,5]

1 Department of Physiotherapy, University of Melbourne, Melbourne; **2** Murdoch Childrens Research Institute, Melbourne; **3** Newborn Research, Royal Women's Hospital, Melbourne; **4** Paediatric Infant Perinatal Emergency Retrieval (PIPER), The Royal Children's Hospital, Melbourne; **5** Department of Obstetrics and Gynaecology, University of Melbourne, Melbourne; **6** Department of Paediatrics, University of Melbourne, Melbourne, Australia.

Correspondence to Alicia Spittle, 7th Floor, Alan Gilbert Building, University of Melbourne, Grattan Street, Parkville, Victoria, Australia, 3052. Email: aspittle@unimelb.edu.au

PUBLICATION DATA

Accepted for publication 10th September 2016.

Published online 00th Month 2016.

ABBREVIATIONS

Bayley-III Bayley Scales of Infant and Toddler Development 3rd edition

HNNE Hammersmith Neonatal Neurological Examination

NNNS Neonatal Intensive Care Unit Network Neurobehavioral Scale

MLPT Moderate-to-late preterm

[Abstract]

AIM To examine the association between newborn neurobehavioural assessments and neurodevelopmental outcomes at two years in infants born moderate-to-late preterm (MLPT).

METHOD 201 MLPT infants (born 32 to 36+6 weeks' gestation) were assessed with the Hammersmith Neonatal Neurological Examination (HNNE) and NICU Network Neurobehavioral Scale (NNNS), with suboptimal performance defined as scores lower than the 10th centile. Development was assessed at two years' corrected age with the Bayley Scales of Infant and Toddler Development 3rd Edition, with delay defined as scores less than 1 standard deviation (SD) below the mean. The relationships between neurobehaviour at term and Bayley-III cognitive, language, and motor scales at two years were examined using linear regression.

RESULTS Increased odds for cognitive delay were associated with suboptimal HNNE total scores (odds ratio [OR] 2.66; 95% confidence interval [CI] 1.14–6.23, p=0.020) and suboptimal NNNS excitability (OR 3.01; 95% CI 1.33–6.82, p=0.008) and lethargy (OR 4.05; 95% CI 1.75–9.31, p=0.001) scores. Suboptimal lethargy scores on the NNNS were associated with increased odds of language (OR 5.64; 95% CI 1.33–23.85, p=0.019) and motor delay (OR: 6.86; 95% CI 1.64–28.71, p=0.08).

INTERPRETATION Suboptimal performance on specific aspects of newborn neurobehavioural assessments is associated with neurodevelopmental delay at two years in children born MLPT.

[First page footer]

DOI:

What this paper adds

- Poor performance on the lethargy and excitability scales of the NICU Network Neurobehavioral Scale (NNNS) is associated with worse cognitive outcomes at two years in infants born moderate-to-late preterm (MLPT).
- Poor performance on the Hammersmith Neonatal Neurological Examination is also associated with worse cognitive outcomes at two years in infants born MLPT.
- Poor performance on the NNNS lethargy scale is associated with motor and language delay at two years in infants born MLPT.

[Main text]
Worldwide there is a growing number of infants born preterm (<37 weeks' gestation), who account for 11% of all births.[1] Lower gestational age at birth is related to cognitive, behaviour, and motor impairments during childhood, with the severity and range of impairments increasing as gestational age at birth decreases.[2] The functional implications of these impairments include poorer school performance and increasing need for special education.[3] While the incidence of neurodevelopmental impairments in infants born moderate-to-late preterm (MLPT; born 32 to 36+6 weeks' gestation) is lower than peers born very preterm (<32 weeks' gestation), the rate of MLPT birth is much greater.[1] Thus, MLPT birth potentially places a significant burden on the health system, not only in the neonatal period but also throughout childhood.[1,3]

At term-equivalent age, infants born MLPT have poorer brain growth and maturation, as seen via magnetic resonance imaging (MRI), compared with peers born at term.[4–6] Despite these differences in neuroimaging findings between MLPT and term-born infants, rates of periventricular cysts, intraventricular haemorrhage, and

ventricular dilatation, which are predictive of cerebral palsy and other adverse neurodevelopmental outcomes, have been reported to be low in both groups.[4] It is not feasible to offer neuroimaging to all MLPT infants given their large numbers; therefore, it is important to have other clinical indicators to identify those most at risk.

Neonatal neurobehavioural assessments provide opportunity for parental education, targeting of follow-up services, and referral to early intervention.[5] Several reviews have identified the Hammersmith Neonatal Neurological Examination (HNNE)[6] and the Neonatal Intensive Care Unit (NICU) Network Neurobehavioral Scale (NNNS)[7] as being valid and reliable assessments of neurobehaviour in the newborn period.[5,8] Compared with peers born at term, children born MLPT have less trunk and leg flexor tone, and poorer head control and quality of movement shortly after birth and/or at term-equivalent age on the HNNE.[9–11] On the NNNS, MLPT infants have poorer arousal, regulation, lethargy, and quality of movement, and higher rates of non-optimal reflexes, stress, and hypotonicity.[11,12] However, it is unclear whether these early neurobehavioural alterations seen in MLPT infants, who are considered at low risk of adverse neurodevelopmental outcomes compared with infants born very preterm,[9] are associated with long-term neurodevelopmental delays. Therefore, the aim of this study was to examine the relationship between neurobehaviour at term-equivalent age and neurodevelopmental outcomes at two years in a prospective longitudinal cohort of infants born MLPT.

