The prevalence and predictive value of weak language skills in children with very low birth weight – a longitudinal study

Suvi Stolt (suvi.stolt@utu.fi), Jaakko Matomäki, Annika Lind, Helena Lapinleimu, Leena Haataja, Liisa Lehtonen

1. Department of Behavioural Sciences and Philosophy, University of Turku, Turku, Finland
2. Department of Pediatrics, Turku University Hospital and University of Turku, Turku, Finland
3. Department of Pediatric Neurology, Turku University Hospital and University of Turku, Turku, Finland

ABSTRACT

Aim: Previous findings regarding the prevalence and predictive value of weak language skills in preterm children with very low birth weight (VLBW) are unclear. This study analysed the prevalence of weak language skills, the predictive value of early weak language skills on later weak language skills, and the sensitivity and specificity of cognitive scores for identifying concurrent weak language skills in a longitudinal sample of VLBW children (n = 141) and their full-term controls (n = 146).

Methods: Data on language skills and cognitive development were gathered at two and five years of age. Weak language skills were defined by the 10th percentile value of the controls.

Results: In VLBW children, the prevalence of weak language skills varied between 16% and 18% at 2 years of age (controls: 8 to 10%) and between 20% and 27% at 5 years of age (controls: 10%). Early weak language skills predicted later weak language skills in VLBW children. Cognitive scores were specific, but their sensitivity for identifying concurrent weak language skills was low.

Conclusion: The prevalence of weak language skills in VLBW children increased during the follow-up period and was higher than the controls. Language-sensitive methods should be used in the clinical follow-up of VLBW children.

INTRODUCTION

The language skills of children born very preterm (<32 gestational weeks) and/or with low birth weight (VLBW) (birth weight ≤1500 g) have been studied mostly at preschool age (1–3). However, there are also studies that provide information on earlier development (4,5). The common finding is that these children as a group exhibit weaker language skills than the full-term controls. The prevalence of weak language skills in very preterm children varies in different studies (3,6,7). Since only a handful of studies provide longitudinal information on language development in this group (7–9), and the findings in these studies provide contradictory information on whether the prevalence of weak language skills in preterm children remains stable over time, or whether it increases or decreases, there is a need for further research.

In clinical work, scores on cognitive development are often used to indicate whether a preterm child needs follow-up in terms of his/her language development. However, cognitive scores only provide a general index for cognitive development, and it is possible that they mask problems in distinct skills (10–13). It is unclear how specific and how sensitive information the general cognitive scores provide on weak language skills in very low birth weight children.

The aims of the present study were to analyse (i) the prevalence of weak language skills in a longitudinal sample of VLBW children and their controls at two and at five years of age (ii) the predictive value of early weak language skills on weak language skills at 5 years of age; and (iii) the specificity and sensitivity of the cognitive scores for identifying concurrent weak language skills in VLBW children.

Key notes

- Previous findings regarding the prevalence and predictive value of weak language skills in preterm children with very low birth weight (VLBW) are unclear.
- Our study found that the prevalence of weak language skills was higher in the VLBW group than controls at 2-years-of-age and that the group difference increased during the 3-year follow-up period.
- Clinicians should use language-sensitive measures to screen language skills in children born with a VLBW.
scores for identifying VLBW children with weak language skills at two and at five years of age.

METHODS

Participants

The participants were VLBW children who were born at Turku University hospital between January 2001 and December 2006. The flow chart of the VLBW children who took part in the present study is presented in Figure 1. Families of 210 VLBW infants were invited to take part in the study in the neonatal intensive care unit. At 2 years of age (corrected age for the VLBW children), the parent-report method was used to assess the children’s language development. The present sample consisted of those VLBW children whose parents had filled in the inventory, and who were living in Finnish-speaking families. At 2 years of age, our study population consisted of 141 VLBW children. At 5 years of age, the number of VLBW children assessed varied between 120 and 131.

