Concurrent and predictive validity of the Infant Motor Profile

Kirsten Heineman\textsuperscript{1,2}, Karin Middelburg\textsuperscript{1}, Arend Bos\textsuperscript{3}, Lieke Eidhof, Mijna Hadders-Algra\textsuperscript{1}

\textsuperscript{1}Department of Paediatrics, Institute of Developmental Neurology
\textsuperscript{2}Department of Neurology
\textsuperscript{3}Department of Paediatrics, Division of Neonatology
University Medical Center Groningen

Submitted
ABSTRACT

Background: The Infant Motor Profile (IMP) is a qualitative instrument to assess motor behaviour of infants aged 3 to 18 months. The IMP consists of five domains: size of motor repertoire, ability to select motor strategies (variability), fluency, symmetry and motor performance.

Objective: To assess inter-observer reliability, concurrent validity of the IMP with the Alberta Infant Motor Scale and the Touwen Infant Neurological Examination, and predictive validity of the IMP for neurological outcome at 18 months.

Design: A longitudinal prospective study was performed in a group of 30 term born and 59 preterm infants. For the concurrent validity part of the study, a second group of term infants was added with cross-sectional assessments.

Methods: Assessments were performed at corrected ages of 4, 6, 10, 12 and 18 months and consisted of the IMP, AIMS and neurological assessment. Socio-economic and perinatal data were collected. Non-parametric statistics were used to analyze the data.

Results: Inter-observer reliability was high (intra-class coefficient 0.95). Correlations between IMP and AIMS scores varied across IMP domains; they were highest for the performance domain of the IMP (Spearman’s rho 0.47-0.84). A clear relationship was found between total IMP score and neurological condition (Kruskal-Wallis p < 0.001). Sensitivity for prediction of abnormal neurological outcome at 18 months was 86 to 100% with specificity ranging from 71 to 78%.

Limitations: Assessors were not blinded with respect to term or preterm status of the infants. Follow-up did not extend beyond the age of 18 months.

Conclusions: Reliability of the IMP is good and concurrent and predictive validity are satisfactory. These findings support the notion that the IMP is a promising and valuable instrument to assess motor behaviour in infancy.
INTRODUCTION

Prediction of neurological outcome in infants with a high risk for developmental motor disorders, such as cerebral palsy (CP) or developmental coordination disorder (DCD), is difficult. It appears that instruments that assess qualitative aspects of motor behaviour, such as the General Movement method (GM\(^1\)) and the Test of Infant Motor Performance (TIMP\(^3\)) are most promising as single clinical neuromotor predictors\(^4\). However, the GM method and TIMP are only applicable until the age of four months. Therefore, we developed the Infant Motor Profile\(^5\), a qualitative assessment of motor behaviour that is applicable throughout infancy until the age of 18 months.

The IMP was developed for three purposes: first it may be used to detect infants with a high risk for developmental motor disorders, such as CP or DCD. Infants born very preterm are especially at risk for these developmental motor disorders\(^6\). Early detection of high-risk infants is important to provide early intervention at young age when plasticity of the brain is still high\(^7,8\). Second, the IMP may be used for evaluation of changes in neuromotor function, e.g. during or after early intervention. The third aim of the IMP is prediction of future developmental outcome.

The IMP is based on ideas of the Neuronal Group Selection Theory (NGST) on motor development\(^9,10,11\). According to NGST, typical motor development starts with the phase of primary variability with exploratory, variable motor behaviour. Children with pre- or perinatally acquired brain damage show more stereotyped motor behaviour with considerably less variation. During development, infants learn to select adaptive motor strategies out of their primary motor repertoire and to adapt motor behaviour to the environment. This phase of adaptive selection is called secondary variability. Children with developmental motor disorders often have problems in selecting adaptive motor strategies\(^10,11\). Two domains of the IMP are based on these principles of motor development; they assess variation of motor behaviour and the ability to select motor strategies. Three additional domains assess movement fluency, movement symmetry and motor performance.

Two types of validity that are important in the development and validation of a new instrument are concurrent and predictive validity. Concurrent validity is the extent to which scores on the new instrument relate to scores on another measure of the same theoretical construct, ideally a ‘gold standard’. However, no gold standard for assessment of neuromotor function in infancy is available. Therefore, concurrent validity of the new instrument with other established instruments is assessed. Predictive validity is defined as the extent to which current scores on the new instrument predict future developmental outcome. A distinction can be made between prediction of major developmental disorders, such as cerebral palsy, and prediction of minor developmental motor problems, such as minor neurological dysfunction and developmental coordination disorder.