METHODS

Participants

Between November 2009 and November 2012, 201 infants born MLPT (born at 32 to 36+6 weeks' gestation) were recruited from the Royal Women's Hospital, Melbourne, Australia.[4,11] Having one parent who could speak English (funding was not available for interpreters) was an inclusion criterion. Infants were excluded if they had a congenital abnormality known to affect neurodevelopmental outcomes. The Human Research Ethics Committees of the Royal Women's and the Royal Children's Hospitals in Melbourne approved this study and parents provided written informed consent for their child to participate in the study.

Perinatal data were recorded during the neonatal period by a research nurse including gestation at birth, sex, birthweight standard deviation (SD) z-score (calculated according to gestational age and sex using the British Growth Reference

norms),[13] multiple birth, use of antenatal corticosteroids, and whether the infant received respiratory support. The Social Risk Index,[14] comprising six aspects of social status (family structure, education of primary caregiver, occupation and employment status of primary income earner, language spoken at home, and maternal age at birth), was collected and used to categorize the infants as higher (>1) or lower social risk (0 or 1).

Term-equivalent age assessments

At term-equivalent age (38–44 weeks' postmenstrual age) participants were asked to return for a follow-up visit at the Royal Children's Hospital where they completed the HNNE and the NNNS. These two standardized neurobehavioural assessments were chosen because they provide complementary information regarding the infant's neurodevelopment and could be completed together within 30 minutes.[5,8,11] One of five trained assessors who were unaware of the infant's previous assessments and masked to the infant's clinical history administered the NNNS first, followed by the additional items needed to complete the HNNE scoring (tendon reflexes, arm and leg traction). All assessors were certified in the NNNS examination and had done training on the HNNE before the commencement of the study. Inter-rater reliability for both assessments was excellent (intra class correlation: ICC>0.80) for all subscales, except for HNNE tone patterns (ICC=0.54) and NNNS asymmetrical reflexes (ICC=0.67).[15] Infants were assessed in a standardized manner according to the assessment procedures, with the infants minimally dressed in a warm, quiet room with the parents or caregivers present.

Hammersmith Neonatal Neurological Examination (HNNE)

The HNNE is a neurological examination that was developed for use with both term and preterm infants in the newborn period.[6] It consists of six subscales (tone, tone patterns, reflexes, spontaneous movements, abnormal neurological signs, and behaviour) and a total score. There are 34 items, with a maximum possible of one point per item, which are totaled to provide an overall optimality score, as well as the six subscale scores. The overall optimality score was designed primarily for use in research and originally was validated on 224 low-risk term infants from England assessed between 6 and 48 hours after birth.[6] Our research team has published optimality scores based upon 200 healthy children born at term and assessed between

38 to 44 weeks' gestation using the same scoring principles.[11] These local optimality scores were used to identify participants with 'suboptimal' scores in the current study (defined as <10th centile), as the original cut-points created by Dubowitz et al. were based on infants within the first 48 hours after birth,[6] and may not be directly applicable to our population, who were assessed much later. Thus, local norms are recommended when possible.[16]

Neonatal Intensive Care Unit Network Neurobehavioral Scale (NNNS)

The NNNS assesses neurological function, behavioural organization, and stress responses in high-risk infants.[7] It consists of 45 items administered to the infant in a state-dependent standardized sequence, along with observations of the infant's behavioural responses during the evaluation, totaling 115 items overall. These items correspond to 13 summary scales, of which we used 12, with the habituation scale excluded because the majority of infants were not in the appropriate behavioural state (i.e. not asleep) at the commencement of the assessment. Higher scores on each of the scales represent more of the construct being measured, with higher scores for the attention, quality of movement, and regulation scales representing better neurobehavioural performance, while higher scores on handling, non-optimal reflexes, asymmetrical reflexes, hypotonicity, hypertonicity, excitability, lethargy, arousal, and the stress/abstinence scale represent poorer neurobehavioural performance.[17] Infants' performance on each scale was classified as 'optimal' or 'suboptimal' based on the same term cohort used for the HNNE classification. 'Suboptimal' classification was given to scores lower than the 10th centile on the NNNS attention, quality of movement, and regulation scales, and to scores greater than the 90th centile for handling, non-optimal reflexes, asymmetrical reflexes, hypotonicity, hypertonicity, excitability, lethargy, arousal, and the stress/abstinence scales.

Outcome measures

Bayley Scales of Infant and Toddler Development – 3rd edition

At two years' corrected age, children were invited for a further follow-up appointment at the Royal Women's or Royal Children's Hospital, Melbourne, where neurodevelopmental outcome was assessed with the Bayley Scales of Infant and Toddler Development 3rd Edition (Bayley-III) by a psychologist trained in the

Bayley-III and unaware of the child's perinatal history. The Bayley-III is a valid, reliable, and standardized measure of developmental delay.[18] It comprises three index scores: (1) the cognitive index which provides an estimate of general cognitive development based on nonverbal activities involving memory, problem solving, and manipulation; (2) the language composite which has expressive and receptive subscales; and (3) the motor composite which has gross motor and fine motor subscales. Age standardized scores were calculated for each index using the test manual, with local reference data used to classify children as having developmental delay, defined as scores more than 1 SD below the mean (cognitive scale: mean 108.9, SD 14.3; language scale: mean 108.2, SD 14.8; motor scale: mean 118.4, SD 16.7).[19] Local reference data were obtained from 220 term-born infants and were used to define delay in the current study because of concerns with the Bayley-III test norms underestimating developmental delay.[16,19]

