# Follow up assessment
# of very preterm infants at five years of age

Martin Jan Koop de Kleine

Cover: 'Waakzaamheid' [Vigilance] by Beatrijs van Oijen, 2004.
Geese are considered symbols of vigilance since their honks alerted the citizens of ancient Rome in time to save their city from invading Gauls in 390 BC.


Follow up assessment of very preterm infants at five years of age.
Thesis Leiden University. With a summary in Dutch.
ISBN no:
NUR no:



Lay-out: Henk en Lars Dinnissen
Printed by Verhagen Grafische Media, Veldhoven, The Netherlands.

# Follow up assessment
# of very preterm infants at five years of age

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 17 maart 2005
klokke 16.15 uur

door

Martin Jan Koop de Kleine
geboren te Meppel
in 1948

**Promotiecommissie**

| | |
|---|---|
| Promotores: | Prof. dr. S.P. Verloove-Vanhorick<br>Prof. dr. L.A.A. Kollée<br>(Universitair Medisch Centrum St Radboud, Nijmegen) |
| Copromotor: | Dr. A.L. den Ouden<br>(Inspectie voor de Gezondheidszorg, Den Haag) |
| Referent: | Prof. dr. J.H. Kok<br>(Academisch Medisch Centrum, Amsterdam) |
| Overige leden: | Prof. dr. S. Bambang Oetomo<br>(Máxima Medisch Centrum, Veldhoven)<br>Prof. dr. F.J. Walther<br>Prof. dr. J.M. Wit |

The study was supported by a grant of the Dutch Health Organisations Praeventiefonds, project number 28-2756, and ZorgOnderzoek Nederland (ZonMw), project number 1010004-20.

Printing of this thesis has been made possible by financial support of:
Friso Kindervoeding
Wetenschapsfonds van Máxima Medisch Centrum
Abbott BV
Nycomed Nederland BV
Vygon Nederland BV
Azimed BV
The support is gratefully acknowledged.

*aan mijn vader*

*voor Beatrijs, Ruben, Merel en Matthijs*

**Contents**

Chapter 1

# Introduction and outline of the study

## Introduction

Survival of very preterm and/or very low birth weight infants increased considerably during the last decades [1-5]. A major part of this improvement is due to the use of steroids prior to preterm birth and the introduction of the treatment with artificial surfactant in the early 1990s [6-8].

As survival increased, so did the awareness that not all very preterm children survive without sequelae [9,10]. Mild developmental disturbances that interfere with the acquisition of every day skills and normal learning appeared to be very frequent and were not detected until school age. The expectation was that such developmental disabilities were temporary side effects and would diminish with increasing perinatal knowledge and experience. Whether this will happen is still an important question as more immature and vulnerable infants survive.

At hospital discharge major determinants of later outcome such as intraventricular haemorrhage and bronchopulmonary dysplasia are obvious, but the prediction of the long term outcome of the individual infant is difficult.

By the age of two years, chronic illnesses, especially chronic lung disease, and handicaps such as cerebral palsy, severe visual problems, deafness and severe developmental delay have been diagnosed in 10 to 20% of these children [11-14]. The incidence and extent of these impairments are stable in later years [12,15].

Diagnosis of mild motor problems, learning disabilities and problem behaviour, however, is mostly delayed until school age [10,16-19]. When with advancing age the demands imposed by society increase, the problem rate increases even further [15]. In a nationwide follow up of very preterm and/or low birth weight infants in the Netherlands, the need for special education increased from 19% at age nine to 28% at age fourteen [12]. Therefore, developing a programme for long term longitudinal follow up of these children is necessary. The first aim of such a follow up programme is the early detection of developmental disturbances that necessitate developmental intervention. The follow up programme should also evaluate the quality of perinatal care and contribute to the scientific knowledge of pathways by which specific risk factors or types of damage in the perinatal period lead to later developmental problems.

In 1993, 89% of all very preterm live-born infants the Netherlands were treated in one of ten tertiary perinatal intensive care centres [20]. Though treatment was similar in the ten centres, the follow up programmes were not standardised and differed between individual centres in the country. Therefore, a national working party on neonatal follow up presented in 1996 a programme for standardised follow up from birth to eight years of age that would offer an equal follow up programme to every individual survivor of neonatal intensive care and yield comparable follow up data for all [21].

For this programme, the Validation Study Group* developed a comprehensive assessment for the age of five years to enable paediatricians to identify, before school age, which children have developmental disturbances that may interfere with acquisition of daily skills and normal life. In addition, this comprehensive assessment should offer paediatricians a standardised tool to evaluate the perinatal care.

The Validation Study Group developed a three-step procedure, consecutively consisting of a questionnaire for the parents, a comprehensive Paediatrician's assessment and detailed standardised and domain-specific assessment batteries. Each step aimed at identifying children with definite developmental impairments in one or more domains, children with possible impairments and children without impairments. Children with definite or possible impairments should enter the next step in order to be referred for domain-specific treatment if necessary; the optimal children should labelled 'optimal' and discharged from further paediatric control.

In the first part of this thesis the comprehensive Paediatrician's assessment is validated in 566 very preterm and/or low birth weight children from three perinatal centres against a battery of standardised tests for cognitive, motor- and language development: the revised Amsterdam child intelligence test (IQ test [22]), the Movement Assessment Battery for Children (movement ABC) [23] and the full Child Behavior CheckList (CBCL) [24,25].

Such a Paediatrician's assessment can be considered as a diagnostic test in a population with a high prevalence of developmental problems. It should have a high sensitivity, a high negative predictive value and a likelihood ratio of a negative test close to zero: all children with a disturbance or a delay must be identified. It is also preferable that such an instrument has a good specificity, meaning that the children without a disturbance or a delay are being identified as such, so that the parents can be reassured and further examinations be prevented. In a population with such a high prevalence of disturbances some over-referral of children with a normal development is acceptable and specificity is less important than sensitivity. The performance of such a test should be determined by comparing the test with a 'gold standard' with 100% sensitivity and specificity. This ideal situation does not occur in medical practice. The 'most used and most accepted test' is often considered as 'gold standard'.

We hypothesised that a paediatrician could identify with the help of this comprehensive assessment at least 80% of the very preterm and low birth weight infants with developmental disturbances in the somatic, motor and behavioural domains that may interfere with normal education and normal life, but would have considerable difficulties detecting children with cognitive problems.

The second part of this thesis concerns a comparison between perinatal care for very preterm infants born in the early 1980s and 1990s and its effect on later outcome. This

was feasible because several parts of the described follow up programme had been used before in the assessment at age five years of a national cohort of very preterm and/or very low birth weight infants born in 1983, in the POPS study [26].

We hypothesised that the increased survival of very preterm infants would be accompanied by improved later outcome, i.e. that, related to the number of live-born very preterm infants, the incidence of children without impairments would increase and the incidence of children with impairments would remain similar.

## Objectives of this study

1. To develop and validate an assessment tool that enables paediatricians to identify, before six years of age, which very preterm and/or very low birth weight survivors of neonatal intensive care have developmental disturbances that may interfere with normal education and normal life.

   Specific research questions were:

   1. Is it possible to develop a paediatric assessment that covers all relevant developmental domains?
   2. Does this paediatric assessment identify, before school age, which survivors of neonatal intensive care, without severe handicaps, have developmental disturbances that may interfere with normal education and normal life?
   3. Which part of this paediatric assessment best identifies these children?

2. To investigate the relation between changes in perinatal care and short term and long term outcome, mortality as well as morbidity, of very preterm infants born in 1983 and 1993 respectively.

   Specific research questions were:

   1. Which changes in perinatal factors have taken place between 1983 and 1993 in the Netherlands?
   2. Have these changes lead to changes in mortality and morbidity?
   3. Is increased survival accompanied by improved later outcome?

## Outline of the study

Chapter 1 presents the objectives and the research questions.

Chapter 2 describes the development and evaluation of a follow up assessment designed to enable paediatricians to identify, before six years of age, which survivors of neonatal intensive care have developmental disturbances that may interfere with normal education and normal life. This Paediatrician's assessment is validated against a battery of standardised tests in 412 very preterm and/or very low birth weight infants.

Chapter 3 compares a paediatric judgement of motor performance, the neurological assessment according to Touwen and a screening of motor development (Denver Developmental Screening Test, DDST) with the movement Assessment Battery for Children in 396 very preterm and/or very low birth weight infants without severe handicaps at five years of age.

Chapter 4 compares a paediatric judgement and the subtests of the Denver Developmental Screening Test (DDST) with a standardised intelligence test in 368 very preterm or very low birth weight infants, without severe handicaps, to evaluate whether a paediatrician is able to identify prematurely born infants at five years of age that should be referred to a child psychologist for extensive standardised cognitive assessment.

Chapter 5 describes the speech and language development of 145 preterm and very low birth weight infants at five years of age. In this population the Dutch Language Screening Test, used by a paediatrician, was validated against standardised language tests performed by a speech therapist.

Chapter 6 presents the perinatal risk factors, mortality and neonatal morbidity of two geographically defined populations of very preterm infants born in 1983 (POPS cohort) and in 1992 - 1994 (Validation cohort). The changes over time between these cohorts are analysed.

Chapter 7 presents the visual, auditive, neuromotor, cognitive and behavioural development at five years of age of two geographically defined populations of very preterm infants, born in 1983 (POPS cohort) and in 1992 - 1994 (Validation cohort). The changes between these cohorts are analysed.

Chapter 8 presents a general discussion of these studies and recommendations for follow up at five years of age of very preterm and/or low for gestational age infants.

* The Validation Study Group consisted of A. Lya den Ouden (TNO Prevention and Health, Leiden and Health Care Inspectorate, Ministry of Health, The Hague), Ina Kloosterboer-Boerrigter (TNO Prevention and Health, Leiden), Anneloes van Baar (University of Tilburg), Louis A.A. Kollée, Maria W.G. Nijhuis-van der Sanden, Boudien J.M. van Kessel-Feddema, Meta Sondaar and Simone Knuijt (Radboud University Nijmegen Medical Centre, Nijmegen), Adri Ilsen, Aleid G. van Wassenaer, Judy M. Briët, Rinske Breur-Pieterse and Karin Koldewijn (Academic Medical Centre, Amsterdam), Martin J.K. de Kleine, Jan Buijs and Marion J. Wit (Máxima Medical Centre, Veldhoven).

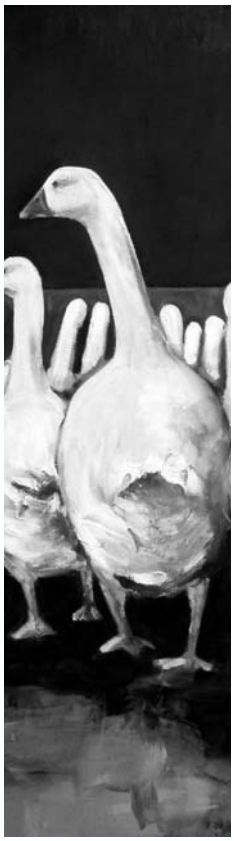
Ronald Brand (Department of Medical Statistics, Leiden University Medical Centre) advised on the statistical methods and analyses and S. Pauline Verloove-Vanhorick (TNO Prevention and Health, Leiden, and Department of Child Health, Leiden University Medical Centre) supervised the study.

## References


1. The Victorian Infant Collaborative Study Group. Improved outcome into the 1990s for infants weighing 500-999 g at birth. Arch Dis Child Fetal Neonatal Ed 1997;77:F91-4.
2. International Neonatal Network Scottish Neonatal Consultants Nurses Collaborative Study Group. Risk adjusted and population based studies of the outcome for high risk infants in Scotland and Australia. Arch Dis Child Fetal Neonatal Ed 2000;82:F118-23.
3. Battin M, Ling EW, Whitfield MF, Mackinnon M, Effer SB. Has the outcome for extremely low gestational age (ELGA) infants improved following recent advances in neonatal intensive care? Am J Perinatol 1998;15:469-77.
4. Lorenz JM. The outcome of extreme prematurity. Semin Perinatol 2001;25:348-59.
5. Lorenz JM, Paneth N, Jetton JR, den Ouden L, Tyson JE. Comparison of management strategies for extreme prematurity in New Jersey and the Netherlands: outcomes and resource expenditure. Pediatrics 2001;108:1269-74.
6. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care. Pediatrics 1998;102:893-9.
7. Crowley P. Prophylactic corticosteroids for preterm birth. Cochrane Database Syst Rev 2000;CD000065.
8. Soll RF. Synthetic surfactant for respiratory distress syndrome in preterm infants (Cochrane Review). In: The Cochrane Library, 1, 2002. Oxford: Update Software. 2002. Oxford, The Cochrane Library.
9. Wolke D. Psychological development of prematurely born children. Arch Dis Child 1998;78:567-70.
10. Foulder-Hughes LA, Cooke RW. Motor, cognitive, and behavioural disorders in children born very preterm. Dev Med. Child Neurol 2003;45:97-103.
11. The Victorian Infant Collaborative Study Group. Outcome at 2 years of children 23-27 weeks' gestation born in Victoria in 1991-92. J Paediatr Child Health 1997;33:161-5.
12. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in the Netherlands in 1983. Early Hum Dev 2000;59:175-91.
13. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. N Engl J Med 2000;343:378-84.
14. Rijken M, Stoelhorst GM, Martens SE, van Zwieten PH, Brand R, Wit JM, Veen S. Mortality and neurologic, mental, and psychomotor development at 2 years in infants born less than 27 weeks' gestation: the Leiden follow-up project on prematurity. Pediatrics 2003;112:351-8.
15. Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very-low-birth-weight infants. N Engl J Med 2002;346:149-57.
16. Wolke D, Meyer R. Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. Dev Med Child Neurol 1999;41:94-109.
17. Marlow N, Roberts L, Cooke R. Outcome at 8 years for children with birth weights of 1250 g or less. Arch Dis Child 1993;68:286-90.
18. Hille ET, den Ouden AL, Saigal S, Wolke D, Lambert M, Whitaker A et al. Behavioural problems in children who weigh 1000 g or less at birth in four countries. Lancet 2001;357:1641-3.
19. Anderson P, Doyle LW. Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. JAMA 2003;289:3264-72.
20. Kollée LA, den Ouden AL, Drewes JG, Brouwers HA, Verwey RA, S.P. [Increase in perinatal referral to regional centers of premature birth in The Netherlands: comparison 1983 and 1993]. Ned Tijdschr Geneeskd 1998;142:131-4.

21. Werkgroep Landelijke Neonatale Follow-up van de Sectie Neonatologie Nederlandse Vereniging voor Kindergeneeskunde. Follow up voor kinderen met perinatale risicofactoren. 1996.
22. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. Intelligentie-meting bij kinderen: empirische en methodologische verantwoording van de geReviseerde Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1987.
23. Henderson SE, Sugden DA. The Movement Assessment Battery for Children. San Antonio: Psychological Corporation, 1992.
24. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991.
25. Verhulst FC, van der Ende J, Koot HM. Manual for the Child BehaviorChecklist/4-18 (Dutch version). Rotterdam: Department of Childpsychiatry, Sophia Children's Hospital/Academic Hospital Rotterdam/Erasmus Universiteit Rotterdam, 1996.
26. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birhtweight infants at five years of age. The collaboration project on preterm and small for gestational age infants (POPS) in the Netherlands. Lancet 1991;338:33-6.