The first, healthy, full-term (at ≥ 37 gestational age) girl and boy born each week at Turku University Hospital between November 2001 and March 2004 were invited to join the control group. The exclusion criteria were as follows: admission to neonatal intensive care unit, small for gestational age (≤−2 SD), presence of syndromes or congenital anomalies and mother’s known use of alcohol or drugs during the pregnancy. A group of 199 children was recruited. At 2 years of age, the parent-report method was used to evaluate children’s language development. The inventory was given to 193 families (97%) and 164 (82%) were returned. The present sample consisted of children whose families had filled in the inventory and who were living in Finnish-speaking families (n = 146). At 5 years of age, the number of children assessed varied between 123 and 137. The study protocol was approved by the Ethical Committee of the Hospital District of Southwest Finland in November 2001. All the parents gave written consent after having been informed about the study.

FinCDI = the Finnish version of the MacArthur-Bates Communicative Development Inventory

Figure 1 The flow chart of the participants. VLBW = very low birth weight children; FinCDI = the Finnish version of the MacArthur-Bates Communicative Inventory.
The background characteristics are shown in Table 1. Bronchopulmonary dysplasia was diagnosed if a child needed supplemental oxygen at 36 gestational weeks. The groups did not differ in terms of mothers’ educational background, \( \chi^2 (2) = 2.48, p = 0.29 \).

**Assessment of language abilities and cognitive development**

**Expressive language skills and cognitive development at 2 years of age**

At 2 years of age, the validated Finnish, toddler version of the MacArthur–Bates Communicative Development Inventory (14,15) and the expressive language score (10) of the Bayley scales of Infant Development version II (Bayley-II) (16) were used to assess the language development of the children. The MacArthur–Bates Communicative Development Inventory measures the development of lexicon and early grammar. The following variables were used: lexicon size and the mean length of the three longest utterances value. Lexicon size is the number of words the mother has estimated that her child uses. The estimation is based on comprehensive word lists. The mean length of the three longest utterances value is calculated in morphemes based on the three longest utterances the child has recently said. The expressive language score was calculated based on the following 15 items of the Bayley-II: the number of pictures (maximum 10) and the number objects (maximum five) named in the testing situation. Although expressive language score has not been validated, it has been shown to give representative information on language delay in high-risk children (10). Cognitive development was assessed using the Bayley-II, and mental developmental index was determined.

**Language skills and cognitive development at 5 years of age**

At 5 years of age, information on language skills was gathered using the standardised edition of the Nepsy II test (17) and the validated Five-to-Fifteen questionnaire (18,19). The Nepsy II is a neuropsychological test. The following subtests of the Nepsy II were used: phonological processing, speeded naming, comprehension of instructions, word list interference and narrative memory. The phonological processing subtest focuses on phonemic awareness and the speeded naming subtest on rapid production of familiar lexical items. The comprehension of instructions subtest assesses the ability to understand instructions of increasing syntactic complexity. The word list interference subtest measures verbal working memory and the narrative memory subtest the memory for logical verbal information. The first three subtests are part of the ‘Language domain’, and the last two subtests are part of the ‘Memory and learning domain’ in the Nepsy II. The two verbal subtests from the ‘Memory and learning domain’ were used as memory skills are necessary when processing language. In these subtests, linguistic material is used in the testing situation. Our procedure follows with slight modifications Korkman et al. (20), as both ‘language’ and ‘memory’ subtests of Nepsy were used to study language skills in Korkman et al.’s study. For the purposes of the present study, the mean value of the five subtests was calculated for each child (mean Nepsy-Language score).

The Five-to-Fifteen questionnaire is a parent-report method consisting of 181 questions on the development of children aged between five and 15 years (18,19). The parents are asked to compare the development of their child with the development of same-aged children based on the statements presented. Parents choose the alternative (0 = does not describe, 1 = describes to some extent, 2 = describes well) that best describes their child’s behaviour. A higher score indicates more problems. In the present study, the mean value of the language domain of the Five-to-Fifteen questionnaire was used (mean Five-to-Fifteen-Language score). This value was calculated based on the following items: five items measuring receptive language, 13 items measuring expressive language and three items measuring communication ability.