In our pilot study\(^5\), we described the Infant Motor Profile and its theoretical background, its domains and items, and details on scoring procedures. In addition, first data on reliability and some data on validity were presented. Intra and inter observer reliability of scoring were satisfactory in the
persons who developed the IMP. Concurrent validity of the IMP with the Alberta Infant Motor Scale (AIMS\textsuperscript{13,14}) and the Touwen Infant Neurological Examination (TINE\textsuperscript{13,14}) was assessed in a relatively small sample of infants.

Aim of the present study is threefold: first to examine inter-observer reliability of a newly trained assessor without prior experience with the IMP. The second aim is to investigate concurrent validity of the IMP with AIMS and TINE in a large sample of infants and assessments. Based on the idea that the AIMS measures motor performance and the IMP assesses various aspects of motor behaviour including motor performance, we expect a moderate correlation between the total IMP score and the AIMS and a high correlation between the performance domain of the IMP and the AIMS. As the IMP assesses several parameters of neurological integrity, we expect strong association between IMP scores and neurological condition assessed with TINE. Third aim of this study is to determine predictive validity of the IMP at 4, 6, 10 and 12 months for neurological outcome at 18 months measured by the Hempel assessment.

\textbf{METHODS}

\textit{Participants}

We included a longitudinal study group of term and preterm infants and a cross-sectional study group of only term infants. The longitudinal study group consisted of 30 term born and 59 preterm infants. The term infants (12 girls and 18 boys) were recruited from amongst colleagues and acquaintances of the researchers. Median gestational age of the longitudinal term group was 40.1 weeks (range 37.6-42 weeks), median birth weight was 3588 grams (range 2730-4470 grams) and there had been no pre or perinatal complications. Fifty-nine infants were born preterm (25 girls and 34 boys) with median gestational age 29.7 weeks (range 25 to 34.7 weeks) and median birth weight of 1285 grams (range 630 to 2180 grams). The preterm infants had been admitted to the neonatal intensive care unit of the Beatrix Children’s Hospital of the University Medical Center (UMC) in Groningen between December 2003 and January 2005. Thirty-five of the preterm infants were singletons and 24 were twins. Nine pairs of twins participated in the study; the remaining six had lost their twin sibling. Neonatal ultrasound was available for 57 of the 59 preterm infants. Serious brain pathology was observed in six infants: one infant had cystic PVL\textsuperscript{15}, four infants had IVH grade IV\textsuperscript{16} and one infant had middle cerebral artery infarction on the right side. The longitudinal study group (term and preterm infants) was assessed at corrected ages 4, 6, 10, 12 and 18 months.

For the concurrent validity part of the study, another group of 116 term born infants (62 girls and 54 boys) was added. They were recruited at Well Child Centers and had cross-sectional assessments at ages 4, 6, 10, 12 or 18 months. Median gestational age was 40.1 weeks (range 37-43 weeks) and median birth weight was 3500 grams (1960-4660 grams). All parents of the infants signed an informed consent form. The project was approved by the Ethics Committee of the UMC in Groningen.
Socio-economic, perinatal and neonatal data were collected for all infants on standardised forms by means of an interview with the parents and consultation of neonatal intensive care unit discharge certificates. Socio-economic status (SES) was operationalized as the sum score of four variables describing educational and professional level of father and mother, all expressed on a scale from 0 (lowest) through 2 (highest). ‘Small for gestational age’ (SGA) was defined as birth weight below 10th percentile\textsuperscript{17}. ‘Signs of fetal distress’ was defined as the presence of at least one of the following factors: meconium staining, cardiotocography abnormalities and acidaemia during delivery (arterial umbilical pH below 7.05). Table I shows socio-economic and neonatal characteristics of term and preterm groups. As no significant differences in characteristics were found between the longitudinal and cross-sectional term groups, we displayed them as one term group.