Chapter 2



# Development and evaluation of a follow up assessment of preterm infants at five years of age

M.J.K. de Kleine, A.L. den Ouden, L.A.A. Kollée, M.W.G. Nijhuis-van der Sanden, M. Sondaar, B.J.M. van Kessel-Feddema, S. Knuijt, A.L. van Baar, A. Ilsen, R. Breur-Pieterse, J.M. Briët, R. Brand, S.P. Verloove-Vanhorick.

Arch Dis Child 2003; 88: 870-5.

## Summary

### Background

Long term follow up shows a high frequency of developmental disturbances in preterm survivors of neonatal intensive care formerly considered non-disabled.

### Aims

To develop and validate an assessment tool that can help paediatricians to identify before six years of age which survivors have developmental disturbances that may interfere with normal education and normal life.

### Methods

A total of 431 very premature infants, mean gestational age 30.2 weeks (S.D. 2.0), mean birth weight 1276 grams (S.D. 332) were studied at age five years. Children with severe handicaps were excluded. The percentage of children with a correctly identified developmental disturbance in the domains cognition, speech and language development, neuromotor development, and behaviour were determined.

### Results

The follow up instrument classified 67% as optimal and 33% as at risk or abnormal. Of the children classified as at risk or abnormal 60% had not been identified at earlier follow up assessments.

The combined set of standardised tests identified a further 30% with mild motor, cognitive or behavioural disturbances. The Paediatrician's assessment had a specificity of 88% (95% C.I. 83 - 93%), a sensitivity of 48% (95% C.I. 42 - 58%), a positive predictive value of 85% (95% C.I. 78 - 91%) and a negative predictive value of 55% (95% C.I. 49 - 61%).

### Conclusions

Even after standardised and thorough assessment, paediatricians may overlook impairments for cognitive, motor and behavioural development.

Long term follow up studies that do not include detailed standardised tests for multiple domains, especially fine motor domain, may underestimate developmental problems.

## Introduction

Survival of very preterm and very low birth weight live-born infants has increased substantially in the 1980s and the 1990s, with some variation depending on the population investigated and mortality definition used [1-6]. Developmental outcome varies even more depending on the age of the child at assessment, the population studied and the definition of handicap used. Follow up during the first two years of life shows the presence of chronic illnesses, especially chronic lung disease, and handicaps such as cerebral palsy and severe developmental delay in 10 to 20% [7,8]. Long term follow up to school age reveals an even higher frequency of developmental impairments. These include motor performance problems, visual and auditory impairments, problems in cognitive and behavioural development and school failure [9-14]. In a nation-wide follow up of very preterm and very low birth weight infants in the Netherlands, the need for special education increased from 19% at age 9 to 28% at age 14 years [11]. Similar figures have been reported from Florida and Cleveland [15,16].

At the moment of hospital discharge, the prediction of the long term outcome of individual infants is difficult and not accurate. By the age of two years major handicaps such as gross motor disturbances, severe mental retardation, and chronic illnesses will have become clear. Diagnosis of learning disabilities, problem behaviour, and mild motor problems, however, is often delayed until school age. Therefore, long term longitudinal follow up of these children is necessary. Such follow up programmes should aim at the early detection of any developmental disturbances that necessitates developmental intervention as well as at the evaluation of perinatal treatment. A third goal is to contribute to the scientific knowledge of pathways by which specific types of damage in the perinatal period lead to developmental problems in childhood or even in adulthood and to recognise, at an earlier age, factors that predict developmental problems at a later age.

In the Netherlands the majority of very preterm infants are treated in one of ten perinatal intensive care centres. After treatment, survivors are referred to their local hospital for the remaining neonatal admission and post-discharge care. In the first years after birth they return a number of times to the centre of initial treatment for follow up assessments. Though treatment is similar in the ten perinatal centres, the follow up programmes are not standardised and post-neonatal care may differ throughout the country. Furthermore, because of a shortage in perinatal care the developmental departments are constantly urged to restrict follow up programmes to the bare minimum. Therefore a national working party on neonatal follow up designed a standardised follow up programme that would offer post-neonatal care as well as standardised follow up figures for all very preterm infants in the Netherlands.

In this paper we present the development of a Paediatrician's assessment for the most important developmental domains at five years of age and the validation of this assessment against a set of standardised test. This assessment should help paediatricians to identify, before the age of six years, which survivors of neonatal intensive care have developmental disturbances that may interfere with normal education and normal life.

## Population and methods

*Study population*

The study population included five year old survivors born before 32 weeks of gestation or weighing less than 1500 grams and treated in one of three Dutch neonatal intensive care units (NICUs). For logistic reasons the inclusion periods varied slightly: from 1 October 1992 to 15 June 1994, in the Radboud University Nijmegen Medical Centre, Nijmegen, from 15 November 1992 to 1 January 1994 in the Academic Medical Centre, Amsterdam, and from 1 January 1993 to 1 January 1995 in the Máxima Medical Centre, Veldhoven. During these time periods 764 infants were consecutively admitted to the three NICUs. One hundred thirty one (17%) died before the age of five years; and 46 patients below 30 weeks gestation and treated in the AMC were excluded because they participated in another study [17]. Twenty one children with known severe cerebral palsy, blindness, severe mental retardation, chromosomal abnormalities, or inborn error of metabolism were excluded, because it was beforehand obvious that they would not be able to perform the tests. As a result, 566 children were eligible for the study. Perinatal data were prospectively collected during admission, stored in the NICU-based databases and retrospectively retrieved for data analysis.

*Paediatrician's assessment*

The Paediatrician's assessment at five years consisted of a questionnaire sent to the parents and a structured assessment by a paediatrician, specially trained for this purpose.

The questionnaire addressed the following domains: socioeconomic status, general health, visual and auditory function, motor development, cognitive development, language and behaviour. Social status of the child was assessed by questions about residence, upbringing and number of siblings, and by questions about employment and education of both parents. General health was assessed by questions about medical consumption (visits to outpatient clinic, number of hospital admissions, number of

operations, and number of diagnostic tests), somatic symptoms of the ear, nose and throat (ENT), respiratory, gastrointestinal, and central nervous systems and the use of medicines. Visual and auditory functions were assessed by questions about hearing without or with hearing aids and vision without or with glasses. Cognitive development was assessed by questions about school performance, including learning problems and the need for remedial teaching or special school. Problem behaviour was addressed by a shortlist of those 15 problem items from the Child Behavior CheckList (CBCL) that discriminated most between normal children and children referred to mental health services in the general population [18].

The Paediatrician's assessment started with a check and further exploration of the data from the questionnaire, measurement of growth parameters and blood pressure. This was followed by a formal physical examination.

Neurological function was assessed according to Touwen [19]. This assessment consists of 40 items that cover tone, reflexes, muscle power, involuntary movements, posture, balance, coordination and gross motor skills. The children were assigned to the following categories: normal (no neurological abnormalities), minor neurological dysfunction (neurological abnormalities without influence on normal posture or movement) or cerebral palsy (neurological abnormalities with abnormal posture or movements).

The Denver Development Screening Test (DDST) was used to assess neuromotor and cognitive functioning [20]. The DDST consists of 105 test items, clustered into four groups, motor function, language, adaptation and social behaviour. Each test item has a cut off point at an age that 90% of Dutch children are able to perform it [20]. A child was classified as abnormal when he scored two or more delays beyond the cut off point in each of two groups or two delays in one group plus one delay in another group without compensation in the same group. He scored at risk when he scored two or more delays in one group or one delay in a group without compensation in the same group. In all other cases he scored normal.

Language was assessed by an age appropriate Dutch Language Screening Test, designed and validated by Gerritsen [21]. This test consists of 39 items covering the use of vocabulary, comprehension, memory and production of language, use of plurals and prepositions and pronunciation. Normal score is equal or less than 17, at risk ranges from 18 to 25 and abnormal from 25 to the maximal score of 52.

The results were summarised into five domains: cognitive development, neuromotor development, language development, behaviour and general health, and a final conclusion: (1) all domains normal, 'optimal', (2) further examination necessary, 'at risk' or (3) treatment necessary or already treated in one or more domains, 'abnormal'.

*Validation instruments*

The combined results of the Paediatrician's assessment were validated against a combined set of standardised tests for cognitive, motor, and language development. This set consisted of the revised Amsterdam child intelligence test (IQ test) [22], the Movement Assessment Battery for Children (movement ABC) [23], and the full Child Behavior CheckList (CBCL) [24,25]. Trained child psychologists and child physiotherapists performed these tests. They were blinded to the findings of the Paediatrician's assessment. The IQ test took one hour, the movement ABC 45 minutes.

The revised Amsterdam child intelligence test has been normalised for Dutch children between 4 and 7 years of age. Children with a score between -2 and -1 standard deviations were considered at risk; those below -2 standard deviations were abnormal.

The movement ABC indicates motor functioning in daily life [23]. Normative data have been collected in the USA, UK, and the Netherlands and yielded similar cut off points for normal American and Dutch children [23,26]. Total scores below or equal to 10.5 (15$^{th}$ centile) were considered normal, from 11.0 through 17.0 (5$^{th}$ centile) at risk and above 17.0 abnormal.

The CBCL is a standardised measure of child behaviour, developed by Achenbach and normalised for Dutch children by Verhulst [25,27]. Total scores up to and including 59 are considered normal, from 60 up to and including 63 intermediate, and from 64 upwards 'clinically important' disturbance of behaviour.

When all tests gave a normal result the final conclusion was 'optimal'; when the results of one or more of the three tests was 'at risk' or 'abnormal', the final conclusion was 'at risk' or 'abnormal' respectively.

*Statistical analysis*

Differences of continuous perinatal risk factors between groups were compared with the Student *t* test and differences between dichotomous risk factor with the chi square test; p values $< 0.05$ were considered statistically significant. Comparison of the Paediatrician's assessment with the formal assessments was expressed as sensitivity, specificity, and percentage correctly classified. Since no arbitrary cut off points were involved and classification into normal or abnormal was fixed by the test definition, ROC curves did not apply. Positive and negative predicted values, likelihood ratios and post-test probabilities estimated the performance of the test in practice.

## Results

Of the 566 eligible children, 135 (23.9%) were not assessed for various reasons (table 1), but language problems played an important role. Non-response rate was 65% (36/55) in children of non-Dutch and 19% (99/511) of Dutch speaking families (p = 0.000). Out of the 431 tested children, 395 completed all the batteries of the Paediatrician's assessment, IQ test, movement ABC and CBCL. The exact number of children that completed each test is included in the tables.

**Table 1:** Eligible and assessed patients.

| | N | % | N | % |
|---|---|---|---|---|
| Cohort 1992 - 1994 | 764 | 100.0% | | |
| Died | 131 | 17.2% | | |
| Excluded because of participation in an other study | 46 | 6.0% | | |
| Excluded because of severe handicap | 21 | 2.8% | | |
| | | | | |
| Eligible | 566 | 74.0% | 566 | 100.0% |
| Not assessed | | | | |
| Address unknown | | | 30 | 5.3% |
| Moved outside the country | | | 7 | 1.2% |
| Impossibility to make a convenient appointment | | | 52 | 9.2% |
| Refusal by the parents | | | 46 | 8.1% |
| Assessed | | | 431 | 76.2% |

Mean gestational age (30.2 (S.D. 2.0) weeks) and mean birth weight (1276 (S.D. 332) grams) were relatively high compared to most follow up studies from the 1990s. This illustrates the selection bias caused by non-inclusion of infants with gestational ages below 30 weeks from one of the three participating hospitals. There were no differences in perinatal data between assessed and non-assessed children, with the exception of multiple births (36% of the assessed versus 21% of the non-assessed children; table 2).

In 412 children a full Paediatrician's assessment was performed. The total assessment took approximately one hour and could be performed in a routine outpatient setting. The paediatricians involved felt it contributed highly to their ability to assess development in preschool children. Most children liked the different parts of the assessment and their parents were very satisfied with the thorough examination. The

assessment resulted in 275 (67%) children classified as optimal, 58 (14%) as at risk and 79 (19%) as abnormal. Of the 137 children classified as at risk or abnormal 82 (60%) had not been identified at earlier follow up assessments.

The IQ test classified 78% of the children as optimal and 22% as non-optimal, the CBCL 77% and 23%. The movement ABC, however, indicated 43% of the children with a motor problem (table 3). The combined results of the three formal assessments indicated significantly more children with developmental disturbances than the Paediatrician's assessment. Only 167 children (42%) scored optimal, 228 (58%) were identified as at risk or abnormal by at least one of the tests (table 4).

The specificity (88%) and positive predictive value (85%) of the Paediatrician's assessment were adequate, but many children with a developmental disturbance were not identified by it, resulting in a sensitivity of 48% and a negative predictive value of 55% (table 4).

Further investigation of the 20 children classified as abnormal or at risk by the Paediatrician's assessment but not by any of the validated tests showed that 10 had delayed language development according to the DDST or the language screening test, in two of them combined with either behaviour or cognitive problems. Three children had an abnormal neurological assessment and four showed abnormal behaviour during the assessment. Three of these children were suspected of a non-optimal cognitive development while their IQs proved to be 89, 95, and 97, respectively. Two children were considered as not optimal without further specification.

Of the 118 children that were incorrectly not identified by the Paediatrician's assessment the majority failed on the movement ABC: 54 scored between the P5 and P15, and 25 below the P5. In 24 of them either behaviour problems or cognitive delay or both accompanied this. One third (N=43) had a high total problem score on the CBCL, of whom 10 also had other problems. Nineteen children proved to have an IQ score of one S.D. or more below the mean, also mostly in combination with other problems.

Fifty four children had an abnormal outcome on both the Paediatrician's assessment and the validated tests. One third of them (18/54) were not identified before the follow up appointment and did not receive any intervention therapy. Furthermore, multiple problems did not enhance earlier recognition or intervention. Sixty five percent (55/85) of the children with a failure on more than one test were not identified previously.

