ACTA PÆDIATRICA
NURTURING THE CHILD

Acta Pædiatrica ISSN 0803-5253

SHORT COMMUNICATION

Cumulative deviance scores can be used as an alternative to the Hammersmith Neonatal Neurological Examination in scientific research

Susanna Stjerna[1], Aulikki Lano[2], Marjo Metsäranta[3], Sampsa Vanhatalo (sampsa.vanhatalo@helsinki.fi)[1]

1.Department of Clinical Neurophysiology, HUS Medical Imaging Center, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland
2.Department of Child Neurology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland
3.Department of Pediatrics, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

Keywords
Correlation studies, Neonatology, Neurological assessment, Newborn, Outcome

Correspondence
Sampsa Vanhatalo, BABA center, Children's Castle Hospital, Helsinki University Central Hospital, PO Box 280, FIN-00029 HUS, Helsinki, Finland.
Email: sampsa.vanhatalo@helsinki.fi

Received
26 September 2014; revised 28 April 2015; accepted 2 June 2015.

DOI:10.1111/apa.13066

Newborn neurological assessments for clinical and research purposes are commonly performed using structured scales with numerous complementary test items (1–4). Nowadays, the widely used Hammersmith Neonatal Neurological Examination has become important in identifying neurological abnormalities or developmental risks. It was designed for easy and rapid detection of neurological deviances and for following up early neurological progress (3). The Hammersmith scores were later quantified for research purposes, and it was shown early on that the number of obvious abnormalities within the Hammersmith Neonatal Neurological Examination items related to compromised development at a later date (5,6). However, statistical use of the Hammersmith scores was not trivial (7). The concept of optimality was later converted into the now widely used optimality score, based on reference population distributions and the 10th and 5th percentiles of the scores within individual items (8–11).

Using optimality scores is practical in clinical research, because it allows a rapid and systematic profiling of the newborn infant by providing a nearly dichotomic assessment at the level of individual test items. As stated above, optimality scores may also have prognostic value, especially with respect to the most severe outcomes (5,6). In this context, it is potentially interesting to note that the direction of neurological deviance, such as the presence of hypotonia as opposed to hypertonia, is developmentally meaningful (12). We decided to create an alternative scoring method for research purposes, to better account for quality of deviance from the expected, normative Hammersmith Neonatal Neurological Examination results. We also reasoned that using neurological assessment scores in correlative studies, together with other research methods such as neuroimaging or neurophysiology, would be of great benefit in preserving more of the original interindividual variability, especially within the subnormal–normal range. It is notable that both of these ideas were discussed when the assessment method was first introduced (3). To this end, we implemented an alternative scoring of raw Hammersmith scores, hereafter called cumulative deviance, which combines both the magnitude and direction of deviance when the infant is compared to optimal performance.

To assess the potential advantages of our alternative cumulative deviance score, compared to the existing optimality score, we used two benchmarking criteria to indicate an improvement. First, we aimed to preserve more of the interindividual variability present in the original Hammersmith scores. Second, we reasoned that a better scoring system could even predict later neurodevelopmental outcome. Finally, we wanted to demonstrate that other factors or covariates, such as gestational age, could be successfully applied as independent variables when using cumulative deviance scores.

To pilot our cumulative deviance scores, we used a previously published cohort (13) that included 43 preterm and 16 full-term infants, who had undergone the Hammersmith Neonatal Neurological Examination at term-equivalent age, as well as the Griffiths Mental Developmental Scales (14) at a corrected age of two years. For this study, we only selected 38 of the 43 preterm infants, as they had data from all the items in both examination scales: 25 were male and their mean gestational age was 26.3 weeks, with a

Table 1 Basic descriptive statistics of Hammersmith scores with cumulative deviance (CD) and optimality scoring (OPT)