The cognitive development was assessed using a short form of the Wechsler Preschool and Primary Scale of Intelligence-Revised test (21). The full-scale intelligence quotient was estimated based on the following subtests: information, sentences, arithmetic, block design, geometric design and picture completion. The verbal intelligence quotient was estimated based on the following subtests: information, sentences and arithmetic.
Statistical analysis
The 10th percentile cut-offs of the results of the control children on different measures were used to define weak language skills at both age points. Comparisons between two categorical variables (maternal educational background, the effect of gender and small for gestational age status on weak language skills) were made using chi-square test. The percentages of the children with weak language skills in the VLBW group and in the control group were compared using Fisher’s exact test. Pearson’s correlation coefficient values were used to analyse the correlations between the results of the language measures at two and five years of age. The predictive value of weak language skills at 2 years of age on weak language at 5 years of age was analysed using the logistic regression analysis. Since the language-related independent variables were correlating with each other, they were analysed separately. Two dependent variables were used in six analyses for both groups; in the first three analyses, the dependent variable was weak language skills estimated based on the mean Nepsy-Language score, and in the next three analyses, the dependent variable was weak language skills estimated based on the mean Five-to-Fifteen-Language score. In VLBW children, the common independent variable in all models was neurological impairment (bilateral hearing loss >40 dB, cerebral palsy, low cognitive development; mental developmental index value <70 at 2 years of age). Weak language based on lexicon size at 2 years of age was used as another independent variable in the first model. Weak language based on the mean value of the three longest utterances was used in the second model. Weak language based on expressive language score at 2 years of age was used in the third model. For the controls, the same analyses with the same dependent and independent variables were run, except that the independent variable of neurological impairment was not used. The standard score 85 was used as a cut-off value when calculating the specificity and sensitivity of the cognitive scores (mental and verbal intelligence quotient at 5 years of age) for identifying concurrent weak language skills.

RESULTS
The prevalence of weak language skills
At 2 years of age, the prevalence of weak language skills varied between 16% and 18% in the VLBW group and between 8% and 10% in the controls (Table 2). In two of the three prevalence figures, there was a significant difference (p < 0.5) between the groups. When children with neurological impairment were excluded from the VLBW group, the prevalence of weak language skills decreased slightly and the differences between the groups were no longer significant.

At 5 years of age, the prevalence of weak language skills varied between 20% and 27% in the VLBW group, and was 10% in the controls. When children with neurological impairment were excluded from the VLBW group, there was still a significant difference between the groups if the prevalence of weak language skills was measured using mean Nepsy-Language score.

The gender did not have an effect on the prevalence of weak language skills in either of the groups at two or at 5 years of age. Nor did the small for gestational age status have an effect on the prevalence of weak language skills in the VLBW group in either of the studied age points.

Correlations between language abilities at two and at five years of age, and the predictive value of early language
Language skills measured at 2 years of age correlated significantly with mean Nepsy-Language score at 5 years of age in both groups (p < 0.001; Table 3). There were also significant correlations between low early language scores and high mean Five-to-Fifteen-Language scores (VLBW group: p < 0.001; controls: between p < 0.01 and p < 0.001).

In VLBW children, a low mean Nepsy-Language score at 5 years of age was predicted significantly by a small lexicon size, by neurological impairment and by a very short utterance length at 2 years of age (Table 4). Furthermore, a small lexicon size, a short utterance length and a low expressive language score at 2 years of age predicted a high mean Five-to-Fifteen-Language score in the VLBW group.
In the control group, a low mean Nepsy-Language score at 5 years of age was predicted only by a short utterance length at 2 years of age.