Statistical analysis

Data were analyzed using Stata version 14 (StataCorp, Texas, USA). The relationships between HNNE (6 subscales and the total score) and NNNS (12 summary scales) scores at term, and Bayley-III cognitive, language composite, and motor composite scores at two years, were examined using linear regression fitted using generalized estimating equations to allow for clustering of multiple births. Logistic regression, also fitted using generalized estimating equations, was used to examine the association between being classified as suboptimal on the HNNE and NNNS and having delay on the Bayley-III cognitive, language, or motor scales at two years. Results are reported as regression coefficients or odds ratios (OR) and their 95% confidence intervals (CI). All analyses were repeated, adjusting for age at term assessment, social risk status, and sex as potential confounders, because these variables have been shown to be independently related to neurobehaviour.[9,11,17]

RESULTS

A total of 201 infants born MLPT were recruited during the first week after birth, with neurobehavioural assessments attempted for all infants. There were 1789 infants born MLPT admitted to the Royal Women's Hospital, Melbourne during the recruitment period. We did not approach all infants because we had to limit the rate of recruitment as a result of workforce issues, including accessibility to the MRI scanner

at term-equivalent age. Of the 201, 187 infants had complete data for both the HNNE and NNNS (Table I). The perinatal characteristics were similar between infants who did and did not have complete HNNE and/or NNNS data, with the exception that gestational age at time of the term assessment was higher for those with missing data. Of the 14 infants who did not have a complete neurobehavioural assessment, one infant was too unsettled to commence the assessment and the remaining had missing data (due to the infant being in an inappropriate behavioural state to assess or coding error).

At 24 months, 197 children (98% follow-up) were assessed with the Bayley-III; 197, 190, and 194 completed the cognitive, language, and motor scales respectively. Of note, the language scale was not administered on infants who did not speak English at home. Two infants had cerebral palsy, one with hemiplegia (Gross Motor Function Classification System [GMFCS][20] level I) and one with quadriplegia (GMFCS level III).

The means and SD for the HNNE, NNNS, and Bayley-III subscales are reported in Table II. Overall, 25% of MLPT infants scored suboptimally on the HNNE total score, with 31% suboptimal for excitability, 20% for lethargy, 26% for regulation, 31% for arousal, and 36% for stress on the NNNS at term-equivalent age. At two years, 19%, 36%, and 42% of those assessed scored more than 1 SD below the mean, and 6%, 4% and 8% scored more than 2 SD below the mean for cognitive, language, and motor scales respectively.

Cognitive outcomes

There was evidence of relationships between continuous scores for several subscales and the total score for the HNNE (Table III) and several NNNS summary scales (Table IV) with outcomes on the Bayley-III cognitive scale using continuous scores. Better (higher) scores for spontaneous movements, abnormal signs, behaviour, and total score on the HNNE and attention and regulation scales on the NNNS were associated with better (higher) cognitive scores, whilst poorer scores for the NNNS excitability, lethargy, and arousal summary scales were associated with lower (worse) cognitive scores. After adjusting for age at time of assessment, social risk, and sex, the direction of the results remained the same; however, the magnitude and strength of evidence for the relationships decreased. After adjustment, there was evidence that the HNNE abnormal signs (co-efficient: 3.00; 95% CI 0.21–5.79, p=0.035), behaviour

(co-efficient: 1.40; 95% CI 0.01–2.79, p=0.048), and total score (co-efficient: 0.62; 95% CI 0.01–1.22, p=0.045) and NNNS attention (co-efficient: 1.52; 95% CI 0.24–2.80, p=0.020), self-regulation (co-efficient: 2.52; 95% CI 0.70–4.34, p=0.007), and lethargy (co-efficient: –1.17; 95% CI –1.89 to –0.44, p=0.002) summary scales were correlated with the cognitive scores.

Suboptimal scores for the HNNE spontaneous movements (OR 3.13; 95% CI 1.29–7.6, p=0.012), abnormal signs (OR: 2.58; 95% CI 1.19–5.61, p=0.016) subscales and HNNE total scores (OR: 2.98; 95% CI 1.42–6.27, p=0.004) were associated with increased odds of cognitive delay on the Bayley-III (Fig. 1). The relationships were generally weaker after adjustment, with only the HNNE total scores independently related to cognitive outcome (OR 2.66; 95% CI 1.14–6.23; p=0.024). The tone patterns subscale was excluded from Figure 1, as there were no children with suboptimal tone patterns with cognitive, language, or motor delay.

There was evidence of a relationship between suboptimal NNNS scores for excitability (OR 3.42; 95% CI 1.67–7.02, p=0.001), regulation (OR 5.15; 95% CI 2.47–10.75, p<0.0001), and arousal (OR 2.14; 95% CI 1.04–4.40, p=0.038) and increased odds of cognitive delay (Fig. 2). After adjustment, only the NNNS excitability (OR 3.01; 95% CI 1.33–6.82, p=0.008) and regulation (OR 4.05; 95% CI 1.75–9.31, p=0.001) scales remained associated with increased odds of cognitive delay.