Table I: Socio-economic and neonatal characteristics of study groups

<table>
<thead>
<tr>
<th></th>
<th>Term</th>
<th>Preterm</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of children</td>
<td>146</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>72 (49)</td>
<td>34 (58)</td>
<td>0.282</td>
</tr>
<tr>
<td>Maternal age at child birth (years), mean ± SD</td>
<td>32±5.1</td>
<td>31.8 ± 5.2</td>
<td>0.826</td>
</tr>
<tr>
<td>SES\textsuperscript{a}, median (range)</td>
<td>6 (0-8)</td>
<td>4 (0-8)</td>
<td>0.002\textsuperscript{f}</td>
</tr>
<tr>
<td>Twins, n (%)</td>
<td>2 (1.4)</td>
<td>24 (41)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gestational age (weeks), median (range)</td>
<td>40.1 (37-43)</td>
<td>29.7 (25-34.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Birthweight (grams), median (range)</td>
<td>3500 (1960-4660)</td>
<td>1285 (630-2180)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Small for gestational age\textsuperscript{b}, n (%)</td>
<td>15 (10)</td>
<td>20 (34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Caesarian section, n (%)</td>
<td>16 (11)</td>
<td>34 (58)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Signs of fetal distress\textsuperscript{c}, n (%)</td>
<td>35 (24)\textsuperscript{d}</td>
<td>30 (51)\textsuperscript{e}</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Apgar score at 5 minutes, median (range)</td>
<td>10 (7-10)</td>
<td>9 (4-10)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Artificial ventilation, n (%)</td>
<td>1 (0.7)</td>
<td>40 (68)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\textsuperscript{a} SES = socio-economic status, sum score of four variables describing educational and professional level of father and mother, all expressed on a scale from 0 (lowest) through 2 (highest).

\textsuperscript{b} Small for gestational age is defined as birthweight compared with gestational age below 10\textsuperscript{th} centile\textsuperscript{17}.

\textsuperscript{c} Presence of at least one of the following factors: meconium staining, CTG abnormalities, acidaemia during delivery (arterial umbilical pH below 7.05).

\textsuperscript{d} Data on signs of fetal distress were available for 144 of the term infants: 107 (73\%) infants showed no signs of fetal distress, 23 (16\%) had meconium staining, 11 (7.5\%) had CTG abnormalities and one had arterial umbilical pH of 6.82. This infant had perinatal asphyxia with meconium aspiration for which ventilation was required. In addition, neonatal convulsions occurred. Brain MRI at day 6 was normal.

\textsuperscript{e} 30 (51\%) of the preterm infants showed signs of fetal distress: 4 (6.8\%) had meconium staining, 22 (37\%) showed CTG abnormalities, 2 (3\%) had acidaemia with low umbilical pH and 2 (3\%) showed a combination of signs of fetal distress, both as a result of placental dysfunction.

\textsuperscript{f} SES of preterm group is lower than term group

Procedures

Assessments were performed at (corrected) ages 4, 6, 10, 12 and 18 months for the longitudinal
term group and the preterm group. The term cross-sectional group was assessed at one (n = 102 infants), two (n = 13 infants) or three (n = 1) of these ages. The actual number of infants assessed at each age is displayed in Table II. Assessments consisted of a video-recording of approximately 15 minutes of spontaneous motor behaviour in supine, prone, sitting, standing, and walking condition, depending on age and functional capacities of the infant. Furthermore, reaching, grasping and manipulation of objects was tested in supine and in (supported) sitting position. The 80 items of the IMP were scored on the basis of the video recording. These constitute the scores in five domains: size of repertoire (variation), ability to select (variability), movement fluency, movement symmetry and motor performance. The mean of the five domain-scores is the total IMP score\textsuperscript{5}. IMP assessments of the longitudinal term and preterm groups were carried out by KRH who knew whether an infant was term or preterm born, but was not aware of any of the perinatal and neonatal details. IMP assessments of the cross-sectional term group were scored by KRH and LE.

Table II: Number of assessments for term and preterm groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of infants</th>
<th>4 mo</th>
<th>6 mo</th>
<th>10 mo</th>
<th>12 mo</th>
<th>18 mo</th>
<th>Number of assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term longitudinal</td>
<td>30</td>
<td>30\textsuperscript{a}</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>150</td>
</tr>
<tr>
<td>Term cross-sectional</td>
<td>116</td>
<td>22</td>
<td>25</td>
<td>26</td>
<td>29</td>
<td>30</td>
<td>131</td>
</tr>
<tr>
<td>Preterm</td>
<td>59</td>
<td>58</td>
<td>57\textsuperscript{b}</td>
<td>54</td>
<td>54</td>
<td>57</td>
<td>280</td>
</tr>
<tr>
<td>Total</td>
<td>205</td>
<td>110</td>
<td>112</td>
<td>110</td>
<td>113</td>
<td>117</td>
<td>561</td>
</tr>
</tbody>
</table>

\textsuperscript{a} For 1 term infant data of neurological examination at 4 months were missing, only AIMS and IMP were assessed.