Specificity and sensitivity of cognitive scores
The mental developmental index of Bayley-II identified most of the children without weak language skills in the VLBW group and almost all of them in the control group at 2 years of age (Table 5). The same was true for the full-scale intelligence quotient and verbal intelligence quotient of the Wechsler’s test assessed at 5 years of age. However, none of the three measures identified the majority of the children with weak language skills, that is, the sensitivity of the values was weak.

DISCUSSION
This study provides longitudinal information on the prevalence of weak language skills in the VLBW children. The percentages of the children with language skills in the VLBW group are slightly lower than those presented earlier (6,7). One reason for the different prevalence figures of weak language skills found in different studies may be the different methods used; the age point when language development is measured may also have an effect on the results. It is also possible that the moderately good cognitive development detected in the present sample of VLBW children (22) may have affected positively the results. In addition, it is possible that the neonatal treatments as well as the neonatal care culture may affect language development in preterm infants. As parent’s presence has been shown to increase the infant’s vocalisation already at 32 weeks of gestation (23), it is possible that the neonatal care culture (i.e. family centred care in the neonatal intensive care unit at the Turku University Hospital) may have contributed to the language development of the present VLBW group.

When preterm children with neurological impairment were excluded, there was still a significant difference between the VLBW group and the controls in the prevalence of weak language skills when measured using...
Language skills and very low birth weight

Stolt et al.

Table 5 The specificity and the sensitivity of the mental developmental index (MDI) of Bayley-II, and full-scale intelligence quotient (FSIQ) and verbal intelligence quotient (VIQ) of the Wechsler’s test for identifying concurrent weak language skills. The information is presented separately for the very low birth weight children (VLBW) and the full-term controls for both age points

<table>
<thead>
<tr>
<th>At 2:0:</th>
<th>Specificity of the MDI</th>
<th>Specificity of the FSIQ</th>
<th>Specificity of the VIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLBW group</td>
<td>Lexicon size</td>
<td>95.7</td>
<td>98.5</td>
</tr>
<tr>
<td></td>
<td>Mean length of the three longest utterances value</td>
<td>93.6</td>
<td>98.5</td>
</tr>
<tr>
<td>Control group</td>
<td>Lexicon size</td>
<td>38.5</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>Mean length of the three longest utterances value</td>
<td>33.3</td>
<td>9.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>At 5:0</th>
<th>Specificity of the FSIQ</th>
<th>Specificity of the VIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLBW group</td>
<td>Mean Nepsy-Language score</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>Mean Five-to-Fifteen-Language score</td>
<td>90.0</td>
</tr>
<tr>
<td>Control group</td>
<td>Mean Nepsy-Language score</td>
<td>43.8</td>
</tr>
<tr>
<td></td>
<td>Mean Five-to-Fifteen-Language score</td>
<td>35.0</td>
</tr>
</tbody>
</table>

The finding that none of the cognitive measures used in the present study were good at identifying children with weak language skills is in line with earlier results (2,10). Accordingly, even if cognitive data provide some information on the risk for language impairment (7), a decision as to whether or not a child needs follow-up regarding his/her language development cannot be based solely on the general cognitive score, but language-sensitive measures should be used. Neither the Bayley-II nor cognitive tests for older children have been developed for the purpose of detecting language deficits. Thus, language-sensitive tools are needed to identify those VLBW children with weak language skills.

The Nepsy II test at 5 years of age. This result is parallel with earlier findings (2,7). For example, in an earlier study (7), there was a significant difference in the percentage of children with a risk for language impairment between the groups of healthy preterm children and full-term controls at three-and-a-half year of age. These findings propose that the high incidence of neurological impairment in the preterm children does not alone explain the increased prevalence of weak language skills in this population.