HNNE	Scoring	Range	Mean	Median	IQR	SD
Posture	CD	16.00	−5.49	−5.25	5.88	3.94
	OPT	7.50	7.67	8.00	2.00	1.60
Tone	CD	7.00	−0.35	−0.50	2.00	1.50
	OPT	3.00	3.54	4.00	1.00	0.75
Refl	CD	7.00	−0.12	0.00	2.00	1.49
	OPT	4.00	4.63	5.00	1.38	1.04
Mov	CD	6.00	−0.24	0.00	1.88	1.16
	OPT	2.50	2.53	3.00	1.00	0.65
Abn	CD	5.00	−0.10	0.00	1.38	1.08
	OPT	2.00	2.02	2.00	0.00	0.48
OR	CD	9.50	0.40	0.50	3.00	2.13
	OPT	6.00	5.67	6.00	1.50	1.22

IQR, Interquartile range; SD, Standard deviation; Posture, Posture tonus; Tone, Tone patterns; Refl, Reflexes; Mov, Movements; Abn, Abnormal Signs; OR, Orientation and behaviour.

range of 23.4–27.9 weeks. We also selected 15 of the 16 full-term infants: 12 were male and their mean gestational age was 40.1 weeks, ranging from 38.6 to 41.7 weeks. Our cumulative deviance score was computed by subtracting the full-term infant referential values (8) from the observed raw Hammersmith item scores. Doing this will only consider the median of normative data, as it yields both negative and positive values depending on the direction of deviance. Acquiring comparable optimality scores was performed earlier (13), with separate references for full-term (8) and preterm (9) infants. In both the cumulative deviance and optimality scores, the item scores were summed over each of the six Hammersmith main dimensions (Table 1) used for subsequent analyses. Basic descriptive statistics of the acquired cumulative deviance and optimality scores are shown in Table 1.

First, we compared the amount of optimality and cumulative deviance score variability with three measures: range, interquartile range (IQR) and standard deviation (Table 1). As expected from its design, cumulative deviance scoring yielded wider overall and interquartile ranges and larger variance within all of the Hammersmith subscales. Second, we used linear regression analysis to study whether optimality and cumulative deviance scores would show relationships to later neurodevelopmental assessment using the five dimensions of the Griffiths Mental Developmental Scale. Using a total of 30 pairwise comparisons, including six Hammersmith dimensions and five Griffiths Mental Developmental Scale dimensions, we found that none of the Hammersmith dimensions that used optimality scoring were significantly linearly associated with any of the Griffiths Mental Developmental Scale dimensions. However, four of the five Griffiths Mental Developmental Scale domains showed significant pairwise relationships with one or more of the Hammersmith dimensions when alternative cumulative deviance scoring was applied. According to binomial statistics, the chance level of this finding was 6%. Refining and interpreting these associations required more elaborate analysis and thus are not discussed here in detail.

We have recently shown a relationship between Hammersmith orientation and behaviour subscale items and Griffiths visuomotor subscale at two years of corrected age (15), which motivated us to compare optimality and cumulative deviance scores to assess their ability for neurodevelopmental prediction. A pairwise comparison of Griffiths visuomotor subscale scores with optimality and cumulative deviance scores (Fig. 1A and B) only showed a significant relationship with cumulative deviance scores. This finding was significant even within the normal–subnormal outcomes group alone, that is after removing the three infants with Griffiths visuomotor subscale scores