Language outcomes

The only evidence for relationships with continuous scores of the Bayley-III language composite was for the non-optimal reflexes scale of the NNNS (Table IV). There was evidence of this relationship (co-efficient: 0.92; 95% CI 0.22–1.62, p=0.010) after adjustment for confounders. Suboptimal performance on the lethargy scale of the NNNS was associated with increased language delay (OR 5.64; 95% CI 1.33–23.85, p=0.019). There was little evidence of relationships between the HNNE and the language composite of the Bayley-III.

Motor outcomes

There was little evidence of relationships between the HNNE and Bayley-III motor composite. For the NNNS continuous scores, there were no relationships with the motor composite on the Bayley-III for both unadjusted and adjusted analyses.

However, suboptimal performance on the lethargy scale of the NNNS was associated with motor delay (OR 6.86; 95% CI 1.64–28.71, p=0.008) with the relationship remaining similar after adjustment.

Discussion

This study has shown that neurobehaviour at term-equivalent age, assessed with the HNNE and NNNS, is associated with cognitive development assessed with the Bayley-III at two years in infants born MLPT. Higher (better) HNNE total scores are associated with higher cognitive scores on the Bayley-III, whilst those infants who have a HNNE total score in the suboptimal range on the HNNE have three times higher odds of having a cognitive delay, even after adjusting for age at assessment, social risk, and sex. The behaviourally focused scales on the NNNS, including the attention, regulation, and lethargy scales, were also associated with cognitive development at two years, with suboptimal excitability and regulation scale scores associated with at least a three times higher odds of cognitive delay, after adjusting for important predictors of outcome.[11,17]

The relationships between early neurobehavioural assessments and language development at two years are less clear. Whilst there was evidence of an association between the NNNS non-optimal reflexes scale and the language composite, the direction of the relationship was in the opposite direction to that expected. This could be due to type-I error with multiple analyses performed. Furthermore, there was little evidence of this relationship when the cut-off of the 90th centile was used. The lethargy subscale of the NNNS was related to both motor and cognitive delay. Infants with suboptimal lethargy scores on the NNNS have poor orientation, lower muscle tone, poor alertness, and poor head control with pull to sit during their newborn examination. These are important components of a clinical assessment and a key finding of this study is that these patterns of behaviour are related to increased risk of motor and language impairment.

There was little evidence of any other relationships between the HNNE and NNNS at term and motor development at two years in infants born MLPT. This is most likely due to the low rates of major brain abnormalities on neuroimaging reported in this cohort.[4] Further, we have classified motor delay using a cut-off of 1 SD below the mean, as this is a relatively low-risk cohort. Stephens et al. found the NNNS handling, quality of movement, hypotonicity, and stress scales were associated

with motor scores on the 2nd edition of the Bayley Scales using a cut-off of more than 2 SD below the mean in a sample of 395 infants born before 36 weeks, of which 5% had cerebral palsy.[21]

We found that only 8% of infants born MLPT had suboptimal behaviour on the HNNE at term-equivalent age, which is lower than the expected rate of 10% in the term-born population.[11] Whilst it may be expected that infants born MLPT would perform worse than their peers born at term, these findings are consistent with research involving infants born preterm compared with those born at term that has reported greater variability in visual behaviour at term-equivalent age, with more preterm infants able to follow in a full circle than those born full-term. These findings are most likely because of infants born MLPT having more opportunities for visual experiences compared with infants born at term assessed at the same gestational age.[9] Nonetheless, our study has shown that infants born MLPT who have better behaviour at term-equivalent age on the HNNE or in regulation and attention on the NNNS have better cognitive scores, whilst poorer behaviour on the NNNS lethargy, excitability, and stress scales is associated with lower cognitive scores.

The NNNS requires specialized training by the administrator and takes approximately 20 to 30 minutes to administer, plus additional time for scoring. It was initially designed for the large multisite 'Maternal Lifestyle Study', which required an assessment with excellent reliability and sensitivity to detect differences in a range of behavioural domains.[7] Whilst the training and administration time may limit the clinical utility of the tool in some settings, it offers an opportunity to assess an infant's neurobehaviour in detail, which can be important for tailoring interventions for those at risk. Several studies of very preterm infants have shown that some domains of early neurobehaviour assessed with the NNNS, including attention, stress, lethargy, hypertonicity, and non-optimal reflexes, are affected by the NICU environment, stress, or developmental support in infants born very preterm.[22] Further research on interventions with a focus on the environment, stress reduction, and developmental care are warranted in infants born MLPT, given that early neurobehaviour on the NNNS is associated with outcome at two years.

The HNNE can be completed in approximately 10 minutes and requires no formal certification process, although training is recommended with an experienced user and a manual is available.[6] The HNNE was designed to be used by health professionals in routine clinical evaluation of newborn infants and therefore needs to

be both quick and easy to administer and score.[5] Our study has demonstrated that the optimality scores are useful not only when used within the first 48 hours as previously reported,[6] but also when used from 38 to 44 weeks' gestation in infants born MLPT. Our study has shown that the total score of the HNNE, rather than subscale scores alone, was most strongly related to cognitive outcome. This is consistent with previous findings using the HNNE where abnormalities in neonatal neurological examinations are often global and non-specific, with isolated findings such as abnormal reflexes not related to outcome.[9,23]