**Table 2:** Risk factors in the study group and the non-response group.

| | Assessed N=431 | Not assessed N=135 | p |
|---|---|---|---|
| Male | 55% | 49% | |
| Multiple births | 36% | 21% | 0.001 |
| Congenital malformations | 10% | 7% | |
| Gestational age, wks (± S.D.) | 30.2 ± 2.0 | 30.1 ± 1.9 | |
| Birth weight, g (± S.D.) | 1276 ± 332 | 1327 ± 321 | |
| Caesarean section | 48% | 47% | |
| Apgarscore at 5' < 7 (N=565) | 17% | 14% | |
| Resuscitation including endotracheal ventilation | 24% | 28% | |
| Positive pressure ventilation | 49% | 49% | |
| Surfactant administration | 19% | 18% | |
| Bronchopulmonary Dysplasia | 14% | 17% | |
| Intraventricular haemorrhage, grade 1-4 | 19% | 26% | |
| Days in NICU (± S.D.) | 32 ± 31 | 32 ± 28 | |

**Table 3:** Results of the separate tests.

| | Paediatrician's assessment | Movement ABC | IQ test | CBCL |
|---|---|---|---|---|
| Number assessed per test | 412 | 404 | 407 | 407 |
| Optimal | 275 (67%) | 228 (56%) | 319 (78%) | 314 (77%) |
| At risk | 58 (14%) | 86 (21%) | 63 (16%) | 37 (9%) |
| Abnormal | 79 (19%) | 90 (22%) | 25 (6%) | 56 (14%) |

**Table 4:** Comparison between the Paediatrician's assessment and combined set of standardised tests (N=395 performed each test).

| **Conclusion of Paediatrician's assesment** | **Conclusion of the combined set of standardised tests** | | |
|---|---|---|---|
| | **At risk or Abnormal** | **Optimal** | **All** |
| At risk of abnormal | 110 | 20 | 135 |
| Optimal | 118 | 147 | 265 |
| All | 228 | 167 | 395 |
| | | *95% confidence interval* | |
| *Sensitivity* | *48%* | *[42; 55]* | |
| *Specificity* | *88%* | *[83; 93]* | |
| *Positive predictive value* | *85%* | *[78; 91]* | |
| *Negative predictive value* | *55%* | *[49; 61]* | |
| *Prior probability (prevalance)* | *0.58* | | |
| *Pre-test odds* | *1.4* | | |
| *Postive test* | | | |
| *Postive likelihood ratio* | *4.03* | *[2.61; 6.21]* | |
| *Post-test probability* | *0.85* | *[0.78; 0.89]* | |
| *Post-test odds* | *5.5* | | |
| *Negative test* | | | |
| *Negative likelihood ratio* | *0.59* | *[0.51; 0.67]* | |
| *Post-test probability* | *0.45* | *[0.41; 0.48]* | |
| *Post-test odds* | *0.8* | | |

## Discussion

Follow up of the first survivors of modern neonatal care that started in the 1970s showed that preterm birth has an effect well into adulthood [28]. Advances in neonatal care since the early days have led to an increase of survival. Developmental sequelae however, are still a major problem, mostly because babies who would previously have been expected to die are now surviving neonatal intensive care [8,29]. Follow up studies have shown that developmental problems increase with age, and seemingly healthy toddlers may still have developmental problems and school failure at a later age [30,31]. This makes long term follow up mandatory, both for timely identification of children in need of extra help and to enable a true evaluation of neonatal intensive care. There is no consensus about what should be measured at what age [30,32]. As a result, outcomes in different studies are not always comparable. To achieve reliable outcome measurements, it is necessary to include validated and multidisciplinary instruments into follow up programmes that assess different domains of development. However, funding for such intensive programmes is usually deficient. We therefore tried to develop a comprehensive follow up instrument, manageable by a paediatrician, that could differentiate between preschool children with optimal, and with suboptimal or abnormal development. This instrument proved suitable for use in an outpatient clinic, although it takes approximately one hour to assess a five year old child. The assessment enabled the participating paediatricians in this study to identify 137 children with developmental problems, of whom 82 (60%) were not identified at earlier assessments.

Studies that succeeded in reaching a virtually 100% follow up rate have shown that the disability rate in children that are hard to follow up is considerably higher than in children that are easy to follow up [33]. This selection bias means that the number of children with developmental problems is even higher.

The specificity of the assessment was quite high, 88%. Only 20 children were assessed ‘false positive’. Ten of them were considered to be at risk and 3 abnormal because of language or behavioural problems and therefore regarded as not optimal. This means that using this instrument as a screening instrument, unnecessary referral for full assessment will be infrequent.

However, only half of the children with abnormal results on one or more of the standard assessments were identified by the Paediatrician’s assessment. The majority of them had either motor problems or behaviour problems or a combination; a minority failed on the IQ test. It could be argued that developmental disturbances that do not give rise to suspicion at school or at home and are not identified by a standardised and thorough Paediatrician’s assessment are not important. On the other hand many of these children encountered delays in more than one domain. Mild impairments at the age of

five are often not disabling at that age and would perhaps not be insuperable when occurring alone. The combination of problems however, robs these children from the potential to compensate and puts them at risk for later learning disabilities and social isolation. Some of them may even become more disabled than their peers with a single, but more severe impairment.

Some studies underlined a relationship between motor dysfunction at an early age and cognitive problems at school age and later [34-38], although other studies did not [39]. The failure to recognise motor disturbances at an early age may be one of the reasons for the seemingly increase in developmental problems with increasing age. On the test of motor impairment (TOMI), the predecessor of the movement ABC used in our study, Powls et al. reported an improvement of motor function in half of the assessed preterm children between the ages of 6 and 8 years. However, at age 12 the percentage of school problems was 34% in children who failed on the TOMI at age 6 versus 5% of the children with optimal results [40].

Motor disturbances that predict later developmental problems may be found at a much earlier age [38,41]. Hadders-Algra and Groothuis reported an association between mildly abnormal general movements at fidgety age (2 - 4 months post term) and the development of attention problems and minor neurological dysfunction at the age of 4 to 9 years [41]. They hypothesised that perinatal hypoxia had resulted in long term changes in the striatic dopaminergic system. These disturbances could have influenced motor fluency and coordination in early childhood and behaviour in later life.

Motor performance is a result of information processing. The more task complexity increases, the more the information processing becomes important. As the movement ABC tests more than only motor domain, it is, in all likelihood, a good detector of the information processing capacity of the nervous system. We therefore assume that a high impairment score in the movement ABC tests does not only indicate a motor problem, but also identifies children who are at risk for developmental problems in other domains. By inclusion of the movement ABC in the Paediatrician's assessment only six children with cognitive delay would have been missed erroneously, five with IQ scores between -2 and -1 standard deviation and one below -2 standard deviations. Therefore we propose to include the movement ABC in the Paediatrician's assessment at age five, and also base the need for further assessment on the results of this test.

Children of non-Dutch speaking families were far more likely not to participate in the assessment than their Dutch peers (65% versus 19%). The majority of these families belonged to a low socioeconomic class, had low educational levels, and experienced cultural barriers when they looked for medical help. Moreover, an insufficient test result in children of non-Dutch speaking families was in some cases interpreted as a language problem and not as a cognitive problem. Sometimes this interpretation was correct,

sometimes it was not. As a result, the tests were often inconclusive in this population. Preterm children in these families are therefore threatened by a combination of biological and social risk factors that may cause long delays before they get the interventions they need.

In conclusion, the systematic and standardised Paediatrician's assessment at age five years did identify a significant number of children in whom the need for intervention was not recognised by the standard care they received. We also found that inclusion of the movement ABC might be necessary to avoid underestimation of a large number of developmental problems. Non-native children were seldom assessed properly.

Long term follow up studies of adolescents and young adults who were born preterm show that a developmental problem may pose a lifelong burden. Neonatal follow up should therefore aim to identify such problems at an early age and provide intervention therapy when needed. Furthermore, the evaluation of perinatal care should include late sequelae. Early predictors of these late sequelae are necessary in the research aimed at advancing perinatal care. We proved this could only be done with an extensive multidisciplinary assessment. Follow up studies that do not include detailed, standardised tests for several domains will underestimate developmental problems in survivors of neonatal intensive care and may be the reason for incomparability of follow up results and a seemingly increase of developmental problems at later ages. Our study indicates that there is no simple way to identify children in need of extra help or to give a true evaluation of neonatal intensive care.

## Acknowledgement

The authors thank Ina Kloosterboer-Boerrigter, Jan Buijs, Karin Koldewijn and Marion Wit for their participation in the Validation Study Group and Astrid Ehrnreich, Lili Jung, Brigit Hamers and Elle van de Bogaart for their help in studying the children. The authors thank Professor Dr. M. Hadders-Algra for her help in setting up this study and training the paediatricians and Professor Dr. M.J. Jongmans for her comments on the study and the manuscript.

The study was supported by a grant of the Dutch Health Organisations Praeventiefonds, project number 28-2756, and ZorgOnderzoek Nederland (ZonMw), project number 1010004-20.

## References


1. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care. Pediatrics 1998;102:893-9.
2. International Neonatal Network, Scottish Neonatal Consultants, Nurses Collaborative Study Group. Risk adjusted and population based studies of the outcome for high risk infants in Scotland and Australia. Arch Dis Child Fetal Neonatal Ed 2000;82:F118-23.
3. Battin M, Ling EW, Whitfield MF, Mackinnon M, Effer SB. Has the outcome for extremely low gestational age (ELGA) infants improved following recent advances in neonatal intensive care? Am J Perinatol 1998;15:469-77.
4. The Victorian Infant Collaborative Study Group. Economic outcome for intensive care of infants of birthweight 500-999 g born in Victoria in the post surfactant era. J Paediatr Child Health 1997;33:202-8.
5. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986;1:55-7.
6. Health Council of the Netherlands. Intensive care at birth. 2000. The Hague, Health Council of the Netherlands; publication no. 2000/08.
7. The Victorian Infant Collaborative Study Group. Outcome at 2 years of children 23-27 weeks' gestation born in Victoria in 1991-92. J Paediatr Child Health 1997;33:161-5.
8. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. N Engl J Med 2000;343:378-84.
9. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birhtweight infants at five years of age. The collaboration project on preterm and small for gestational age infants (POPS) in the Netherlands. Lancet 1991;338:33-6.
10. The Victorian Infant Collaborative Study Group. Neurosensory outcome at 5 years and extremely low birthweight. Arch Dis Child Fetal Neonatal Ed 1995;73:F143-6.
11. Hille ET, den Ouden AL, Bauer L, van den Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: perinatal risk factors and predictors at five years of age. Collaborative Project on Preterm and Small for Gestational Age (POPS) Infants in The Netherlands. J Pediatr 1994;125:426-34.
12. Vohr BR, Msall ME. Neuropsychological and functional outcomes of very low birth weight infants. Semin Perinatol 1997;21:202-20.
13. Taylor HG, Klein N, Schatschneider C, Hack M. Predictors of early school age outcomes in very low birth weight children. J Dev Behav Pediatr 1998;19:235-43.
14. Schaap AH, Wolf H, Bruinse HW, Smolders-de Haas H, Van Ertbruggen I, Treffers PE. School performance and behaviour in extremely preterm growth-retarded infants. Eur J Obstet Gynecol Reprod Biol 1999;86:43-9.
15. Resnick MB, Gomatam SV, Carter RL, Ariet M, Roth J, Kilgore KL, et al. Educational disabilities of neonatal intensive care graduates. Pediatrics 1998;102:308-14.
16. Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very-low-birth-weight infants. N Engl J Med 2002;346:149-57.
17. van Wassenaer AG, Kok JH, de Vijlder JJ, Briet JM, Smit BJ, Tamminga P, et al. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks' gestation. N Engl J Med 1997;336:21-6.

18. Verhulst FC, Akkerhuis GW, Althaus M. Mental health in Dutch children: (I). A cross-cultural comparison. Acta Psychiatr Scand Suppl 1985;323:1-108.
19. Touwen BC. Examination of the child with minor neurological dysfunction. Clinics in Developmental Medicine 1989;11:353-63.
20. Cools ATM, Hermans JMA. Dos-handleiding. Denver Ontwikkeling Screeningstest. Lisse: Swets & Zeitlinger, 1979.
21. Gerritsen FME. VTO Taalscreening 3- tot 6-jarigen: de ontwikkeling van een taalscreeningsinstrument voor gebruik in de Jeugdgezondheidszorg. Amsterdam/Lisse: Swets & Zeitlinger, 1988.
22. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. Intelligentie-meting bij kinderen: empirische en methodologische verantwoording van de geReviseerde Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1987.
23. Henderson SE, Sugden DA. The Movement Assessment Battery for Children. San Antonio: Psychological Corporation, 1992.
24. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991.
25. Verhulst FC, van der Ende J, Koot HM. Manual for the Child BehaviorChecklist/4-18 (Dutch version). Rotterdam: Department of Childpsychiatry, Sophia Children's Hospital/Academic Hospital Rotterdam/Erasmus Universiteit Rotterdam, 1996.
26. Smits-Engelsman BCM. Movement ABC; Nederlandse Handleiding [Dutch manual Movement ABC]. Lisse: Swets & Zeitlinger, 1998.
27. Verhulst FC, Berden GFMG, Sanders-Woudstra JAR. Mental health in Dutch children: (II) the prevalence of psychiatric disorder and relationship between measures. Acta Psychiatr Scand 1985;72:1-45.
28. Steward AL, Rifkin L, Amess PN, Kirkbride V, Townsend JP, Miller DH, et al. Brain structure and neurocognitive and behavioural function in adolescents who were born very preterm. Lancet 1999;353:1653-7.
29. Lorenz JM. Survival of the extremely preterm infant in North America in the 1990s. Clin Perinatol 2000;27:255-62.
30. McCormick MC. The outcomes of very low birth weight infants: are we asking the right questions? Pediatrics 1997;99:869-76.
31. Wolke D. Psychological development of prematurely born children. Arch Dis Child 1998;78:567-70.
32. Jobe AH. Predictors of outcomes in preterm infants: which ones and when? J Pediatr 2001;138:153-6.
33. Wariyar UK, Richmond S. Morbidity and preterm delivery; importance of 100% follow-up. Lancet 1989;8634:387-8.
34. Diamond A. Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. Child Dev 2000;71:44-56.
35. Botting N, Powls A, Cooke RW, Marlow N. Cognitive and educational outcome of very-low-birthweight children in early adolescence. Dev Med Child Neurol 1998;40:652-60.
36. Hadders-Algra M, Touwen BC. Minor neurological dysfunction is more closely related to learning difficulties than to behavioral problems. J Learn Disabil 1992;25:649-57.
37. Hadders-Algra M. Two distinct forms of minor neurological dysfunction: perspectives emerging from a review of data of the Groningen Perinatal Project. Dev Med Child Neurol 2002;44:561-71.
38. Gillberg IC, Gillberg C. Children with preschool minor neurodevelopmental disorders. IV: Behaviour and school achievement at age 13. Dev Med Child Neurol 1989;31:3-13.
39. D'Eugenio DB, Slagle TA, Mettelman BB, Gross SJ. Developmental outcome of preterm infants with transient neuromotor abnormalities. Am J Dis Child 1993;147:570-4.

40. Powls A, Botting N, Cooke RW, Marlow N. Motor impairment in children 12 to 13 years old with a birthweight of less than 1250 g. Arch Dis Child Fetal Neonatal Ed 1995;73:F62-6.
41. Hadders-Algra M, Groothuis AMC. Quality of general movements in infancy is related to neurological dysfunction, ADHD, and aggressive behavior. Dev Med Child Neurol 1999;41:381-91.