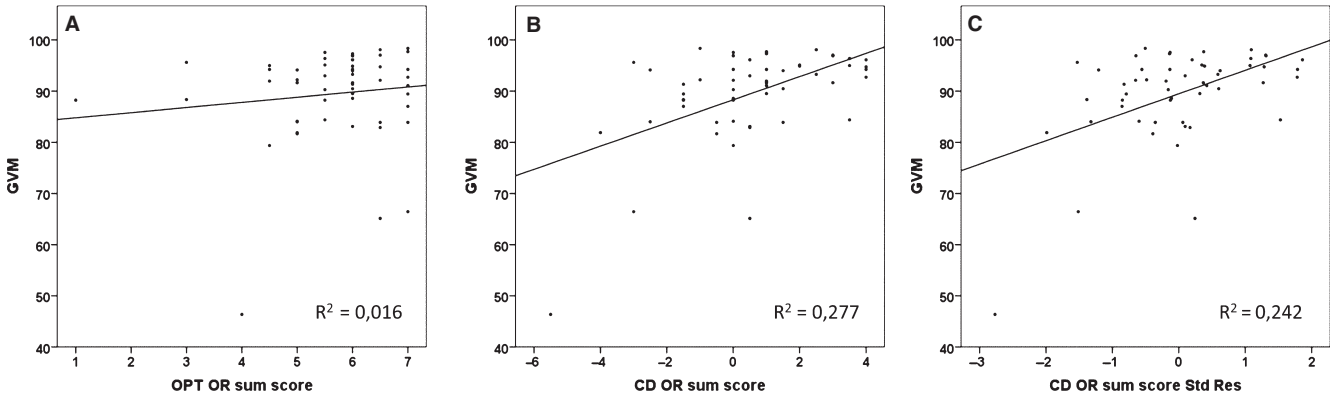

Figure 1 An example of using the cumulative deviance scores in a regression study compared to optimality scores. Each point in figures represents the Hammersmith orientation and behaviour (*X*-axis) and Griffiths visuomotor (*Y*-axis) scores of one individual infant. (A) Optimality scores and Griffiths visuomotor scores have no significant linear relationship. (B) Cumulative deviance scores and Griffiths visuomotor scores have a significant relationship. Notably, most patients are within a normal/subnormal range (Griffiths visuomotor scores >80), and the relationship remains statistically significant for the cumulative deviance scores even after removing the three poorly performing (Griffiths visuomotor scores <70) patients (cumulative deviance score $R^2 = 0.129$ $p = 0.004$ and optimality scores $R^2 = 0.038$ $p = 0.089$). (C) Graph showing the significant relationship of standardised residuals of cumulative orientation and behaviour scores and Griffiths visuomotor scores after the shared variance with gestational age has been removed.

under 70. While this finding is interesting, it is only exploratory and future studies with larger populations are obviously needed to refine the role of the Hammersmith Neonatal Neurological Examination in neurodevelopmental prediction.

In the scientific research context, it is practical to apply the same examination methods for all infants and to treat clinical details, including gestational age, as independent variables in the analyses. Unlike optimality scores, cumulative deviance scoring is independent of gestational age, which can be included in the analysis later, if necessary, using the specific research question. To demonstrate this effect, we adjusted the sum of the Hammersmith orientation and behaviour cumulative deviance compound scores for gestational age with linear regression. Here, gestational age was indeed significantly related to the orientation and behaviour cumulative deviance compound score ($\beta = 0.24$, $t = 2.17$, $p = 0.03$), as expected. The remaining standardised variance of the orientation and behaviour cumulative deviance scores was then added into a regression model to predict the Griffiths visuomotor subscale scores at the corrected age of two years (Fig. 1C), yielding a highly significant association ($\beta = 0.47$, $t = 3.98$, $p < 0.001$). This procedure demonstrated how cumulative deviance scoring also allowed us to study the effect of gestational age separately and shows that the association with later visuomotor outcome was significant in the whole cohort (Fig. 1C).

Taken together, our findings firstly suggest that cumulative deviance scores could be used as an alternative scoring method to the Hammersmith Neonatal Neurological Examination in scientific research. Cumulative deviance scores could also reflect the quality of deviance that would be likely to have pathophysiological significance, as we knew that the range of expected values would become narrower and shift from left to right in the scale as a function of gestational age (8–11) Secondly, cumulative deviance scores yield better differentiation of infants by preserving more of the interindividual variance – wider ranges, larger interquartile ratios and standard deviations – and could thereby enhance the possible detection of significant associations with other assessments or measurements. Thirdly, our implementation of cumulative deviance was based on using the same normative full-term infant references (7) for all infants, which allowed direct comparisons between groups. Cumulative deviance scores also allow us to use independent variables such as gestational age. Finally, our findings suggest that cumulative deviance scores might even offer a way to define the as yet elusive relationship (15) between early neurological assessment and later neurodevelopmental outcome.