The prevalence of weak language skills among VLBW children increased during the 3-year follow-up period. Only few earlier studies provide longitudinal information on the prevalence of weak language in preterm population. In a recent meta-analysis, the group difference between preterm and full-term children in complex language functions increased during the period from three to twelve years of age (8). Although this meta-analysis focused on cross-sectional reports only, its result is parallel with the present finding. The same tendency was found in a recent follow-up study of preterm children (born at <33 weeks, birth weight ≤1600 g) (7). In this study, within a 1-year period from two-and-a-half years of age to three-and-a-half years of age, the percentage of the preterm children with a risk for language problems increased. These results together contradict the earlier findings that very preterm children and/or VLBW children catch up their peers in terms of their language development (9,24). This discrepancy may be explained by inclusion of only receptive language in the earlier studies (9,24), or by different age points at data collection (9).

In the present cohort, weak language ability at 2 years of age was a significant predictor of weak language skills 3 years later especially in the VLBW children. Thus, our result showed that a considerable amount of the VLBW infants with weak language skills at 5 years of age can be identified already at the age of 2 years. To our knowledge, there is only one earlier study reporting the predictive value of early language on later language in preterm children (7) showing that the small lexicon size and a very short utterance length at two-and-a-half years of age both predicted a very short utterance length at three-and-a-half years of age in healthy preterm children. It is interesting to think about the present findings in the light of the literature regarding late-talking toddlers. Late-talking toddlers are children with slow expressive language development during the first 2 years, but with normal cognitive development, and with no other diagnosis or medical problem which would explain the early slow expressive language development. In the previous follow-up study (25), the expressive lexicon size measured at the beginning of the third year correlated with different language measures at the age of six, eight and 9 years. In addition, late-talking toddlers also had poorer scores on the majority of language measures through age of 9 years than the controls. Thus, as also argued in earlier studies (25, see also (7)), the early expressive vocabulary acquisition may be a good measure of general language development at the age of 2 years. It is possible, that very slow early expressive lexical development reflects the difficulties such as weak processing skills, weak articulatory skills or weak phonological representation skills (25). Furthermore, our result also lifts up early, very short utterance length as a predictor for further language development in VLBW children as also shown by others (7).

The finding that none of the cognitive measures used in the present study were good at identifying children with weak language skills is in line with earlier results (2,10). Accordingly, even if cognitive data provide some information on the risk for language impairment (7), a decision as to whether or not a child needs follow-up regarding his/her language development cannot be based solely on the general cognitive score, but language-sensitive measures should be used. Neither the Bayley-II nor cognitive tests for older children have been developed for the purpose of detecting language deficits. Thus, language-sensitive tools are needed to identify those VLBW children with weak language skills.
A limitation of the present study was that receptive language was not analysed at the age of 2 years. This should be taken into consideration when the results are applied to a clinical context. Children with both weak early receptive and expressive skills have been shown to have the poorest prognosis in terms of later language development, at least in children with a hereditary risk for dyslexia (26). Furthermore, it has been shown that VLBW children acquire their receptive lexicons during the second year at a slower rate than full-term controls (27). VLBW children also had weaker receptive language ability at 2 years of age than the controls in the same study.

CONCLUSIONS
This study showed that a considerable number of VLBW children suffered from weak language skills at two and five years of age and that the percentage of VLBW children with weak language skills increased during the 5-year follow-up period. When children with neurological impairment were excluded, there was still a significant difference between the VLBW group and the controls in the percentage of children with weak language skills at the age of 5 years. The early weak language was a significant predictor of later weak language skills especially in the VLBW group. The findings propose that a considerable amount of VLBW children with weak language skills can be identified already at the age of 2 years. Based on the results, language-sensitive measures should be used to screen VLBW children in clinical work.

ACKNOWLEDGEMENTS
The present study was supported by the Academy of Finland. This study is part of the PIPARI study. The members of the PIPARI study group are Satu Eklblad, Eeva Ekholm, Leena Haataja, Mira Huhtala, Pentti Kero, Riikka members of the PIPARI study group are Satu Ekblad, Eeva