The strengths of our study include the use of a longitudinal prospective cohort with a local normative data set used for classification of suboptimal performance, along with excellent follow-up rates.[11,19] All assessors were masked to the child's clinical history, which has been a limitation of previous research. However, our study does have some limitations. Our study has purposefully not examined the specificity or sensitivity of the HNNE and NNNS, because it has been recommended that they are not used as diagnostic tools in clinical practice. Rather, these assessments offer the opportunity to identify those at increased risk of later neurodevelopmental impairments, and to target those infants born MLPT who require further follow-up.[5] Further, we have conducted multiple comparisons, which increases the risk of chance findings. It would be beneficial to follow up these children longer term, as the predictive validity of the Bayley-III is moderate for school-age outcomes.[16,24]

Our study has important implications for both clinical and research practice, as we have shown that two neurobehavioural assessment tools, the HNNE and NNNS, are associated with cognitive outcomes at two years in MLPT infants. The standardized clinical examinations considered in this study are relatively low-cost (in comparison with MRI), non-invasive methods for identifying infants at risk of cognitive delay. Early detection of MLPT infants with developmental delay is important to facilitate early intervention and optimize outcomes for this vulnerable but sizeable population.[25]

ACKNOWLEDGEMENTS

This work was supported in part by the Australian National Health and Medical Research Council (NHMRC) (Project Grant ID 1028822); Centre of Clinical Research Excellence Grant ID 546519; Centre of Research Excellence Grant ID 1060733; Senior Research Fellowship ID 1081288 to PJA.; Early Career Fellowship

ID 1053787 to JLYC, ID 1053767 to AJS; Career Development Fellowship ID 1108714 to AJS and ID 1053609 to KJL; Australian Postgraduate Scholarship to JEO, Murdoch Children's Research Institute, Clinical Sciences Theme Grant, the Victorian Government Operational Infrastructure Support Program. The authors have stated that they had no interests which may be perceived as posing a conflict or bias.

REFERENCES

1. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012; **379**: 2162–72.

2. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet* 2008; **371**: 261–9.

3. Chan E, Leong P, Malouf R, Quigley MA. Long-term cognitive and school outcomes of late-preterm and early-term births: a systematic review. *Child Care Health Dev* 2016; **42**: 297–312.

4. Walsh JM, Doyle LW, Anderson PJ, Lee KJ, Cheong JL. Moderate and late preterm birth: effect on brain size and maturation at term-equivalent age. *Radiology* 2014; **273**: 232–240.

5. Wusthoff CJ. How to use: the neonatal neurological examination. *Arch Dis Child Educ Pract Ed* 2013; **98**: 148–53.

6. Dubowitz L, Mercuri E, Dubowitz V. An optimality score for the neurologic examination of the term newborn. *J Pediatr* 1998; **133**: 406–16.

7. Lester BM, Tronick EZ, Brazelton TB. The Neonatal Intensive Care Unit Network Neurobehavioral Scale procedures. *Pediatrics* 2004; **113**: 641–67.

8. Brown N, Spittle A. Neurobehavioral evaluation in the preterm and term infant. *Curr Pediatr Rev* 2014; **10**: 65–72.

9. Romeo DM, Luciano R, Corsello M, et al. Neonatal neurological examination of late preterm babies. *Early Hum Dev* 2013; **89**: 537–45.

10. Romeo DM, Ricci D, Brogna C, et al. Neurological examination of late-preterm infants at term age. *Eur J Paediatr Neurol* 2011; **15**: 353–60.

11. Spittle AJ, Walsh J, Olsen JE, et al. Neurobehaviour and neurological development in the first month after birth for infants born between 32–42 weeks' gestation. *Early Hum Dev* 2016; **96**: 7–14.
12. Barros MC, Mitsuhiro S, Chalem E, Laranjeira RR, Guinsburg R. Neurobehavior of late preterm infants of adolescent mothers. *Neonatology* 2011; **99**: 133–9.
13. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat Med* 1998; **17**: 407–29.
14. Roberts G, Howard K, Spittle AJ, Brown NC, Anderson PJ, Doyle LW. Rates of early intervention services in very preterm children with developmental disabilities at age 2 years. *J Paediatr Child Health* 2008; **44**: 276–80.
15. Eeles AL, Olsen JE, Walsh JM, et al. Reliability of neurobehavioral assessments from birth to term equivalent age in preterm and term born infants. *Phys Occup Ther Pediatr* 2016 Mar 22: 1–11 [Epub ahead of print].
16. Spencer-Smith MM, Spittle AJ, Lee KJ, Doyle LW, Anderson PJ. Bayley-III Cognitive and language scales in preterm children. *Pediatrics* 2015; **135**: e1258–65.
17. Fink NS, Tronick E, Olson K, Lester B. Healthy newborns' neurobehavior: norms and relations to medical and demographic factors. *J Pediatr* 2012; **161**: 1073–79.
18. Bayley N. Bayley Scales of Infant and Toddler Development. San Antonio, TX: The Psychological Corporation, 2005.
19. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW, Victorian Infant Collaborative Group. Underestimation of developmental delay by the new Bayley-III Scale. *Arch Pediatr Adolesc Med* 2010; **164**: 352–56.
20. Palisano RJ, Hanna SE, Rosenbaum PL, et al. Validation of a model of gross motor function for children with cerebral palsy. *Phys Ther* 2000; **80**: 974–85.
21. Stephens BE, Liu J, Lester B, et al. Neurobehavioral assessment predicts motor outcome in preterm infants. *J Pediatr* 2010; **156**: 366–71.
22. Smith GC, Gutovich J, Smyser C, et al. Neonatal intensive care unit stress is associated with brain development in preterm infants. *Ann Neurol* 2011; **70**: 541–49.