\textsuperscript{b} For 1 preterm infant AIMS score could not be determined at 6 months, because assessment in prone position was not performed.

Reliability assessment

LE, who was involved in this study as a master student, was trained in the assessment of the IMP. During a training period of five weeks, 100 video's of term and preterm infants at various ages were assessed. After this training period, inter observer agreement between LE and KRH was investigated on a sample of another 25 video's consisting of five video's at each of the five assessment ages (4, 6, 10, 12 and 18 months). The five video's at each age consisted of two randomly selected video's of the term infant group and three of the preterm group.

Concurrent validity

At all ages the AIMS\textsuperscript{12} and Touwen Infant Neurological Examination (TINE\textsuperscript{13,14}) were assessed, in order to investigate concurrent validity of the IMP with these instruments. The AIMS was scored on the basis of the video-recording of spontaneous motor behaviour. Total AIMS scores, instead of centiles, were used in the data processing, as the Canadian reference values seem currently inappropriate for Dutch children\textsuperscript{18}. Reliability of the AIMS is good, but predictive validity for major developmental...
Concurrent and predictive validity of the Infant Motor Profile

disorders is only moderate\textsuperscript{19,20}. The Touwen Infant Neurological Examination (TINE\textsuperscript{13,14}) was performed at (corrected) ages of 4, 6, 10 and 12 months. In TINE, neurological signs are organized according to age-specific norms into clusters of dysfunction. Five clusters are distinguished: reaching and grasping, gross motor function, brain stem function, visuomotor function and sensorimotor function (consisting of reflexes and muscle tone). Neurological condition is classified as abnormal if there is a distinct neurological syndrome, such as a hemisyndrome, irrespective of the number of deviant clusters. An infant is classified as having minor neurological dysfunction (MND) in case of presence of more than two clusters of dysfunction. Two forms of typical neurological condition are distinguished: normal – suboptimal when one or two clusters are deviant and normal when no clusters fulfil criteria for dysfunction\textsuperscript{14}. Reliability of TINE is good. Predictive validity is good for major developmental motor disorders such as cerebral palsy and moderate for minor motor disorders\textsuperscript{4,14}. Predictive validity for major developmental motor disorders such as cerebral palsy is good, but for minor motor disorders moderate at best\textsuperscript{4,14}.

**Predictive validity**

For assessment of the predictive validity of the IMP, neurological outcome at the (corrected) age of 18 months was determined with the Hempel assessment\textsuperscript{21}. This method is suitable for children of pre-school age, from 18 months until four years of age. Similar to the TINE, the Hempel assessment classifies neurological signs into clusters of dysfunction, namely fine motor dysfunction, gross motor dysfunction, dysfunctional muscle tone regulation, reflex abnormalities and visuomotor dysfunction. Neurological condition is classified into four categories: abnormal, complex MND (denoting the presence of more than one dysfunctional cluster), simple MND (one cluster of dysfunction) or normal (no deviant clusters or the isolated presence of reflex abnormalities)\textsuperscript{22}.

**Statistical analyses**

To analyze inter-observer reliability, intra-class correlation coefficients (ICCs) for a two-way mixed effects model with associated 95% confidence intervals were used. Differences in IMP scores between term and preterm groups and between the four neurological conditions were analyzed by means of the non-parametric Mann-Whitney U test and Kruskal-Wallis test respectively. Relations between IMP scores and AIMS scores and correlation between IMP scores throughout infancy and neurological outcome at 18 months were assessed with Spearman’s rank correlation with associated confidence intervals. Interpretation of Spearman’s correlation coefficient was as follows: $\rho < 0.50$ weak relationship, $0.50 \geq \rho \geq 0.75$ moderate relationship, $\rho > 0.75$ good relationship\textsuperscript{23}. To assess predictive validity of the IMP scores for the outcome at 18 months, a cut-off score below the 5\textsuperscript{th} percentile was used. Throughout the analyses, differences and correlations with a $p$-value $< 0.05$ were considered to be statistically significant (two-tailed testing).
RESULTS

At all ages total IMP scores did not significantly differ between girls and boys. The preterm group consistently showed lower total IMP scores than the term group (Figure 1, Mann-Whitney U test p values < 0.001). This was also the case for the scores on the domains size of repertoire (p < 0.001 at all ages), fluency (p< 0.001 at 4, 6, 12 and 18 months, p = 0.028 at 10 months) and performance (p = 0.001 at 4 months and p < 0.001 at 6, 10, 12 and 18 months). Scores on adaptive selection (variability) were significantly lower for preterm infants compared to term infants from age 10 months onwards (4 months p = 0.90, 6 months p = 0.16, 10 months p<0.001, 12 months p = 0.03, 18 months p = 0.001). Symmetry scores were significantly lower for the preterm group at ages 4 and 18 months (p-values 0.02 and 0.04 respectively), but not at 6, 10 and 12 months.