Chapter 3

# Assessment of motor development of very preterm and low birth weight infants at five years of age

M.J.K. de Kleine, M.W.G. Nijhuis-van der Sanden, A.L. den Ouden.

Submitted.

## Summary

Early detection of mild motor disturbances in very preterm and low birth weight infants is important because they interfere with the acquisition of every day's skills. To test motor performance problems in daily life we used the movement Assessment Battery for Children (movement ABC) in 396 five year old very preterm and low birth weight infants.

In a programme to develop a paediatric assessment of motor development we compared a paediatric judgement, a neurological examination (Touwen) and a screening of motor development (Denver Developmental Screening Test, DDST) with the movement ABC.

The movement ABC detected clinical important motor disorders in 22.0% and borderline disturbances in 21.5% of the children.

Compared to the movement ABC, the sensitivity of the paediatric judgement was 0.19, Touwen 0.62 and DDST 0.52; the negative predictive values were 0.61, 0.74 and 0.69.

We conclude that comparison of outcome on the above mentioned tests indicated that different tests detected different but sometimes overlapping motor capacities in children with non-optimal motor performance. As the paediatric assessments miss too many children with motor problems in daily life, the movement ABC should be added to the assessment of the motor development of very preterm and low birth weight infants at five years of age.

## Introduction

Motor development is often impaired in very preterm (VPT) and very low birth (VLBW) infants. This may range from severe and handicapping conditions like cerebral palsy in five to ten percent to motor performance problems in up to fifty percent [1-7]. While severe neurological handicaps are mostly detected early, milder functional motor problems are often not diagnosed before school age.

Early detection of functional motor problems is important, because these problems interfere with acquisition of everyday skills and social-emotional development [8,9]. They may thereby play a role in the development of behaviour problems that are often found in preterm born children [10]. Early intervention may improve the quality of life of these children. Moreover, developmental motor problems may be a marker of cerebral damage [11] or influenced by environmental factors [12]. Studying the specific markers for motor development problems is the first step to evaluate and improve the quality of perinatal care and follow up programmes in children at risk for developmental problems. Therefore, systematic, long term follow up of all very preterm and/or low birth weight infants is necessary.

In regular clinical practice systematic assessments are seldom performed because of shortage of budget and staff. Therefore, we developed, in a multicentre programme, a comprehensive Paediatrician's assessment of very preterm and low birth weight infants at five years of age and tested whether a standardised assessment by a paediatrician could accurately identify which very preterm and very low birth infants have functional problems in several domains at the age of five years. We found that paediatricians may overlook developmental disturbances especially in the motor domain [7].

Because motor problems are most present in very preterm and low birth weight children, the aim of this study is to compare the elements of this Paediatrician's assessment: paediatric judgement, neurological examination according to Touwen [13] and screening of motor development with the Denver Developmental Screening Test (DDST) [14], all assessed by the paediatrician, with the movement Assessment Battery for Children (movement ABC) [15], assessed by a paediatric physical therapist, and to analyse why these disturbances are overlooked. We choose the movement ABC as gold standard, because this test indicates motor function in daily life by testing those functions that children need at school every day, and is most widely used to diagnose developmental motor coordination disorder in several circumstances [16-18], as well as in follow up studies of preterm born children [4,6,8,19-28].

## Patients and Methods

*Study population*

The total study population consisted of 764 children of less than 32 weeks of gestation and/or weighing less than 1500 grams that were born in a time period between October 1992 and December 1994 and treated in three Dutch neonatal intensive care units: the Radboud University Nijmegen Medical Centre, the Academic Medical Centre, Amsterdam, and the Máxima Medical Centre, Veldhoven. Mortality before the age of five years was 17% (N=131). Forty-six patients with gestational ages below 30 weeks (6%) were excluded in the Academic Medical Centre because they participated in another study [29], and 21 (3%) were excluded because severe handicaps were present such as severe cerebral palsy, blindness, severe mental retardation, chromosomal abnormalities or inborn error of metabolism and it was obvious that they could not perform the tests at five years of age. Of the 566 eligible children, 431 (76%) responded and 396 (70%) had a full neurological and motor assessment by the paediatrician and a motor performance test by the paediatric physical therapist [7]. Mean gestational age was 30.2 ± 1.9 weeks and mean birth weight 1270 ± 326 grams; 55% were males and 49% had been ventilated. In 15% of them bronchopulmonary dysplasia had been diagnosed (Bancalari classification) and in 18% an intraventricular or periventricular haemorrhage (Papile classification). Both parents had had lower education in 18%; at least one of the parents had had middle education in 45% and higher education in 36%.

The perinatal data of the responding and non-responding children did not differ, with exception of a higher percentage of multiple births in the responding children [7].

*Method*

In each hospital trained paediatricians and paediatric physical therapists conducted the motor assessments. Appointments were limited to fixed time for each professional (paediatrician 60 minutes, psychologist 60 minutes and the physical therapist 45 minutes). The professionals were blinded for each other's findings; the physical therapist was not informed about the medical history of the children, but the medical file and a parental questionnaire did inform the paediatrician. The age of the children was not corrected for gestational age. Because motor performance may be related to cognitive and behavioural factors, the study included also a detailed questionnaire of school results, an IQ test [30] and the Child Behaviour Checklist (CBCL) [31]. The overall design of the study has been published in more detail elsewhere [7].

The Paediatrician's motor assessment consisted of an age-specific neurological

examination according to Touwen [13], the DDST [14] and an overall judgement that classified the child's neuromotor development as abnormal, at risk or optimal, directly noted down after the total examination.

**Table 1:** Neuromotor function examination according to Touwen.

| | |
|---|---|
| 1. Tone and posture | Resistance to passive movements in neck, trunk, or extremities; dorsoflexion of ankle; abnormal posture during sitting, standing, or supine; head, neck, or trunk control, downward drift of arms. |
| 2. Muscle power | Grasp of fingers; push away examiners hands with feet. |
| 3. Reflexes | Knee jerks, plantar response, foot grasp response, arm jerks (in case of tone abnormalities). |
| 4. Coordination | Finger-nose test, diadochokinesia; kicking examiner's hand and balance with feet; gait pattern, walking on heels, tiptoe, standing on one leg, hopping, standing with arms extended and eyes closed. |
| 5. Involuntary | Dyskinetic movements observed at rest and during voluntary movements; sitting, standing or supine motor restlessness; involuntary movements of tongue or eyelids, while standing with arms extended and eyes closed. |

The modified Touwen examination is often used by paediatricians in the Netherlands [32] and addresses five neurological domains: tone and posture, muscle power, reflexes, coordination and balance, and involuntary movements (table 1). Besides the overall judgement, a formal score was calculated on paper by the paediatrician. In addition the score was also calculated by computer, using the same algorithm as in the clinical examination: Coordination and balance (4; table 1) were considered normal if the child did not fail on more than two tests; in the other four domains (1, 2, 3 and 5; table 4) only one failure was accepted. If one or more of the domains was abnormal or there was definite asymmetry, the child was categorised as neurological abnormal. These abnormal neurological scores were divided into Minimal Neurological Dysfunction (MND) if there was no abnormality of posture and movement and cerebral palsy if posture or movement was abnormal. Cerebral palsy was

further divided into disabling cerebral palsy if sitting, standing, walking, or the use of hands was impaired and non-disabling cerebral palsy if there was no functional impairment. Children without neurological abnormalities but with delayed development in motor skills were categorised as gross motor retardation (GMR). GMR and MND were called 'at risk' and cerebral palsy was called 'abnormal'.

The Denver Development Screening Test (DDST; table 2) was used by the paediatrician to detect developmental problems [14]. According to the test instruction each of the four domains was classified as 'delayed' when the child scored two or more delays beyond the cut off point or one delay without a positive score as compensation in the same group. A child was classified as 'at risk' when one domain group was delayed, and 'abnormal' when two or more domains were delayed.

*Motor Function Assessment*

The outcomes of the paediatric judgement, the Touwen and the DDST were compared to a standardised norm-referenced motor assessment, the movement ABC, performed by trained paediatric physical therapists (table 3). The movement ABC indicates motor functioning in daily life [15] and was first published in 1966 by Stott, Moyes and Henderson as the Test of Motor Impairment (TOMI) [33]. It was revised in 1984 [34] and again in 1992 [15]. Normative data have been collected in the USA [35], UK [15], Hong Kong [36] and the Netherlands [37] and yielded similar cut off points for normal American and Dutch children. Because such normative data are present a control group was not necessary. The age band for age 4 - 6 years contains eight items, divided into three sections: manual dexterity, ball skills, and static and dynamic balance. There are two sorts of tasks at each item level: time related (scored in seconds) and error-related (scored by number of 'good' attempts). The aim of the test is to assess children's motor competence. The tester ascertains that the child understands the task before commencing. The raw score of the best attempt on each item is converted into a scaled score. The way this is done varies from task to task, but is well described in the test-manual. Per item scaled interval scores are provided: 0 = good and 5 = very poor. Both the preferred and non-preferred hands and feet perform some specific items; the child's scores are then summed and divided by two.

Simply summing the item scores for the three sections of the test produces a profile of the child's performance. The section scores for Manual Dexterity and for Static and Dynamic Balance vary from 0 to 15, the section score for Ball Skills from 0 to 10. These three section scores are then summed again to produce a Total Motor Impairment Score; ranging from 0 to 40. The three section scores and the Total Impairment Score can then be interpreted using the centile normative data tables.

Total scores below or equal to 10.5 (15th centile) were considered normal, from 11.0 through 17.0 (5th centile) at risk and above 17.0 abnormal.

To compare the results between the paediatric judgement, the Touwen examination, the DDST and the movement ABC we dichotomised the results. The 'at risk' and 'abnormal' group were grouped together as non-optimal and the normal group was named 'optimal'.

**Table 2:** The Denver Development Screening Test (DDST) [14]: items, relevant at the age of five years, arranged in four groups: social behaviour, adaptation, speech and language, and motor behaviour.

| | |
|---|---|
| 1. Social behaviour | Dress oneself with help |
| | Fasten buttons |
| | Dress oneself without help |
| 2. Adaptation | Building a tower of eight bricks |
| | Building a bridge of bricks |
| | Draws a circle |
| | Draws a cross |
| | Draws a person |
| | Draws a square |
| 3. Speech and language | Tells his/her first name |
| | Tells his/her first name and family name |
| | Tells the place he/she lives |
| | Recognises the colours red, green, blue and yellow |
| 4. Motor behaviour | Heel-to-toe walking forwards |
| | Heel-to-toe walking backwards |
| | Jumping at one place |
| | Jumping forwards |
| | 1 Second at one foot |
| | 5 Seconds at one foot |
| | 10 Seconds at one foot |
| | Hopping |
| | Throwing a ball |
| | Catching a ball |

**Table 3:** The Movement Assessment Battery for Children [15]: items, arranged in three sections, manual dexterity, ball skills and balance, at the age of five years.

| | |
|---|---|
| 1. Manual dexterity | Putting coins (12) in a slit with both hands separately as fast as possible (number of seconds)<br>Threading beads (12) as fast as possible (number of seconds)<br>Drawing between two lines as accurate as possible (number of failures) |
| 2. Ball skills | Catching a seed bag (number good of 10 trials)<br>Accuracy task: rolling a ball through a goal (number good of 10 trials) |
| 3. Static and Dynamic Balance | Standing at one foot (number of seconds)<br>Hopping over a cord (number of good attempts of three)<br>Heel-to-toe walking forwards over a line (number of good steps) |

*Statistical analysis*

Calculations for statistical analysis were performed in SPSS 11 (SPSS Inc.). To test if motor performance scores in the ex-premature population differed significantly from the distribution in the normal population a chi square test was performed; p values $< 0.05$ were considered statistically significant. Comparisons of the Paediatrician's assessments with the movement ABC were expressed as percentage correctly classified, Cohen's kappa, sensitivity and specificity. Positive predictive value (PPV), negative predictive value (NPV) and likelihood ratios estimated the performance of the assessments in practice [38].

## Results

Three hundred ninety six children were fully assessed by all motor tests and included in the analyses.

The movement ABC classified 22.0% children as abnormal and 21.5% as at risk, compared to 5% and 10% of a normal population (Spearman chi square $p < 0.001$; table 4).

The paediatrician classified 4.5%, 13.9% and 9.6% as abnormal and 4.8%, 22.2% and 24.2% as at risk, depending on the assessment used (table 4). Of the 55 children that were classified as abnormal on the Touwen examination 44 children were classified as disabling CP and 11 as non-disabling CP; of the 88 children that were scored as at risk, 21 were classified as MND and 67 GMR.

Sensitivity and negative predictive value were 0.19 and 0.61 for the paediatric judgement, 0.62 and 0.74 for the Touwen examination and 0.52 and 0.69 for the DDST (tables 5 and 6). Sensitivity and negative predictive value of the Touwen examination did not improve when it was combined with the DDST (sensitivity 0.56 [95% confidence limits 0.48; 0.63] and negative predictive value 0.73 [95% confidence limits 0.67; 0.78]).

Of the 396 children, 153 (38.6%) were classified as optimal by all instruments and 243 (61.4%) as non-optimal by at least one instrument. The overlap between the instruments varied (figure 1).

To get insight in the relations between the Touwen examination, the DDST and the movement ABC and the reasons for the differences, we analysed the ‘false’ positive and ‘false’negative results further and studied the subscores and the relation to the results of the IQ test and the Child Behaviour Checklist.

The 68 children (17%), central in the Venn diagram (figure 1) that scored non-optimal by all test were the most impaired children with also impairments in the language (N=42), cognitive (N=41) and to a lesser extend the behavioural domain (N=23).

The 44 children (11%), at the top of the Venn diagram, that scored non-optimal by the movement ABC and optimal by the Touwen examination and the DDST showed motor problems, but seldom cognitive impairment or behavioural problems. However, 19 had already school problems.

The 38 children (9.6%) that scored non-optimal by the movement ABC and Touwen examination and optimal by the DDST had predominantly motor problems and school problems (N=23), and less frequently language (N=12), cognitive (N=9) or behavioural (N=8) problems. These differences are explicable by the test instruction of the DDST allowing children enough time to meet the test demands.