23. Wusthoff C. How to use: the neonatal neurological examination. *Arch Dis Child Educ Pract Ed* 2013; **98**: 148–53.
24. Spittle AJ, Spencer-Smith MM, Eeles AL, et al. Does the Bayley-III Motor Scale at 2 years predict motor outcome at 4 years in very preterm children? *Dev Med Child Neurol* 2013; **55**: 448–52.
25. Spittle A, Orton J, Anderson PJ, Boyd R, Doyle LW. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. *Cochrane Database Syst Rev* 2015; **11**: CD005495.

Table I: Characteristics of the study sample

Perinatal characteristics	**Study sample *n*=197**	**Infants with missing data *n*=14**
Birthweight (g) – mean (SD)	2161 (463)	2160 (319)
Gestational age (weeks) – mean (SD)	34.4 (1.2)	34.2 (0.9)
Male gender – *n* (%)	96 (48%)	10 (63%)
Multiple birth – *n* (%)	74 (37%)	8 (50%)
Antenatal corticosteroids – *n* (%)	121 (60%)	7 (44%)
Respiratory distress at birth – *n* (%)	28 (14%)	3 (19%)
Small for gestational age – *n* (%)	20 (10%)	1 (7%)
Higher social risk – *n* (%) *n*=195	66 (34%)	6 (38%)
Gestational age at term assessment, weeks– mean (SD) *n*=200	41.5 (1.4)	42.2 (1.9)
Gestational age at 24 month assessment, months– mean (SD) *n*=197	25.3 (1.3)	25.3 (1.3)

SD, standard deviation.

Table II: Means, standard deviations, and percentage of infants with suboptimal assessments at term-equivalent and two years of age

Term Assessment		***n***	**Mean**	**SD**	**Suboptimal – *n* (%)**
HNNE	Tone	197	8.52	1.40	53 (27)
	Tone patterns	197	4.42	0.47	15 (8)
	Reflexes	193	5.53	0.63	17 (9)
	Spontaneous movements	194	2.43	0.64	27 (14)
	Abnormal signs	196	2.53	0.57	46 (23)
	Behaviour	194	6.17	0.97	14 (7)
	Total	189	29.7	2.64	48 (25)
NNNS	Attention	179	5.52	1.18	17 (10)
	Handling	182	0.59	0.29	14 (7)
	Excitability	187	4.20	2.59	58 (31)
	Lethargy	187	3.24	1.83	37 (20)

	Regulation	187	5.28	0.88	51 (26)
	Arousal	187	4.42	0.63	58 (31)
	Stress	187	0.15	0.07	68 (36)
	Quality of movement	187	4.69	0.73	50 (27)
	Hypertonicity	187	0.10	0.38	14 (7)
	Hypotonicity	187	0.14	0.40	24 (13)
	Reflexes – asymmetrical	187	1.04	1.12	19 (10)
	Non-optimal reflexes	187	4.48	2.25	16 (9)
Two year assessment		***n***	**Mean**	**SD**	**Delay – *n* (%)**
Bayley	Cognitive	196	102.5	13.9	37 (19)
	Language	190	99.1	17.7	68 (36)
	Motor	194	103.5	16.0	82 (42)

SD, standard deviation; HNNE, Hammersmith Neonatal Neurological Examination; NNNS, NICU Network Neurobehavioral Scale.

Table III: Associations between Hammersmith Neonatal Neurological Examination subscale and total scores at term-equivalent age and Bayley-III cognitive, language, and motor composite scores at 2 years of age

	Cognitive			Language			Motor		
Subscale	***n***	**Coeff**	**95% CI**	***n***	**Coeff**	**95% CI**	***n***	**Coeff**	**95% CI**
Tone	193	0.57	–1.20, 2.34	186	1.64	–1.54,4.83	190	1.07	–0.73, 2.88
Tone patterns	193	–0.54	–4.43, 3.35	186	–1.80	–9.78, 6.17	190	–1.49	–7.07, 4.10
Reflexes	189	1.86	–1.20, 4.85	182	0.34	–3.52, 4.18	186	0.60	–2.65, 3.85
Spontaneous movements	190	2.99	**0.37, 5.60**	183	–0.88	–4.38, 2.62	187	–2.20	–4.85, 0.44
Abnormal signs	192	3.54	**0.23, 6.85**	185	–0.76	–5.39, 3.87	189	2.21	–1.96, 6.38
Behaviour	190	1.98	**0.49, 3.46**	183	–1.30	–2.92, 3.12	187	–0.21	–2.59, 2.17
Total	185	0.79	**0.07, 1.52**	178	0.23	–0.88, 1.36	182	0.19	–0.60, 0.98

Results presented are from unadjusted linear regression models with higher scores representing better performance; SD, standard deviation; Coeff, coefficient; CI, confidence interval; bold text refers to *p* values <0.05.