Reliability

Interobserver reliability of the total IMP score yielded an intraclass correlation coefficient (ICC) of 0.95 (95% confidence interval 0.89-0.98), indicating good reliability. Reliability of IMP domains was moderate to good with ICC from 0.74 to 0.99 (Table III).

Concurrent validity

Correlations between AIMS scores and total IMP scores were weak to moderate at all ages. 

Figure 1: Differences in total IMP scores between term and preterm infants at 4, 6, 10, 12 and 18 months. Data are presented as median values (horizontal bars), interquartile ranges (boxes) and ranges (vertical lines). Open circles and asterisks represent outliers. FT = full term group, PT = preterm group. At all ages differences in total IMP scores between the two groups were significant (Mann-Whitney U test p < 0.001).
Concurrent and predictive validity of the Infant Motor Profile

(Spearman's rho 0.36 to 0.55, see Table IV). The performance domain of the IMP showed the strongest correlations with the AIMS scores, especially at the age of 10 and 12 months (Spearman's rho 0.84 and 0.81 respectively, Table IV). Correlations between the other domains and the AIMS were weak (Spearman's rho 0.01-0.41, Table IV). Preterm infants had significantly lower AIMS scores than term infants at ages 4, 10, 12 and 18 months (Mann-Whitney U test, p-values 0.001, < 0.001, < 0.001 and <0.001 respectively). No difference was found at age 6 months (p = 0.32).

We found a clear relationship between the total IMP score and neurological condition at all ages: infants with a normal neurological condition had highest IMP scores and infants with an abnormal neurological condition had lowest IMP scores (Figure 2, Kruskal-Wallis p < 0.001 at all ages). The domains size of repertoire, fluency, symmetry and performance were highly significantly related to neurological condition at all ages (p-values for variability, fluency and performance all < 0.001, except for fluency at 10 months p = 0.008; p-values for symmetry respectively 0.025, 0.006, < 0.001, 0.025 and < 0.001 at 4, 6, 10, 12 and 18 months). Scores on the domain ability to select (variability) were significantly different between neurological conditions at ages 10 and 12 months (p=0.021 and 0.008 respectively), but not at ages 4, 6 and 18 months (p= 0.49, 0.46 and 0.06 respectively).

Predictive validity

Neurological condition at 18 months was determined with the Hempel examination. Of the longitudinal term group, 23 children had normal neurological condition and 7 had simple MND. None of the term children showed complex MND or abnormal neurological condition at 18 months. Of the preterm group, eleven children had a normal neurological condition, 7 had simple MND, 31 had complex MND and neurological condition of 8 infants was considered as abnormal (14% of preterm group), of which four infants had a unilateral spastic CP and four had bilateral spastic CP. Two preterm children did not have follow-up at 18 months.

For the total longitudinal group of infants, correlation between total IMP scores throughout infancy and neurological outcome at 18 months was moderate, with Spearman's rho's of -0.62
Figure 2: Relationship between total IMP scores and neurological condition at the various ages. Data are presented as median values (horizontal bars), interquartile ranges (boxes) and ranges (vertical lines). N = normal neurological condition, N-sub = normal suboptimal neurological condition, MND = minor neurological dysfunction, A = abnormal neurological condition, mo = months, n = number of infants, S-MND = simple MND, C-MND = complex MND. At all ages differences in total IMP scores between the four neurological conditions were significant (Kruskal-Wallis p < 0.001).
Table IV: Spearman’s correlation coefficients of IMP scores and AIMS scores per age