**Table 4:** Classification scores by the paediatric judgement, the Touwen examintion, the Denver Development Screening Test (DDST) and the Movement Assessment Battery for Children in 396 very preterm and/or low birth weight survivors of neonatal intensive care units at five years of age.

| | **Paediatric judgement** | **Touwen** | **DDST** | **Movement ABC** |
|---|---|---|---|---|
| Abnormal | 18 (4.5%) | 55 (13.9%) | 38 (9.6%) | 87 (22.0%) |
| At risk | 19 (4.8%) | 88 (22.2%) | 96 (24.2%) | 85 (21.5%) |
| Optimal | 359 (90.7%) | 253 (63.9%) | 262 (66.2%) | 224 (56.6%) |

**Table 5:** Comparison between the neurological assessment by the paediatrican and the movement ABC in 396 very preterm and/or low birth weight survivors of neonatal intensive care at five years of age.

| **Paediatrician's judgement** | **Movement ABC** | | |
|---|---|---|---|
| | **At risk or Abnormal** | **Optimal** | **All** |
| At risk of abnormal | 32 | 5 | 37 |
| Optimal | 140 | 219 | 359 |
| All | 172 | 224 | 396 |
| | | *95% confidence interval* | |
| *Correctly classified* | *63%* | | |
| *Sensitivity* | *0.19* | *[0.13; 0.25]* | |
| *Specificity* | *0.98* | *[0.95; 0.99]* | |
| *Positive predictive value* | *0.86* | *[0.70; 0.95]* | |
| *Negative predictive value* | *0.61* | *[0.56; 0.66]* | |
| *Prevalence* | *0.43* | | |
| *Postive likelihood ratio* | *8.3* | *[3.3; 21]* | |
| *Negative likelihood ratio* | *0.83* | *[0.77; 0.90]* | |

**Table 6:** Comparison between the Touwen examination and the movement ABC in 396 very preterm and/or low birth weight survivors of neonatal intensive care at five years of age.

| **Touwen** | **Movement ABC** | | |
|---|---|---|---|
| | **At risk or Abnormal** | **Optimal** | **All** |
| At risk of abnormal | 106 | 37 | 143 |
| Optimal | 66 | 187 | 253 |
| All | 172 | 224 | 396 |
| | | *95% confidence interval* | |
| *Correctly classified* | *74%* | | |
| *Sensitivity* | *0.62* | *[0.54; 0.69]* | |
| *Specificity* | *0.83* | *[0.78; 0.88]* | |
| *Positive predictive value* | *0.74* | *[0.66; 0.81]* | |
| *Negative predictive value* | *0.74* | *[0.68; 0.79]* | |
| *Prevalence* | *0.43* | | |
| *Postive likelihood ratio* | *3.7* | *[2.7; 5.1]* | |
| *Negative likelihood ratio* | *0.46* | *[0.36; 0.56]* | |

The 22 children that scored non-optimal by the movement ABC and DDST and optimal by the Touwen examination were predominantly boys (N=17) with poor performance on the ball skills, sometime cognitive impairments (N=11) and frequently already school problems (N=14).

Thirty-seven children (27 [6.8%] + 10 [2.5%]) were identified by the Touwen examination as either cerebral palsy (N=13), MND (N=3) or GMR (N=21) and by the movement ABC as normal. Although their total movement ABC scores were within the normal range (below 10.5), 15 of them had a disharmonic profile with relatively poor performance on only one of the movement ABC sections: manual dexterity (N=1), ball skills (N=9) or balance (N=5). With one exception, they all had normal intelligence; seven had clinical important behavioural problems. Eighteen were boys, nineteen girls.

Thirty four children (8.6%) that were optimal by the movement ABC and the Touwen examination and ‘false’ positive by the DDST had normal motor performance. All but one had normal intelligence. Seventeen of them had impairments in the social behaviour domain and 17 in the speech and language subtest of the DDST. Eleven

(32%) had remedial teaching or attended special school, 12 of them had a relatively poor performance on only one of the movement ABC sections: ball skills (7) or balance (5).

**Table 7:** Comparison between the DDST and the movement ABC in 396 very preterm and/or low birth weight survivors of neonatal intensive care at five years of age.

| **DDST** | **Movement ABC** | | |
|---|---|---|---|
| | **At risk or Abnormal** | **Optimal** | **All** |
| At risk of abnormal | 90 | 44 | 134 |
| Optimal | 82 | 180 | 262 |
| All | 172 | 224 | 396 |
| | | *95% confidence interval* | |
| *Correctly classified* | *68%* | | |
| *Sensitivity* | *0.52* | *[0.45; 0.60]* | |
| *Specificity* | *0.80* | *[0.74; 0.85]* | |
| *Positive predictive value* | *0.67* | *[0.58; 0.75]* | |
| *Negative predictive value* | *0.69* | *[0.63; 0.74]* | |
| *Prevalence* | *0.43* | | |
| *Postive likelihood ratio* | *2.7* | *[2.0; 3.6]* | |
| *Negative likelihood ratio* | *0.59* | *[0.50; 0.70]* | |

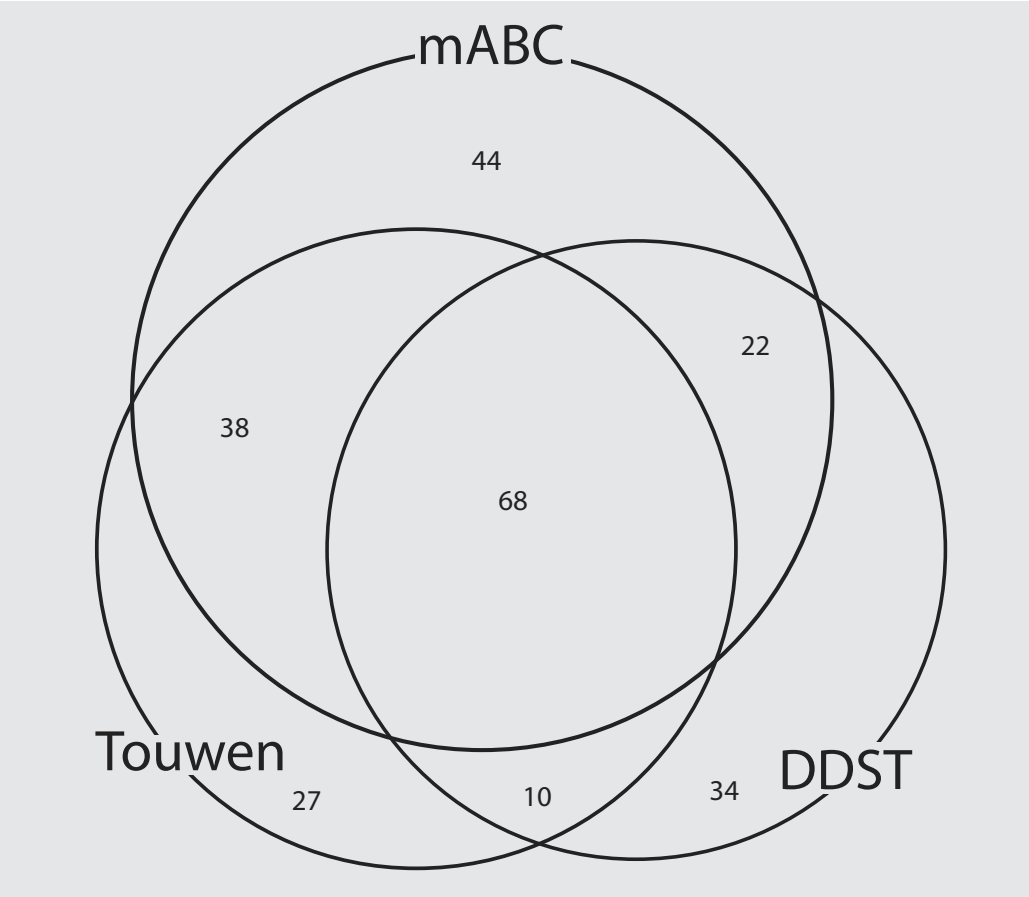



**Figure 1.**
*Venn diagram of 243 (61.4%) out of 396 very preterm and/or low birth weight survivors of neonatal intensive care at five years of age that were classified as non-optimal by at least one instrument. Each circle represents the number of children classified as non-optimal by respectively the movement ABC, Touwen's examination or DDST.*

## Discussion

In a programme to validate a comprehensive Paediatrician's assessment of the development of very preterm and low birth weight children at the age of five years, we evaluated the neurological examination according to Touwen and the Denver Developmental Screening Test in a population without severely handicapped children. These children were excluded because they could not perform the selected tests and were already treated in a rehabilitation programme.

Assessment and classification of children who experience difficulty in the acquisition of motor skills can be done in several ways [17]. The most common classification systems are the DSM IV and the ICD-10. The DSM IV labels them as Developmental Coordination Disorders. The DSM IV applies four criteria: (1) performance of daily activities that require motor coordination is substantially below that expected, given the person's chronological age and measured intelligence (2) the disturbance significantly interferes with academic achievement or activities of daily living, (3) the disturbance is not due to a general medical condition and (4) if mental retardation is present, the motor difficulties are in excess of those usually associated with it. The ICD-10 classifies the above-mentioned problems as Specific Developmental Disorders of Motor Function (SDD-MF).

The movement ABC is most widely used to make the first item of the DSM IV operational [16,17], and the few existing longitudinal studies performed show a relationship between abnormal movement ABC scores at the age of six years and later academic achievement in preterm infants [20]. The test has acceptable validity and reliability: inter-rater reliability ranges from 0.70 to 0.89, while test-retest reliability is 0.75. The test is very useful in identifying children with general motor difficulties and international comparison is possible [15,36,37,39].

The movement ABC has been used to quantify developmental motor coordination disorders in children after cardiac surgery [40], neonatal cerebral infarction [28], meningitis [41] in children with conditions such as Down's syndrome [42], Turner syndrome [43], ADHD and Asperger syndrome [18] and in recent years also in follow up studies of preterm infants [4,6,8,21-28]. For these reasons we used the movement ABC as gold standard to assess functional motor performance problems related to daily life situations.

With the movement ABC, we found in a population very preterm and low birth weight infants without very severe handicap an incidence of clinical important motor disorders of 22.0%, which is four times the incidence in the normal population [18] and another 21.5% borderline disturbances, about twice the incidence in the normal population. These incidences are slightly lower than the 20 to 30% of clinical important motor disorders in very preterm infants and 50 to 55% when borderline motor problems are included, found by several other authors [4,8,21,22,27,44]. The relative low incidence in our population compared to other studies can be explained by the exclusion of the children with severe handicap in our study and the loss of 46 children with gestational ages below 30 weeks in one of the participating hospitals.

Although for research purposes generally standardised tests are used, in practice most preterm infants are seen by a general paediatrician in a routine outpatient clinic. Therefore, we simulated this situation by asking the paediatricians, jointly trained for this study to minimise inter-observer variability, to describe their general impression

directly after the examination before formally counting the results of the Touwen and DDST on paper. The Paediatrician's judgement did identify fewer children with motor problems than the examination according to Touwen after counting and calculating with a fixed algorithm, the DDST and the movement ABC. This was not surprising. Most paediatricians assess few children without illness and problems and therefore their internal standard of normality drifts over time. Moreover, most paediatric assessment instruments have no rigid structure of taking the test and counting the results. The judgement of the observed results is also biased by the involvement in direct care of these patients, awareness of their history, and empathy with the parents.

For these reasons we trained the paediatricians to use the modified Touwen as a standardised neurological assessment. Although the paediatrician could use the Touwen scores in his judgement, the sensitivity of the Touwen examination improved from 0.19 to 0.62 using a fixed algorithm. The latter figure compares favourably to the 0.46 found by Jongmans et al. [4].

The agreement between the different parts of the paediatric assessment and the movement ABC varied. Although there is an influence of the absence of standardisation and objective counting when the Paediatrician's assessment is compared to the movement ABC, also on theoretical considerations a partial difference is not surprising. Within the individual child, functional motor capacity is not an isolated event, but emerges from physical abilities and conditions and executes through the cooperative effort of many brain structures and processes. When we study overt goal-directed motor behaviour in children, we are always looking at the result of a long chain of processes that produces and coordinates the observed movements. Therefore, motor performance problems are also related to cognitive, behavioural, visual-perceptual and health factors. The integrity of the central nervous system is an important structure for the development of perceptual-motor skills in preterm infants and many studies have focused on the relationship between cerebral abnormalities and developmental outcome [45-51]. The Touwen examination within the Paediatrician's assessment focuses on the integrity of the central nervous system. Therefore, we added the DDST to the Paediatrician's assessment to detect more general motor performance problems of the children. While in the Touwen tone and posture, muscle power, reflexes, coordination and balance, and involuntary movements are tested, the DDST screens social necessary and learned skills. However, we have to conclude that adjustment of the DDST did not improve the sensitivity of the Paediatrician's assessment compared to the movement ABC. Apparently the DDST is not sensitive enough to select children with motor performance problems.

Only 153 children scored optimal on all instruments and 68 children scored non-optimal on all tests, therefore in 221 children agreement was present, and in 175

children disagreement was present.

The most striking finding in the children with non-optimal movement ABC was the high percentage of children with poor school performance. Sixty-four percent of all children with non-optimal movement ABC had already school problems at the age of five years. We suppose that in these children learning capacity is decreased. Especially in the age of five years a lot of school activities focus on learning more complex motor tasks and skills i.e. drawing and painting, cutting with a scissor, ball skills, shoelace bowing etcetera. When learning capacity is decreased in the preterm population, it will be seen in motor tasks at this age of five years. However, when the children get older, more learning problems in other domains (social, cognitive) will appear, possibly related to concentration and attention problems or related to dysfunctional brain structures. The longitudinal development may be threatened as has been reported by the group of Marlow et al. [8,20].

From the 37 children identified by the Touwen examination as abnormal or at risk, and by the movement ABC as normal, 15 had a disharmonic profile with relatively poor performance on only one of the movement ABC sections: manual dexterity (N=1), ball skills (N=9) or balance (N=5). With one exception, all these children had normal intelligence and apparently these children had learned to function well with some constraints of their neuromotor system. By a lot of learning and training motor tasks can be learned and performance scores brought up within the normal range. These children met the accuracy and velocity requirements of the movement ABC at this age, but possible they become problems when they get older (ceiling effect). However, in clinical decision making these children are no problem: they might be referred for further assessment, but only longitudinal follow up is necessary to detect whether they learn new motor skills as well as their normal peers when they get older and the task demands increase.

Thirty four children were optimal by the movement ABC and the Touwen examination and ‘false’ positive by the DDST. All but one had normal intelligence; however, 17 of them had impairments in the social behaviour domain and 17 in the speech and language subtest of the DDST. Apparently the DDST is not sensitive for specific motor problems but more sensitive for general problems in social and language development. Eleven children had already remedial teaching or attended special school. They appeared to have more problems either in learning complex motor tasks, concentrating their attention and environmental interacting. This was confirmed by the fact that again 12 of them had a relatively poor performance on only one of the movement ABC sections: ball skills (7) or balance (5). Ball skills and to a lesser extend balance, are highly trained motor skills in which children need to interact with rather variable environmental factors and concentrate on the demanded task.

## Conclusion

The main goal of a paediatric screening assessment is to determine whether the quality and quantity of the child's performance is normal or abnormal for the age. After detecting an abnormal score, a more detailed multidisciplinary assessment is necessary to unravel which developmental aspects are the most determining factors.