Table IV: Associations between NICU Network Neurobehavioral Scale scores at term-equivalent age and Bayley-III cognitive, language and motor composite scores at 2 years of age

	Cognitive			Language			Motor		
Subscales	***n***	**Coeff**	**95% CI**	***n***	**Coeff**	**95% CI**	***n***	**Coeff**	**95% CI**
Attention[a]	176	1.92	0.59, 3.25	171	0.55	–1.39, 2.48	173	1.17	–0.82, 3.16
Handling[b]	179	–3.14	–9.5, 3.25	174	–2.58	–10.47, 5.30	176	–6.11	–15.4, 3.20
Excitability[b]	184	–0.69	**–1.29, –0.09**	179	0.11	–0.67, 0.90	181	–0.72	–1.83, 0.38
Lethargy[b]	184	–1.24	**–1.96, –0.53**	179	–0.41	–1.42, 0.60	181	–0.28	–1.37, 0.82
Regulation[a]	184	3.38	**1.52, 5.25**	179	0.55	–1.93, 3.04	184	3.36	–0.01, 6.72
Arousal[b]	184	–2.91	**–5.5, –0.30**	179	0.45	–3.31, 4.22	181	–2.56	–7.07, 1.94
Stress[b]	184	–7.97	–32.59, 26.67	179	16.98	–6.91, 40.87	181	–23.32	–65.77, 19.13
Quality of movement[a]	184	2.38	–0.37, 5.12	179	–0.89	–5.41, 3.63	181	1.05	–2.34, 4.44
Hypertonicity[b]	184	0.81	–4.32, 5.94	179	3.12	–1.56, 7.80	181	2.50	–1.70, 6.70
Hypotonicity[b]	184	0.57	–4.15, 5.29	179	0.78	–5.11, 6.67	181	2.23	–4.28, 8.78
Reflexes – asymmetrical[b]	184	0.93	–1.15, 3.00	179	1.48	–0.49, 3.45	181	0.61	–1.62, 2.83
Non–optimal reflexes[b]	184	–0.24	–0.9, 0.4	179	0.74	**0.06, 1.41**	181	0.50	–0.16, 1.17

Results presented are from unadjusted linear regression models with [a]higher scores representing better performance and [b]higher scores representing poorer performance. SD, standard deviation; Coeff, coefficient; CI, confidence interval; bold text refers to *p* values <0.05.

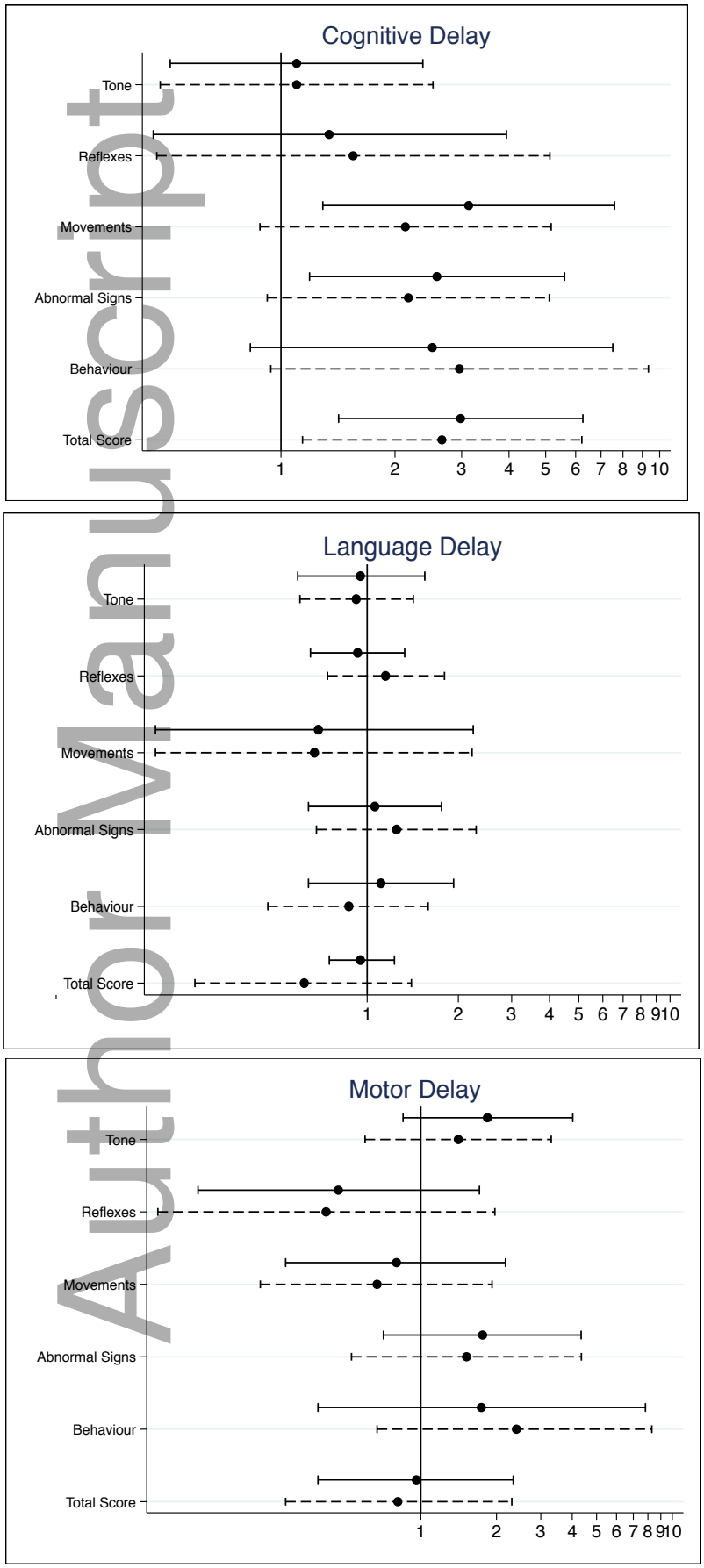
Figure 1: Odds ratio for the association between being suboptimal on the Hammersmith Neonatal Neurological Examination (HNNE) and delayed outcome on the Bayley-III at 2 years

Results are odds ratios and 95% confidence intervals from separate logistic regression models presented on a log scale.
<Straight line> Unadjusted <dotted line> Adjusted for sex, age at assessment, and social risk.