<table>
<thead>
<tr>
<th></th>
<th>AIMS 4 mo</th>
<th>AIMS 6 mo</th>
<th>AIMS 10 mo</th>
<th>AIMS 12 mo</th>
<th>AIMS 18 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total IMP score</strong></td>
<td>0.43**</td>
<td>0.34**</td>
<td>0.55**</td>
<td>0.43**</td>
<td>0.36**</td>
</tr>
<tr>
<td></td>
<td>(0.26-0.57)</td>
<td>(0.19-0.52)</td>
<td>(0.40-0.66)</td>
<td>(0.27-0.57)</td>
<td>(0.19-0.51)</td>
</tr>
<tr>
<td><strong>IMP domain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of repertoire</td>
<td>0.39**</td>
<td>0.29**</td>
<td>0.36**</td>
<td>0.37**</td>
<td>0.41**</td>
</tr>
<tr>
<td></td>
<td>(0.22-0.54)</td>
<td>(0.11-0.45)</td>
<td>(0.19-0.51)</td>
<td>(0.19-0.52)</td>
<td>(0.25-0.55)</td>
</tr>
<tr>
<td>Ability to select</td>
<td>0.07</td>
<td>0.01</td>
<td>0.22*</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(-0.11-0.26)</td>
<td>(-0.18-0.19)</td>
<td>(0.04-0.39)</td>
<td>(-0.07-0.29)</td>
<td>(-0.14-0.22)</td>
</tr>
<tr>
<td>Fluency</td>
<td>0.27**</td>
<td>0.24*</td>
<td>0.08</td>
<td>0.15</td>
<td>0.23*</td>
</tr>
<tr>
<td></td>
<td>(0.09-0.43)</td>
<td>(0.06-0.41)</td>
<td>(-0.11-0.26)</td>
<td>(-0.03-0.33)</td>
<td>(0.05-0.39)</td>
</tr>
<tr>
<td>Symmetry</td>
<td>0.23*</td>
<td>0.05</td>
<td>0.13</td>
<td>0.12</td>
<td>0.32**</td>
</tr>
<tr>
<td></td>
<td>(0.05-0.40)</td>
<td>(-0.13-0.24)</td>
<td>(-0.05-0.31)</td>
<td>(-0.07-0.30)</td>
<td>(0.14-0.47)</td>
</tr>
<tr>
<td>Performance</td>
<td>0.56**</td>
<td>0.56**</td>
<td>0.84**</td>
<td>0.81**</td>
<td>0.47**</td>
</tr>
<tr>
<td></td>
<td>(0.42-0.67)</td>
<td>(0.41-0.67)</td>
<td>(0.78-0.89)</td>
<td>(0.73-0.86)</td>
<td>(0.31-0.60)</td>
</tr>
</tbody>
</table>

Spearman’s correlation coefficients with associated 95% confidence intervals between brackets, * p < 0.05, ** p < 0.01.

Table V: Predictive validity of total IMP score at ages 4, 6, 10 or 12 months for neurological outcome at 18 months

<table>
<thead>
<tr>
<th></th>
<th>Total IMP 4 mo</th>
<th>Total IMP 6 mo</th>
<th>Total IMP 10 mo</th>
<th>Total IMP 12 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction of CP at 18 mo</strong> (IMP score &lt; p5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>88</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>Specificity</td>
<td>71</td>
<td>78</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>PPV</td>
<td>26</td>
<td>29</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>NPV</td>
<td>100</td>
<td>98</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>Accuracy</td>
<td>73</td>
<td>79</td>
<td>79</td>
<td>78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Total IMP 4 mo</th>
<th>Total IMP 6 mo</th>
<th>Total IMP 10 mo</th>
<th>Total IMP 12 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction of deviant neurological outcome (complex MND or CP) at 18 mo</strong> (IMP score &lt; p5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>63</td>
<td>55</td>
<td>56</td>
<td>63</td>
</tr>
<tr>
<td>Specificity</td>
<td>85</td>
<td>94</td>
<td>93</td>
<td>98</td>
</tr>
<tr>
<td>PPV</td>
<td>77</td>
<td>88</td>
<td>87</td>
<td>96</td>
</tr>
<tr>
<td>NPV</td>
<td>75</td>
<td>72</td>
<td>73</td>
<td>78</td>
</tr>
<tr>
<td>Accuracy</td>
<td>76</td>
<td>76</td>
<td>77</td>
<td>83</td>
</tr>
</tbody>
</table>