We compared a paediatric judgement, a neurological examination and a screening of motor development with the movement Assessment Battery for Children and found that these tests apparently address interconnected but different motor capacities. The prevalence of developmental motor problems at the age of five years is high in very preterm and low birth weight infants. Many of them have also problems in other domains such as vision, hearing, cognition and behaviour and a striking high percentage had already problems at school at this early age.

In a population with such a high incidence of developmental problems it is necessary to detect as many children with developmental problems as possible. Some over-referral is acceptable. Therefore, tests should aim at high sensitivity, high negative predicting values and negative likelihood ratios close to zero.

As all tests were performed in the same population, the prevalence of impaired children was the same in all tests and assessments can be chosen on the basis of the negative predictive values. The paediatric judgement, Touwen examination and DDST had negative predictive values 0.61, 0.74 and 0.69. Although the Touwen examination performed best, even this test will miss too many children with developmental motor problems.

Therefore, we recommend using the movement ABC for the assessment of developmental motor problems in very preterm and low birth weight infants. A standardised neurological examination, such as the Touwen examination, however, should also be a part of the Paediatrician's assessment of these children to detect pure neurological impairments.

As over one half of all very preterm and low birth weight infants show developmental problems in several domains, often combined, routine assessment of all very preterm and low birth weight infants before six years of age is necessary. All developmental domains must be covered by standardised and norm-referenced test instruments, with rigid structure of taking the tests by trained examiners and counting the results. General judgement by paediatricians in a routine outpatient clinic, even when these paediatricians are well trained, is insufficient.

## Acknowledgements

The authors thank Ronald Brand, Department of Medical Statistics, Leiden University Medical Centre, for advice and Ina Kloosterboer-Boerrigter (TNO Prevention and Health, Leiden), Louis Kollée, Boudien van Kessel-Feddema, Meta Sondaar and Simone Knuijt (Radboud University Nijmegen Medical Centre, Nijmegen), Adri Ilsen, Aleid van Wassenaer, Judy Briët, Rinske Breur-Pieterse and Karin Koldewijn (Academic Medical Centre, Amsterdam), Jan Buijs and Marion Wit (Máxima Medical Centre, Veldhoven) for their participation in the Validation Study Group.

The study was supported by a grant of the Praeventiefonds, project number 28-2756 and ZorgOnderzoek Nederland (ZonMw), project number 10010004-20.

## References


1. Vohr BR, Msall ME. Neuropsychological and functional outcomes of very low birth weight infants. Semin Perinatol 1997;21:202-20.
2. The Victorian Infant Collaborative Study Group. Neurosensory outcome at 5 years and extremely low birthweight. Arch Dis Child Fetal Neonatal Ed 1995;73:F143-6.
3. The Victorian Infant Collaborative Study Group. Outcome at 2 years of children 23-27 weeks' gestation born in Victoria in 1991-92. J Paediatr Child Health 1997;33:161-5.
4. Jongmans M, Mercuri E, de Vries L, Dubowitz L, Henderson SE. Minor neurological signs and perceptual-motor difficulties in prematurely born children. Arch Dis Child Fetal Neonatal Ed 1997;76:F9-14.
5. Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very-low-birth-weight infants. N Engl J Med 2002;346:149-57.
6. Foulder-Hughes LA, Cooke RW. Motor, cognitive, and behavioural disorders in children born very preterm. Dev Med Child Neurol 2003;45:97-103.
7. de Kleine MJK, den Ouden AL, Kollée LA, Nijhuis-van der Sanden MW, Sondaar M, Kessel-Feddema BJ et al. Development and evaluation of a follow up assessment of preterm infants at 5 years of age. Arch Dis Child 2003;88:870-5.
8. Botting N, Powls A, Cooke RW, Marlow N. Cognitive and educational outcome of very-low-birthweight children in early adolescence. Dev Med Child Neurol 1998;40:652-60.
9. Hadders-Algra M, Groothuis AMC. Quality of general movements in infancy is related to neurological dysfunction, ADHD, and aggressive behavior. Dev Med Child Neurol 1999;41:381-91.
10. Hille ET, den Ouden AL, Saigal S, Wolke D, Lambert M, Whitaker A et al. Behavioural problems in children who weigh 1000 g or less at birth in four countries. Lancet 2001;357:1641-3.
11. Skranes JS, Vik T, Nilsen G, Smevik O, Andersson HW, Brubakk AM. Cerebral magnetic resonance imaging and mental and motor function of very low birth weight children at six years of age. Neuropediatrics 1997;28:149-54.
12. Bracewell M,Marlow N. Patterns of motor disability in very preterm children. Ment Retard Dev Disabil Res Rev 2002;8:241-8.
13. Touwen BC. Examination of the child with minor neurological dysfunction. Clinics in Developmental Medicine 1989;11:353-63.
14. Cools ATM, Hermans JMA. Dos-handleiding. Denver Ontwikkeling Screeningstest. Lisse: Swets & Zeitlinger, 1979.
15. Henderson SE, Sugden DA. The Movement Assessment Battery for Children. San Antonio: Psychological Corporation, 1992.
16. Geuze RH, Jongmans M, Schoemaker M, Smits-Engelsman B. Developmental coordination disorder. Hum Mov Sci 2001;20:1-5.
17. Henderson SE, Henderson L. Toward an understanding of developmental coordination disorder: terminological and diagnostic issues. Neural Plast. 2003;10:1-13.
18. Gillberg C, Kadesjo B. Why bother about clumsiness? The implications of having developmental coordination disorder (DCD). Neural Plast 2003;10:59-68.
19. Marlow N, Roberts BL, Cooke RW. Motor skills in extremely low birthweight children at the age of 6 years. Arch Dis Child 1989;64:839-47.
20. Marlow N, Roberts L, Cooke R. Outcome at 8 years for children with birth weights of 1250 g or less. Arch Dis Child 1993;68:286-90.
21. Powls A, Botting N, Cooke RW, Marlow N. Motor impairment in children 12 to 13 years old with a birthweight of less than 1250 g. Arch Dis Child Fetal Neonatal Ed 1995;73:F62-6.

22. Powls A, Botting N, Cooke RW, Marlow N. Handedness in very-low-birthweight (VLBW) children at 12 years of age: relation to perinatal and outcome variables. Dev Med Child Neurol 1996;38:594-602.
23. Katz KS, Dubowitz LM, Henderson S, Jongmans M, Kay GG, Nolte CA et al. Effect of cerebral lesions on continuous performance test responses of school age children born prematurely. J Pediatr Psychol 1996;21:841-55.
24. Katz-Salamon M, Gerner EM, Jonsson B, Lagercrantz H. Early motor and mental development in very preterm infants with chronic lung disease. Arch Dis Child Fetal Neonatal Ed 2000;83:F1-6.
25. Torrioli MG, Frisone MF, Bonvini L, Luciano R, Pasca MG, Lepori R et al. Perceptual-motor, visual and cognitive ability in very low birthweight preschool children without neonatal ultrasound abnormalities. Brain Dev 2000;22:163-8.
26. Cooke RW, Foulder-Hughes L. Growth impairment in the very preterm and cognitive and motor performance at 7 years. Arch Dis Child 2003;88:482-7.
27. Erikson C, Allert C, Carlberg EB, Katz-Salamon M. Stability of longitudinal motor development in very low birthweight infants from 5 months to 5.5 years. Acta Paediatr. 2003;92:197-203.
28. Mercuri E, Barnett A, Rutherford M, Guzzetta A, Haataja L, Cioni G et al. Neonatal cerebral infarction and neuromotor outcome at school age. Pediatrics 2004;113:95-100.
29. van Wassenaer AG, Kok JH, de Vijlder JJ, Briet JM, Smit BJ, Tamminga P et al. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks' gestation. N Engl J Med 1997;336:21-6.
30. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. Intelligentie-meting bij kinderen: empirische en methodologische verantwoording van de geReviseerde Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1987.
31. Verhulst FC, van der Ende J, Koot HM. Manual for the Child BehaviorChecklist/4-18 (Dutch version). Rotterdam: Department of Childpsychiatry, Sophia Children's Hospital/Academic Hospital Rotterdam/Erasmus Universiteit Rotterdam, 1996.
32. Schreuder AM, Veen S, Ens-Dokkum MH, Verloove-Vanhorick SP, Brand R, Ruys JH. Standardised method of follow-up assessment of preterm infants at the age of 5 years: use of the WHO classification of impairments, disabilities and handicaps. Report from the collaborative Project on Preterm and Small for gestational age infants (POPS) in The Netherlands, 1983. Paediatr Perinat Epidemiol 1992;6:363-80.
33. Stott DH. A general test of motor impairment for children. Dev Med Child Neurol 1966;8:523-31.
34. Stott DH, Moyes FA, Henderson SE. Test of Motor Impairment - Henderson Revision. Guelph, Ontario, Canada: Brook Educational Publishing, 1984.
35. Henderson SE, Hall D. Concomitants of clumsiness in young schoolchildren. Dev Med Child Neurol 1982;24:448-60.
36. Chow SM, Henderson SE, Barnett AL. The Movement Assessment Battery for Children: a comparison of 4-year-old to 6-year-old children from Hong Kong and the United States. Am J Occup Ther 2001;55:55-61.
37. Smits-Engelsman BCM. Movement ABC; Nederlandse Handleiding [Dutch manual Movement ABC]. Lisse: Swets & Zeitlinger, 1998.
38. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-based Medicine: How to practice and teach EBM. Edinburgh: Churchill Livingstone, 2000.
39. Chow SM, Henderson SE. Interrater and test-retest reliability of the Movement Assessment Battery for Chinese preschool children. Am J Occup Ther 2003;57:574-7.
40. Karl TR, Hall S, Ford G, Kelly EA, Brizard CP, Mee RB et al. Arterial switch with full-flow cardiopulmonary bypass and limited circulatory arrest: neurodevelopmental outcome. J Thorac Cardiovasc Surg 2004;127:213-22.

41. Stevens JP, Eames M, Kent A, Halket S, Holt D, Harvey D. Long term outcome of neonatal meningitis. Arch Dis Child Fetal Neonatal Ed 2003;88:F179-84.
42. Spano M, Mercuri E, Rando T, Panto T, Gagliano A, Henderson S et al. Motor and perceptual-motor competence in children with Down syndrome: variation in performance with age. Europ J Paediatr Neurol 1999;3:7-13.
43. Nijhuis-van der Sanden RW, Smits-Engelsman BC, Eling PA. Motor performance in girls with Turner syndrome. Dev Med Child Neurol 2000;42:685-90.
44. Hall A, McLeod A, Counsell C, Thomson L, Mutch L. School attainment, cognitive ability and motor function in a total Scottish very-low-birthweight population at eight years: a controlled study. Dev Med Child Neurol 1995;37:1037-50.
45. Allin M, Matsumoto H, Santhouse AM, Nosarti C, AlAsady MH, Stewart AL et al. Cognitive and motor function and the size of the cerebellum in adolescents born very pre-term. Brain 2001;124:60-6.
46. Barnett A, Mercuri E, Rutherford M, Haataja L, Frisone MF, Henderson S et al. Neurological and perceptual-motor outcome at 5 - 6 years of age in children with neonatal encephalopathy: relationship with neonatal brain MRI. Neuropediatrics 2002;33:242-8.
47. Bozynski ME, DiPietro MA, Meisels SJ, Plunkett JW, Burpee B, Claflin CJ. Cranial sonography and neurological examination at term and motor performance through 19 months of age. J Dev Behav Pediatr 1993;14:112-6.
48. Fily A, Truffert P, Ego A, Depoortere MH, Haquin C, Pierrat V. Neurological assessment at five years of age in infants born preterm. Acta Paediatr. 2003;92:1433-7.
49. Mercuri E, Jongmans M, Henderson S, Pennock J, Chung YL, de Vries L et al. Evaluation of the corpus callosum in clumsy children born prematurely: a functional and morphological study. Neuropediatrics 1996;27:317-22.
50. Pasman JW, Rotteveel JJ, de Graaf R, Stegeman DF, Visco YM. The effect of preterm birth on brainstem, middle latency and cortical auditory evoked responses (BMC AERs). Early Hum Dev 1992;31:113-29.
51. Pasman JW, Rotteveel JJ, Maassen B, de Graaf R, Visco Y. Diagnostic and predictive value of auditory evoked responses in preterm infants: I. Patient characteristics and long-term neurodevelopmental outcome. Pediatr Res 1997;42:665-9.

Chapter 4

# Recognition of cognitive developmental problems in very preterm infants by paediatricians

M. Sondaar, B.J.M. van Kessel-Feddema, A.L. van Baar, M.J.K. de Kleine, J.M. Briët, A.L. den Ouden.

Submitted.

## Summary

### Objective

To evaluate if a standardised paediatric assessment enables identification of very preterm and/or low birth weight infants at five years of age who should be referred to a child psychologist for a cognitive assessment.

### Methods

The outcome of a standardised paediatric assessment, consisting of a questionnaire, the Denver Development Screening Test (DDST) and a Dutch Language Screening Test, was compared with the results of a validated intelligence test in 368 very preterm and/or very low birth weight infants at five years of age.

### Results

The agreement between paediatric assessment of cognition and IQ was mostly poor (kappa < 0.40). The total DDST score and information on school performance were found to be the most sensitive (0.72 and 0.79) part of the paediatric assessment. Negative predictive value was high for total DDST (0.93), Dutch Language Screening Test (0.91) and school performance (0.94).

### Conclusion

A systematic and standardised paediatric assessment only poorly identifies preterm and/or very low birth weight infants with cognitive developmental problems. A low score on the DDST and information on extra assistance at school or grade retention best indicate which children should be referred for further standardised cognitive assessment by a professional child psychologist. To prevent school problems in very preterm and/or very low birth weight infants cognitive delays should be identified with formal assessment at an early age.

## Introduction

Preterm infants are at risk for developmental and school problems [1-3]. Long term follow up shows that up to 55% of these children experience cognitive and learning difficulties during school age. In a recent meta-analysis of case-control studies Bhutta et al. showed that the mean intelligence quotient (IQ) of very preterm infants during school age was approximately 10 points below that of healthy controls, depending on gestational age and birth weight [4]. Systematic developmental and cognitive assessment of all survivors of neonatal intensive care is necessary to evaluate perinatal care as well as to identify as early as possible which children may need extra stimulation or intervention. In practice however, these extensive assessments are seldom performed because of constraints in budget and staff. Therefore, a national working party on neonatal follow up designed a multidisciplinary and standardised follow up programme that would offer standardised follow up for all very preterm infants in the Netherlands. For this programme we evaluated a newly developed standardised assessment tool for the age of five years that would help the paediatrician to identify which very preterm and low birth weight survivors of neonatal intensive care have developmental disturbances that may interfere with normal education and normal life [5]. Such an instrument could be sufficient in itself for evaluation of perinatal care and secondary prevention purposes, or it could serve as a filter by identifying children that need further assessment.