FUNDING STATEMENT

The study was supported financially by Academy of Finland grant 253130, Juselius Foundation, Ylppö foundation, Päivikki and Sakari Sohlberg Foundation, Suomen Kulttuurirahasto, and by Foundation for Pediatric Research.

References

1. Prechtl HFR. The neurological examination of full-term newborn infant. *Clinics in developmental medicine*, 2nd ed. London: SIMP/Heinemann, 1977: 63.
2. Amiel-Tison C, Grenier A. *Neurological assessment during the first year of life*. New York: Oxford University Press, 1986.
3. Dubowitz LMS, Dubowitz V, Mercuri E. The neurological assessment of preterm and full-term newborn infant. *Clinics in developmental medicine*, 2nd ed. London: Mac Keith Press, 1999: 148.
4. Brazelton TB, Nugent KJ. Neonatal behavioral assessment scale. *Clinics in developmental medicine*. London: Mac Keith Press, 1995: 137.
5. Dubowitz L, Dubowitz Y, Palmer P, Miller G, Fawer C, Levene M. Correlation of neurological assessment in the preterm newborn infant with outcome of 1 year. *J Pediatr* 1984; 105: 452–6.
6. Molteno C, Grosz P, Wallace P, Jones M. Neurological examination of the preterm and full-term infant at risk for developmental disabilities using the Dubowitz Neurological Assessment. *Early Human Dev* 1995; 41: 167–76.
7. Prechtl HFR. The optimality concept. *Early Human Dev* 1980; 4: 201–5.
8. Dubowitz L, Mercuri E, Dubowitz V. An optimality score for the neurologic examination of the term newborn. *J Pediatr* 1998; 133: 406–16.
9. Mercuri E, Guzzetta A, Laroche S, Ricci D, vanHaastert I, Simpson A, et al. Neurologic examination of preterm infants at term age: comparison with term infants. *J Pediatr* 2003; 142: 647–55.
10. Romeo DM, Ricci D, Brogna C, Cilauro S, Lombardo ME, Romeo MG, et al. Neurological examination of late-preterm infants at term age. *Eur J Paediatr Neurol* 2011; 15: 353–60.
11. Ricci D, Romeo DMM, Haataja L, van Haasert IC, Cesarini L, Maunu J, et al. Neurological examination of preterm infants at term equivalent age. *Early Hum Dev* 2008; 84: 751–61.
12. Prechtl HFR. Neurological findings in newborn infants after pre- and perinatal complications. In Jonxis JHP, Visser HKA, Troelstra JA, editors. *Aspects of prematurity and dysmaturity*. Leiden: Stenfert-Kroese, 1968: 303–21.
13. Nevalainen P, Rahkonen P, Pihko E, Lano A, Vanhatalo S, Andersson S, et al. Evaluation of somatosensory cortical processing in extremely preterm infants at term with MEG and EEG. *Clin Neurophysiol* 2014; 25: 1388–2457.
14. Brandt I, Sticker EJ. *Griffiths Entwicklungsskalen zur Beurteilung der Entwicklung in den ersten beiden Lebensjahren*. Göttingen: Beltz Test GmbH, 2001.
15. Stjerna S, Sairanen V, Gröhn R, Andersson S, Metsäranta M, Lano A, et al. Visual fixation in human newborns correlates with extensive white matter networks and predicts long-term neurocognitive development. *J Neurosci* 2015; 35: 4824–9.