mo = months, p5 = 5th percentile
Sensitivity = true positives/(true positives + false negatives)
Specificity = true negatives/(true negatives + false positives)
PPV = positive predictive value = true positives/all positives
NPV = negative predictive value = true negatives/all negatives
Accuracy = (true positives + true negatives)/all subjects
(CI -0.74—0.47), -0.57 (CI -0.70—0.41), -0.67 (CI -0.77—0.53) and -0.65 (CI -0.76—0.50) respectively at ages 4, 6, 10 and 12 months. Predictive validity of total IMP scores (with cut-off score below 5th percentile) for abnormal neurological condition (CP) at 18 months varied with age of assessment and showed sensitivity from 86 to 100% and specificity from 71 to 78% (Table V). Predictive validity for deviant neurological outcome (complex MND or CP) at 18 months demonstrated sensitivity from 55 to 63% and specificity of 85 to 98% (Table V). If instead of 5th percentile the 15th percentile was used as cut-off point, sensitivity increased to 68 to 89%, but specificity decreased to 72 to 87%. Accuracy was around 80% and did not differ between level of cut-off point at 5th or 15th percentile.

**DISCUSSION**

Our study showed a good reliability of the IMP. Concurrent validity of the IMP with the AIMS met the a priori theoretical expectations: the IMP performance domain correlated best with the AIMS. Concurrent validity of the IMP with TINE and Hempel examination was very good. Predictive validity of the IMP for neurological outcome at 18 months was satisfactory.

Strengths of this study are its predominantly longitudinal character and the low attrition rate (3.5% over all longitudinal assessments, see Table II). A weakness of the study is that the assessors were not blind with respect to term or preterm status of the infant. This may have influenced scoring. However, the assessors were unaware of any details of the child’s clinical history or results of neonatal ultrasounds. Besides, if assessors had been blinded to term or preterm status, it would have been difficult to conceal preterm status, because the infant’s appearance usually discloses preterm birth. A second weakness of this study is the relatively short duration of follow-up. At 18 months, clinical signs of CP often are not yet fully expressed and signs of minor developmental motor disorders may not be present until school age.

The term group showed a relatively high percentage of infants with signs of fetal distress, consisting mainly of meconium staining which is fairly common in term deliveries. The group of preterm infants included in this study may be considered as a representative sample of a Dutch neonatal intensive care unit population in a tertiary referral centre with respect to gestational age, birthweight, frequency of Apgar score at 5 minutes below 7 (15%) and ventilatory support (68%). Our preterm sample showed a relatively high percentage of infants small for gestational age and Caesarean deliveries. The latter could be due to the increased tendency over the years to deliver very preterm infants by Caesarean section. The preterm group had lower socio-economic status than the term group, in accordance with social disadvantage being considered a risk factor for preterm birth.

Interobserver reliability of total IMP score and the domains size of repertoire, symmetry and performance was good to very good. Reliability of the IMP domains ability to select and fluency was moderate. Interobserver reliability in this study was higher than in our pilot study, probably as a result of the more precise definitions and descriptions of the IMP items which had been developed.
Concurrent and predictive validity of the Infant Motor Profile in the meantime. With good training, IMP scoring can be learnt reliably without prior experience in infant motor development.

The study showed clear differences in IMP scores between term and preterm infants for total IMP scores and the domains size of repertoire (variability), fluency and performance at all ages. Preterm birth is a major risk factor for developmental motor disorders\(^6\). Preterm infants are at risk for brain lesions, which according to NGST lead to reduced variability of motor behaviour\(^11\). Loss of fluency of motor behaviour is one of the first signs of non-optimal neurological condition\(^1\). Delayed acquisition of motor milestones, in the IMP reflected as lower performance scores, can be a sign of developing CP\(^31\). For the domain ability to select (variability), scores significantly differed between term and preterm infants from 10 months onwards. In typically developing infants, the ability to select suitable motor strategies gradually emerges at function-specific ages after the first half year of life\(^{10}\). Infants with developmental motor disorders often have problems in processing afferent, sensory information and fine-tuning and adapting motor behaviour\(^{11,32,33}\). This could delay or hamper development of the ability to select, as adaptive selection relies on afferent feedback. Scores on the domain symmetry differed between term and preterm infants at ages 4 and 18 months. The asymmetries observed at 4 months could be transitory neurological findings that resolve spontaneously\(^{34}\), whereas the asymmetries observed at later age of 18 months could be signs of the development of unilateral spastic CP\(^{24}\).