In this paper we present the results of a comparison of the paediatric assessment in 368 very preterm and very low birth weight infants, not diagnosed with cognitive impairment before, with a formal cognitive assessment by a child psychologist. We studied whether the paediatric follow up instrument enables the paediatrician to identify the children that should be referred to a child psychologist for a more extensive cognitive assessment. We also evaluated which items of the paediatric assessment most clearly identify the children that need further cognitive assessment.

## Methods

*Study population*

The total study population consisted of 764 infants of less than 32 weeks of gestation or weighing less than 1500 grams that were born in a time period between October 1992 and December 1994. The infants were treated in three Dutch neonatal intensive care

units: the Radboud University Nijmegen Medical Centre, Nijmegen; the Academic Medical Centre, Amsterdam; and the Máxima Medical Centre, Veldhoven. Mortality before the age of five years was 17.2% (N=131). Forty-six patients with gestational ages below 30 weeks (6.0%) were excluded in the AMC because they participated in another study [6], and 21 (3%) were excluded because severe handicaps were present such as cerebral palsy, blindness, severe mental retardation, chromosomal abnormalities or inborn error of metabolism and it was obvious that they could not perform the tests at five years of age. Of the 566 eligible children 431 (76%) responded and 400 (71%) had a full standardised assessment by a paediatrician and an intelligence test by a psychologist [5].

Because the objective was to verify whether the Paediatrician's assessment could identify possible cognitive defects in children not yet known to have developmental difficulties, children with an already existing diagnosis of cognitive impairment and children in special education were excluded from the analyses (N=32; 8%).

*Procedure*

At five years of age the children were assessed on the same day by a paediatrician and child psychologist. Appointments were scheduled in a random order and were limited to a fixed time. The professionals were blinded for each other's findings. Only the paediatrician was aware of the detailed medical and perinatal history. Age was not corrected for preterm birth.

The institutional medical ethical review board of each of the participating hospitals approved the study and written parental consent was obtained.

*Paediatric assessment*

The follow up instrument consisted of a questionnaire sent to the parents and a standardised assessment by a paediatrician specifically trained for this purpose. The number of paediatricians was limited to a maximum of two per centre. For the present study we analysed the questions addressing school performance, origin (Dutch or non-Dutch) and educational level of mother (low, middle, high). School performance was defined in two categories: normal (no problems) or delayed (grade retention, remedial teaching or other forms of extra help).

The Paediatrician's assessment started with a check and further exploration of the data from the parental questionnaire and continued with paediatric examination, neurological assessment and assessments of cognition and behaviour [7]. For this part of the study the parental questionnaire and assessments of development and cognition were used. As part of the follow up instrument the paediatrician performed the Dutch

version of the Denver Development Screening Test (DDST) [8]. The DDST addresses four domains, adaptive behaviour, social development, language development and motor development, in children between 1 month and 6 years. It consists of 105 test items with a cut off point for each item at an age that 90% of Dutch children are able to perform it. Each domain is separately categorised as normal (no delays or one delay compensated by the performance of one item that is usually scored by less than 90% of a normal population) or abnormal (one delay without compensation or two or more delays). The overall DDST classification is normal (all four domains normal), at risk (one domain abnormal) or abnormal (two or more domains abnormal). Language and speech were furthermore assessed by a formal Dutch language screening test, the Dutch 'Taal Screenings Test' (TST) [9]. The TST is more elaborate than the language development part of the DDST. It consists of nine subtests used to examine different language abilities (naming objects with the same function, plural forms of nouns, repeating sentences, pointing out parts of the body, repeating words, complete sentences with conjunctions, knowledge of prepositions, analogies and antithesis and understanding and insight). At five years of age language development is categorised as normal (≤ 17 errors), moderately delayed (18 - 25 errors) or severely delayed (≥ 26 errors).

At the end of the assessment the paediatrician gave an overall judgement of cognitive development based on the performance at all tests and the overall impression of the child. The overall judgement of cognitive development was classified as: (1) normal, (2) re-assessment necessary in due time or (3) referral for further examination or treatment necessary.

*Intelligence test*

Trained child psychologists administered the short version of the 'Revised Amsterdam KinderIntelligentie Test'(RAKIT), a Dutch intelligence test devised for children between 4 years and 2 months and 11 years and two months of age [10]. This short version takes approximately 50 minutes and has a correlation of 0.93 with the full-scale test [11]. The subtests measure verbal capacities, perceptual and executive capacities, word fluency, memory and reasoning. The subtests are designed in a way to prevent cultural bias, which was done by reducing the influence of language. The concurrent validity with the Revised Wechsler Intelligence Scale for Children (WISC-R) is 0.86 for total IQ which makes international comparison possible [12]. The mean score is 100 with a standard deviation (S.D.) of 15. All scores ≥ 85 (-1 S.D.) are classified as normal, while all scores below -1 standard deviation (< 85) are classified as impaired. Scores between 70 and 85 were classified as dubious and scores below 70 as deviant.

*Statistical Analysis*

Statistical analyses were performed in SPSS 10.0 (SPSS Inc.). Differences in the normal distribution were compared with one-sample Kolmogorov-Smirnov, differences between means were compared with the Student *t* test; two-tailed comparisons were used and p values < 0.05 were considered statistically significant. Agreement in 2 x 2 tables was tested with Cohen's Kappa. The following criteria for kappa were used: < 0.40 = poor; 0.40 - 0.59 = fair; 0.60 - 0.74 = good; and ≥ 0.75 = excellent [13]. The diagnostic efficiency of the assessments by the paediatrician was defined with diagnostic efficiency statistics: the sensitivity, the specificity, the positive predictive value (PPV), the negative predictive value (NPV), the likelihood ratio of a positive test (LR+) and the likelihood ratio of a negative test (LR-) [14].

## Results

*Participants*

A total of 368 patients were included in this study. Mean gestational age was 30.2 ± 1.9 weeks, mean birth weight 1272 ± 329 grams. Of the study group, 54% were male, 35% multiple births, 48% had been artificially ventilated, 14% sustained a sepsis and 18% had an intraventricular haemorrhage in the neonatal period. Mean duration of neonatal intensive care was 28 ± 24 days.

Mean age at the time of assessment was 5 years and 2 months ( ± 1.9 months). Only 2% were from non-Dutch origin, 27% had a mother with a low and 20% with a high educational level.

*Intelligence test*

The RAKIT revealed that 83% of the children in this study had an IQ within the normal range, 13% had dubious scores (70 - 84) and 4% scored deviant (< 70). Mean IQ was 98.5 ± 14.6 with a range 56 - 142. The IQ scores were normally distributed (Kolmogorov-Smirnov 0.811, df 367, p = 0.53). In comparison to the Dutch population the difference was nearly significant (T-test -1.932, df 367, p = 0.05).

**Table 1:** Agreement between paediatric assessments of cognition and a formal intelligence test in 368 very preterm and/or low birth weight infants at five years of age.

| | **Impaired IQ N=61 (17%)** | **Normal IQ N=307 (83%)** | **Total N=368** | **Kappa** |
|---|---|---|---|---|
| DDST | | | | |
| At risk/Abnormal | 44 | 65 | 109 (30%) | 0.387 |
| Normal | 17 | 242 | 259 (70%) | |
| DDST-Adaptation | | | | |
| Abnormal | 17 | 4 | 21 (6%) | 0.360 |
| Normal | 44 | 303 | 347 (94%) | |
| DDST-Language | | | | |
| Abnormal | 19 | 24 | 43 (12%) | 0.265 |
| Normal | 42 | 283 | 325 (88%) | |
| DDST-Motor | | | | |
| Abnormal | 11 | 12 | 23 (6%) | 0.188 |
| Normal | 50 | 295 | 345 (94%) | |
| DDST-Social | | | | |
| Abnormal | 22 | 43 | 65 (18%) | 0.215 |
| Normal | 39 | 264 | 303 (82%) | |
| TST * | | | | |
| Moderately/severely delayed | 33 | 29 | 62 (17%) | 0.450 |
| Normal | 27 | 278 | 305 (83%) | |
| School performance | | | | |
| Delayed | 48 | 90 | 138 (37%) | 0.318 |
| Normal | 13 | 217 | 230 (63%) | |
| Overall judgement of cognitive development | | | | |
| Re-evaluation/Referral | 21 | 9 | 30 (8%) | 0.395 |
| Normal | 40 | 298 | 338 (92%) | |

* *One less tested with the TST*

*Outcome of cognitive developmental screening by the paediatrician*

On the DDST 70% of the children were classified as normal (table 1). Most delays were found on the language (12%) and social (18%) scales of the DDST. The TST identified slightly more language problems than the language scale from the DDST, classifying 17% of the children as moderately or severely delayed. School performance was delayed in 37% of the children. The overall judgement of the cognitive development by the paediatrician showed that they considered 92% of the children normal.

*Agreement between paediatric assessment of cognitive development and IQ test*

There was fair agreement (kappa = 0.45) between the TST and the IQ (table 1). The agreement between the other parts of the paediatric assessment of cognitive development and the IQ was poor (kappa < 0.40).

*Diagnostic efficiency of the paediatric assessments*

The complete DDST performed well with a sensitivity of 0.72, a negative predictive value of 0.93 and a likelihood ratio of a negative test of 0.35. The information of school performance agreed best with the IQ with a sensitivity of 0.79, a negative predictive value of 0.94 and a likelihood ratio of a negative test of 0.30. The TST had a negative predictive value of 0.91. All diagnostic efficiency values are presented in table 2.

**Table 2:** Diagnostic efficiency statistics for paediatric assessments of cognition in 368 very preterm and/or low birth weight infants at five years of age.

| | **Sensitivity** | **Specificity** | **PPV** | **NPV** | **LR+** | **LR-** |
|---|---|---|---|---|---|---|
| DDST | 0.72 | 0.79 | 0.40 | 0.93 | 0.41 | 0.35 |
| DDST-Adaptation | 0.28 | 0.99 | 0.81 | 0.87 | 21.39 | 0.73 |
| DDST-Language | 0.31 | 0.92 | 0.44 | 0.87 | 0.98 | 0.75 |
| DDST-Motor | 0.18 | 0.96 | 0.48 | 0.86 | 0.61 | 0.85 |
| DDST-Social | 0.36 | 0.86 | 0.34 | 0.87 | 0.58 | 0.74 |
| Dutch Language Screening Test | 0.55 | 0.91 | 0.53 | 0.91 | 0.82 | 0.50 |
| School performance | 0.79 | 0.71 | 0.35 | 0.94 | 0.68 | 0.30 |
| Overall judgement of cognitieve development | 0.34 | 0.97 | 0.70 | 0.88 | 11.74 | 0.68 |

## Discussion

The objective of this study was to investigate whether a standardised paediatric follow up instrument for very preterm and low birth weight infants can distinguish which children need further assessment of cognitive development at five years of age. We could not find other studies in the literature that compared a paediatric assessment to a standardised IQ score.

The results of our study show that the agreement between the paediatric assessment of cognition and the formal intelligence test is poor, except for a fair agreement between the TST and the formal intelligence test. The sensitivity of the overall judgement of cognition by the paediatrician and of the DDST subscales was low and underestimated the real need for formal cognitive testing. The complete DDST and school performance provided the best identification. Paediatricians often use the development of language as an indicator of cognition, but the sensitivity of the TST was rather low. Knuijt et al documented that the performance of this test may be improved by lowering the cut off point from 18 to 17 [15]. In a population at high risk for developmental problems, it is necessary to detect as many of the children with developmental problems as possible; some over-referral is acceptable. Therefore, test developers should aim at high sensitivity, high negative predicting values and likelihood ratios of negative tests close to zero. The complete DDST and school performance were found to have the highest sensitivity and negative predicting values. School performance apparently is a good indicator of cognitive difficulties, but the aim of early detection of cognitive problems is to detect them before they result in school problems. The importance of problematic school performance as a sign of cognitive problems was also found in an international study of four countries [3].

The results of the IQ test differ favourably from other studies on cognitive development of preterm children [2,16]. This partly results from our inclusion criteria; we studied a group of relatively healthy preterm infants. In line with the objective of the validation study, children known to be seriously handicapped were not invited. To investigate whether a paediatrician can identify the children that need further cognitive assessment, we also excluded children diagnosed with cognitive impairment before. Furthermore it is also known that actually examined children perform better than non-response children, who also have worse perinatal characteristics [17,18]. Therefore, our results do not reflect cognitive outcome for the total population of preterm survivors of neonatal intensive care.

Another possible explanation is that the intelligence test used underestimates mild developmental delay. Flynn has pointed out that IQ increases approximately 0.3 per year [19,20]. As a result standardised tests get outdated in time [21]. The RAKIT was

standardised in 1984. With a yearly increase of 0.3 point the mean IQ in a contemporary control group may well be 5 points higher than the norm we used. This change however, forms a problem for all standardised tests. Use of norm scores follows convention and allows comparability between studies. Furthermore, results of all kinds of assessments are influenced by estimation errors and intra-individual variation in accomplishments. This can be solved by repeated assessment and by the use of multiple sources and instruments in pre-treatment assessment of individual children with developmental disturbances.

The paediatric follow up instrument investigated in this study is found to be insufficient for a final evaluation of perinatal care. The decision to refer a preterm born child for further assessment may be based on the results of the combination of the DDST and the information on school performance. The information on school performance can be obtained easily and without further constraint and costs. Any child born after less than 32 weeks of gestation with school problems should immediately be referred for further examination. When the aim is to remediate or even prevent school problems, however, referral of the child when such school problems are obvious is too late and even then some children will be missed: the prevalence of cognitive impairment in preterm children with normal school performance was 6%. Since 37% of the children in this study already have school assistance or grade retention a preschool cognitive assessment is advisory. We therefore recommend not to wait until cognitive deficits have led to problems at school, but to test all very preterm and very low birth weight infants at five years of age with a thorough, formal cognitive assessment before school-entry at six years of age.

## Acknowledgements

The authors thank Louis A.A. Kollée, Maria W.G. Nijhuis-van der Sanden and Simone Knuijt (Radboud University Nijmegen Medical Centre, Nijmegen), Adri Ilsen, Rinske Breur-Pieterse and Karin Koldewijn (Academic Medical Centre, Amsterdam), Jan Buijs en Marion Wit (Máxima Medical Centre, Veldhoven), Ina Kloosterboer-Boerrigter (TNO Praevention and Health, Leiden) for their participation in the Validation Study Group.

The study was supported by a grant of the Praeventiefonds, project number 28-2756 and ZorgOnderzoek Nederland (ZonMw), project number 10010004-20.