Figure 2: Odds ratio for the association of Neonatal Intensive Care Unit Network Neurobehavioral Scale (NNNS) suboptimal summary scale scores and delayed outcome on the Bayley-III at 2 years

Results are odds ratios and 95% confidence intervals from separate logistic regression models presented on a log scale.
<Straight line> Unadjusted <dotted line> Adjusted for sex, age at assessment, and social risk

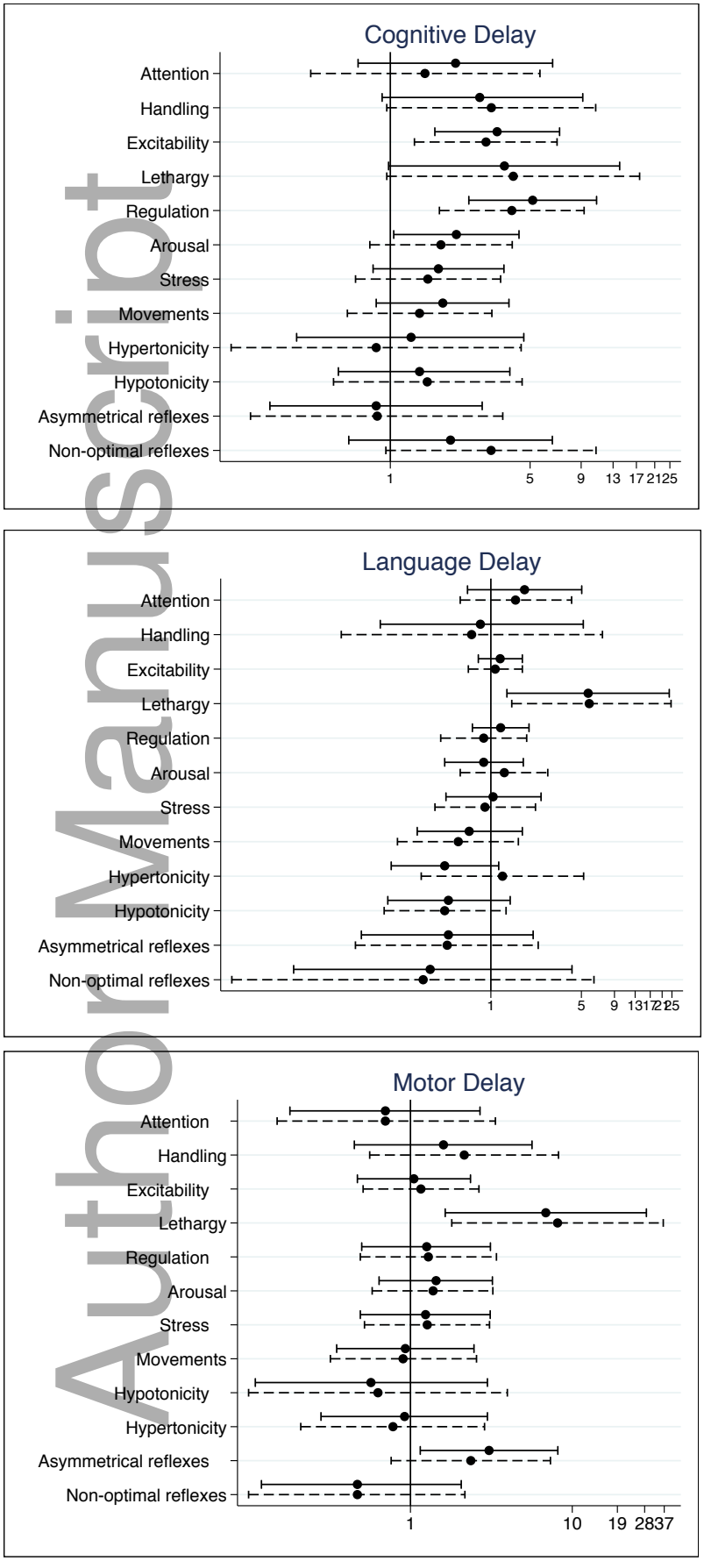
Figure 1. Odds Ratio for the association between being sub-optimal on the Hammersmith Neonatal Neurological Examination (HNNE) and delayed outcome on the Bayley-III at 2 years



Results are odds ratios and 95% confidence intervals from separate logistic regression models presented on a log scale

Unadjusted Adjusted for sex, age at assessment and social risk

Figure 2. Odds Ratio for the association of Neonatal Intensive Care Unit Network Neurobehavioral Scale (NNNS) sub-optimal summary scale scores and delayed outcome on the Bayley-III at 2 years



Results are odds ratios and 95% confidence intervals from separate logistic regression models presented on a log scale

Unadjusted Adjusted for sex, age at assessment and social risk