As we expected, the total IMP score only correlated to a moderate extent with the AIMS. Correlation between AIMS and IMP was highest for the performance domain of the IMP, especially at the ages of 10 and 12 months. Both the performance domain of the IMP and the AIMS assess motor achievements in a quantitative way. From the age of 10 months onwards, motor development is characterized by a rapid gain in motor milestones, which is reflected in both IMP and AIMS scores. After the age of 14 months, the discriminative power of the AIMS is diminished\(^{12,35}\). The IMP domain size of repertoire assesses another aspect of motor behaviour than the AIMS. The finding that these parameters were weakly but significantly correlated at all ages indicates that both assess different aspects of the same underlying construct, being neuromotor integrity. The IMP domain ability to select was only related to the AIMS score at age ten months, the first age in the present study at which selection of adaptive motor strategies was present to some extent. The IMP domain fluency was weakly related to the AIMS. Loss of movement fluency is one of the first signs of non-optimal neurological condition\(^1\), but it is not specific for serious developmental motor disorders that are associated with low motor performance. The IMP domain symmetry was related to the AIMS score at the ages of 4 and 18 months, the latter probably representing the infants with a developing unilateral CP, which besides a low symmetry score also leads to reduced motor performance\(^{31}\).

Concurrent validity of IMP and the age-specific neurological examination (TINE and Hempel) was very good: at all ages infants with normal neurological condition had higher IMP scores than infants with minor neurological dysfunction or abnormal neurological condition. Differences in IMP scores were especially found between infants with (complex) MND or abnormal neurological
condition and infants with normal or normal – suboptimal neurological condition (or simple MND at 18 months). Complex MND, in contrast with simple MND, has clinical relevance and is associated with pre- or perinatal adversities. These differences in IMP scores between infants with different neurological conditions were found at all ages for the domains: size of repertoire (variation), fluency, symmetry and performance, supporting the notion that these are indeed parameters of neuromotor integrity. Scores on the domain ability to select were related to neurological condition at ages 10 and 12 months, but not at 4, 6 and 18 months. This is in analogy with the differences we found between term and preterm infants, except that they did differ in adaptive selection scores at 18 months. The data indicated that the absence of a relation between neurological condition and adaptive selection at 18 months was brought about by relatively good scores of infants with an abnormal neurological condition. Motor behaviour of these children was characterized by a limited motor performance; nevertheless they demonstrated, within the skills which they had developed, a relatively good ability of adaptive section out of their limited motor repertoire.

Sensitivity of IMP scores throughout infancy for predicting CP at 18 months was very high. Positive predictive value of the IMP for neurological outcome at 18 months was low, but negative predictive value was high, implying that IMP scores above the 5th percentile almost certainly excluded abnormal neurological outcome at 18 months. It is important to realize that predictive values of a test strongly depend on the prevalence of the disorder, e.g. CP, in the study population. Therefore, predictive values observed in our sample cannot be extrapolated to the general population.

Prediction of developmental outcome at an early age is difficult and will never be perfect, because change is one of the main characteristics of the developing brain. To optimize prediction of neuromotor outcome in children at high risk for developmental motor disorders, it is probably best to combine multiple, complementary tools, such as neurological examination, assessment of milestones and assessment of qualitative aspects of motor behaviour in addition to neuroimaging and neurophysiological techniques.

**CONCLUSION**

The Infant Motor Profile is a qualitative assessment of motor behaviour based on the Neuronal Group Selection Theory on motor development. Interobserver reliability of the IMP is good. Concurrent validity of the IMP with the Alberta Infant Motor Scale was especially high for the performance domain of the IMP. Concurrent validity of IMP with Touwen Infant Neurological Examination was very good. With respect to the three purposes for which the IMP was developed, we can conclude that the IMP is well able to discriminate between typically developing infants and infants with high risk for developmental motor disorders. The ability of the IMP to evaluate motor function over time should be further explored by applying the IMP in intervention studies. Prediction of neurological outcome is very difficult, due to change being one of the main characteristics of the developing nervous system. In our study population, predictive validity of the IMP was satisfactory. Future studies will aim at generating norm-scores and determining clinical applicability.
ACKNOWLEDGEMENTS

We thank Nicole Arink, Hylco Bouwstra, Dineke Dijkhuizen and Anne Hoekstra for performing part of the assessments. The study was supported by a Junior Scientific Masterclass grant of the post-graduate school Behavioral and Cognitive Neurosciences (BCN), University of Groningen.
REFERENCES


Concurrent and predictive validity of the Infant Motor Profile


29. Haque KN, Hayes AM, Ahmed Z, Wilde R, Fong CY. Caesarean or vaginal delivery for preterm very-low-birth weight (< or =1,250 g) infant: experience from a district general hospital in UK. Arch Gynecol Obstet. 2008;277:207-12.