## References


1. Anderson P, Doyle LW. Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. JAMA 2003;289(24):3264-72.
2. Davis DW. Cognitive outcomes in school-age children born prematurely. Neonatal Netw 2003;22(3):27-38.
3. Saigal S, den Ouden AL, Wolke D et al. School-age outcomes in children who were extremely low birth weight from four international population-based cohorts. Pediatrics 2003;112(4):943-50.
4. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 2002;288(6):728-37.
5. de Kleine MJK, den Ouden AL, Kollée LAA et al. Development and evaluation of a follow up assessment of preterm infants at 5 years of age. Arch Dis Child 2003;88(10):870-5.
6. van Wassenaer AG, Kok JH, de Vijlder JJ et al. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks' gestation. N Engl J Med 1997;336:21-6.
7. Touwen BC. Examination of the child with minor neurological dysfunction. Clinics in Developmental Medicine 1989;11:353-63.
8. Cools ATM, Hermans JMA. Dos-handleiding. Denver Ontwikkeling Screeningstest. Lisse: Swets & Zeitlinger, 1979.
9. Gerritsen FME. VTO Taalscreening 3- tot 6-jarigen: de ontwikkeling van een taalscreeningsinstrument voor gebruik in de Jeugdgezondheidszorg. Amsterdam/Lisse: Swets & Zeitlinger, 1988.
10. Bleichrodt N, Drenth PJD, Zaal JN, Resing WCM. Revisie Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1984.
11. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. Intelligentie-meting bij kinderen: empirische en methodologische verantwoording van de geReviseerde Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1987.
12. Pijl YJ, Hofman RH, Bleichrodt N, Resing WCM, Lutje-Spelberg H, de Bruijn E, Zaal JN. Vergelijkbaarheid van de WISC-R en RAKIT. Amsterdam: Research Instituut voor Onderwijs in het Noorden, Vrije Universiteit, 1984.
13. Cichetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessments instruments in psychology. Psychological assessment 1994;6: 284-90.
14. Bouter LM, Van Dongen MCIM. Epidemiologisch onderzoek: opzet en interpretatie. 4th ed. Houten: Bohn Stafleu Van Loghum, 2000.
15. Knuijt S, Sondaar M, de Kleine MJK, Kollée LAA. Validation of a Dutch Language Screening Instrument for 5-year-old preterm infants. Acta Paed 2004;73:1372-7.
16. Wolke D, Meyer R. Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. Dev Med Child Neurol 1999;41:94-109.
17. Wolke D, Sohne B, Ohrt B, Riegel K. Follow-up of preterm children: important to document dropouts. Lancet 1995;345:447.
18. Tin W, Fritz S, Wariyar U, Hey E. Outcome of very preterm birth: children reviewed with ease at 2 years differ from those followed up with difficulty. Arch Dis Child Fetal Neonatal Ed 1998;79:F83-7.
19. Flynn JR. Massive IQ gains in 14 nations: what IQ test really measure. Psychological Bulletin 1987;101:171-91.
20. Flynn JR. Searching for Justice: the discovery of IQ gains over time. American Psychologist 1999;54:5-20.

21. Wolke D, Ratschinski G, Ohrt B, Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. Eur J Pediatrics 1994;153(12):906-15.

Chapter 5

# Validation of a Dutch language screening instrument for five year old preterm infants

S. Knuijt, M. Sondaar, M.J.K. de Kleine, L.A.A. Kollée.

Acta Paediatr 2004; 93: 1372-7.

## Summary

### Aim

The validation of the Dutch Taal Screenings Test (TST), a language screening test, which is included in a follow up instrument developed to enable paediatricians to assess five year old preterm infants for their motor, cognitive and speech and language development.

### Methods

The speech and language development of 145 five year old infants born before 32 weeks of gestation and/or with a birth weight of less than 1500 grams was assessed by a paediatrician using the TST and by a speech therapist using standardised language tests.

### Results

All correlations between the instruments were significant. Using the original cut off point of the TST for abnormal speech and language development (18 points), the paediatrician will only identify 62% of the children who need speech therapy. For this group of children a cut off point of 17 is more effective. The positive predictive value of the TST improved from 77% to 82% by using a parent and school questionnaire to evaluate in a more subjective way the speech and language development.

### Conclusion

Using the TST, paediatricians will be able to identify speech or language problems in five year old preterm infants.

## Introduction

Speech and language is an important area of child development and can be influenced by many factors. Preterm delivery may negatively influence speech and language development [1-6]. Survival of very preterm infants has increased significantly [7]. Currently, much research is focused on the developmental outcome of these children and the potential risk factors for adverse outcomes. Unfortunately, ambiguous results over speech and language development have been reported, often due to methodological differences between studies. Some of these studies have been based on written questionnaires or personal interviews in which parents evaluated the speech and language development of their children, while others were based on standardised tests performed by professional speech therapists. The latter more often demonstrated a delay in speech and language development, compared to those based on questionnaires, probably because of parents' tendency to overestimate the capabilities of their children [1,3,8-11]. In addition, some studies were based on psychological testing only, not enabling correct interpretation of several speech and language aspects, such as articulation, fluency, grammar and receptive language [6,12-15].

Follow up evaluations of preterm infants are usually performed by paediatricians. They have to decide which children need further investigation of motor, cognitive or speech and language development. Therefore, the paediatrician needs an accurate screening instrument. Recently, a multicentre follow up (validation) study of a group five year old, very low birth weight infants was performed in the Netherlands in order to validate a follow up instrument that was developed for paediatricians screening motor, cognitive and speech and language skills in preterm infants at five years of age. To assess speech and language development, the Dutch Taal Screenings Test (TST) [16] was included in the follow up instrument.

In this study, to validate the TST in this group of preterm infants, the results of the TST performed by the paediatrician were compared with those of the standardised language tests performed by a speech therapist. The main goal was to examine the diagnostic efficiency and usefulness of the TST in accurately identifying preterm infants with a language delay. In addition, the issue was addressed whether the parents' and the paediatricians' personal impressions about speech and language development can be used to prevent over-referral.

## Methods

*Participants*

In a single centre (the Radboud University Nijmegen Medical Centre), the children were assessed by a speech therapist. The study population consisted of five year old neonatal intensive care unit (NICU) graduates, admitted between October 1992 and May 1994 and born before 32 completed weeks of gestation and/or with a birth weight of less than 1500 grams. Ten children with known severe cerebral palsy, blindness, severe mental retardation, chromosomal abnormalities or inborn errors of metabolism, in whom it was obvious that they would not be able to perform the tests, were excluded. Thirty seven infants could not be evaluated because they could not be traced, the parents refused participation, the language spoken in the family was non-Dutch or the available data were insufficient. As a result 145 (75.5%) were evaluated (table 1). The study population consisted of 78 boys and 67 girls with a mean chronological age of 5;01 years (years;months) (range 4;10 - 5;06). Mean gestational age was 30 weeks (range 25 - 38) and mean birth weight 1250 grams (range 570 - 2250).

The protocol was approved by the institutional Ethical Committee and the parents signed an informed consent.

**Table 1:** Number of children included and reasons for refusal and exclusion.

| | **N** |
|---|---|
| NICU graduates | 192 |
| Excluded because of severe handicaps | 10 |
| Moved | 13 |
| Parents refused participation | |
| Already treated somewhere else | 3 |
| Other reasons | 11 |
| Non-Dutch | 3 |
| Insufficient data | 7 |
| Total included | 145 |

*Procedure*

Each child was assessed within a half day time frame by a paediatrician, a psychologist, a physiotherapist and a speech therapist. Appointments were scheduled in random order. Although parents were present at the time of assessment by the paediatrician and speech therapist, they were instructed not to help their child. The paediatrician took approximately 10 minutes for language screening during a one hour physical examination and developmental screening session. The speech therapist used standardised language tests that took 45 minutes to complete. Both examiners were blind for each other's findings. Additionally, the speech therapist was not informed about the medical history. The final recommendation whether speech therapy was necessary or not was formulated during a multidisciplinary evaluation session.

*Follow up instrument*

The Dutch Taal Screenings Instrument (TSI), designed and validated by Gerritsen [16], was used by the paediatrician to assess speech and language development. The TSI contains a language screening test (TST), a school questionnaire and a parent questionnaire.

The TST has nine subtests, which examine different language abilities: naming objects with the same function, generating plural forms of nouns, repeating sentences, pointing out parts of the body, repeating words, completing sentences with conjunctions, knowledge of prepositions, generating analogies and antitheses, and understanding and insight. Based on the total number of errors, children were categorised as 'normal' (≤ 17 errors), 'moderately delayed' (18 - 25 errors) or 'severely delayed' (≥ 26 errors) [11].

Four normalised questions taken from the TSI school questionnaire, regarding vocabulary, pronunciation, receptive and expressive language respectively, were used by the paediatrician. They were awarded a maximum of 12 points. A score of 11 or 12 was regarded normal [16].

The parents' opinion about their child's speech and language development was evaluated based on four normalised questions from the TSI parent questionnaire. Again, a minimal score of 11 points was regarded as normal [16].

*Standardised language tests*

The speech therapist used the Dutch 'Reynell Test voor Taalbegrip' [17] and the Dutch 'Schlichting Test voor Taalproductie' [18], respectively, to assess speech and language development.

The Reynell Test is a standardised test that examines the receptive language development of Dutch-speaking children between 1;02 - 6;03 years of age. In this test, expressive language is not required since the children can point out their answers.

The Schlichting Test is a standardised test that examines the language expression of Dutch-speaking children of the same age group. The tests for syntactic and lexical development were applied. The first test elicited grammar structures and the second test measured the active vocabulary by instructing children to name objects.

The numbers of correct answers resulting from both the Reynell and the Schlichting tests were transformed into age-independent standard quotient scores. These scores had a mean of 100 and a standard deviation of 15, with extreme values at -3 and +3 S.D.. A score of less than -2 S.D. was considered to indicate deviant development, a score between -1 and -2 S.D. at risk and a score of -1 S.D. and higher reflected normal development.

*Statistics*

Statistical analyses were conducted with SPSS 10.0 [19]. Descriptive statistics were performed for the prevalence of language delay. To test the significance between means, Student *t* test was used. Chi square tests were used to establish differences between variables. To study the diagnostic efficiency of the TST, diagnostic efficiency statistics were implied, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and likelihood ratio [20]. Sensitivity is defined as the proportion of true-positive test results, i.e. an abnormal TST result obtained in a population of subjects with true language delays, where as specificity is the proportion of true-negative test results obtained in a population without language delays. The relationship between these values is the likelihood ratio: the chance of an abnormal TST result in the population of all subjects with language delays confirmed by language testing (i.e. sensitivity) divided by the chance of an abnormal TST result in the population of all subjects without any delays (i.e. 1-specificity).

Other characteristics describing diagnostic value include the positive predictive value (PPV), defined as the proportion of all children that failed the TST, which indeed had a language delay, confirmed with language testing. The negative predictive value (NPV) is the percentage of children that passed the TST and were tested normal in the language tests as well. All values were calculated with 2 x 2 tables with the Fisher exact test. Significance was set at $p < 0.05$.

## Results

*Follow up instrument*

An overview of all outcome measures is shown in table 2. Of the 145 children included, 121 (83%) scored within the normal range on the TST, 16 (11%) showed a moderate delay and 8 (6%) children showed a severe delay. The number of children in each category was as expected from a normally distributed population.

*Standardised language tests*

The mean quotient score of all children on the test for receptive language development was 101 (S.D. 14.87), which is not significantly different from the standard ($p = 0.439$). The test for syntactic development shows similar results. The mean quotient score was 97 (S.D. 15.31), again not significantly different from the standard ($p = 0.053$). The mean quotient score on the test for lexical development was 94 (S.D. 15.34), which differs significantly from the standard ($p < 0.001$). In 21 children, a asynchronous profile was observed, meaning a score below the standard on only one of the language tests.

*Validity analyses*

Content validity

The nine subtests of the TST, i.e. the items clustered according to the different language aspects, correlated significantly with the overall score of the TST ($p < 0.005$). All but two of the subtests correlated significantly with the language test that was supposed to assess the same language ability. Naming objects with the same function, testing vocabulary and sentence comprehension did not correlate with the test of lexical development ($p = 0.644$). The sentence completion task, composed to test the receptive and expressive language abilities, did not correlate significantly with the test for receptive language ($p = 0.052$), but correlated significantly with the tests for syntactic development ($p < 0.001$) and lexical development ($p = 0.013$).

The correlation between each of the three standardised language tests and the score on the TST was significant ($p < 0.001$). The mean result of the three language tests was calculated to represent the overall language test result. All correlations between this overall score and each of the nine subtests of the TST were significant ($p < 0.005$). It is remarkable that the children with an asynchronous profile on the language tests scored ‘normal’ on the TST.

The correlation between the parent questionnaire and the standardised language tests, as well as between the school questionnaire and the language tests, was significant ($p < 0.001$). Thus, both questionnaires gave a good impression of the speech and language development of the individual child.

**Table 2:** Overview of the outcome measures.

| | N | % |
|---|---|---|
| **Follow up instrument** * | | |
| Normal | 121 | 83% |
| Moderate delay | 16 | 11% |
| Severe delay | 8 | 6% |
| **Standardised language development** | | |
| *Receptive language development* † | | |
| Normal (> -1 S.D.) | 121 | 84% |
| At risk (-2 < x < -1 S.D.) | 19 | 13% |
| Abnormal (> -2 S.D.) | 3 | 2% |
| Missing | 2 | 1% |
| *Syntactic development* ‡ | | |
| Normal (> -1 S.D.) | 111 | 77% |
| At risk (-2 < x < -1 S.D.) | 25 | 17% |
| Abnormal (> -2 S.D.) | 3 | 2% |
| Missing | 6 | 4% |
| *Lexical development* ‡ | | |
| Normal (> -1 S.D.) | 108 | 75% |
| At risk (-2 < x < -1 S.D.) | 18 | 12% |
| Abnormal (> -2 S.D.) | 12 | 8% |
| Missing | 7 | 5% |
| *Mean quotient score* | | |
| Normal (> -1 S.D.) | 116 | 80% |
| At risk (-2 < x < -1 S.D.) | 26 | 18% |
| Abnormal (< -2 S.D.) | 3 | 2% |

* *TST*
† *Reynell Test*
‡ *Schlichting Test*

Criterion validity

In order to evaluate the diagnostic efficiency of the TST, the result of this test was compared with the mean score of the three standardised language tests. A score below 1 S.D. was considered to indicate speech and language delay.

First, the TST scores identifying moderate (18 to 25 errors) and severe language delay (26 errors or more) as originally defined for the categorisation of the test results were used to calculate the diagnostic efficiency statistics. When 26 errors in the TST were considered abnormal, the paediatrician identified 8 (28%) of the 29 children with a language delay confirmed by language testing. However, he would not refer any child without a language delay for speech therapy; therefore the PPV is 100%. Considering 18 errors or more to be abnormal, the paediatrician identified 18 (62%) of the 29 children with a language delay. However, in this situation he would refer 6 children (25%) without a confirmed language delay to a speech therapist. Therefore the PPV is 75% (table 3).

**Table 3:** Diagnostic efficiency statistics for different TST cut off points compared with standardised language tests.

| **Cut off** | **Value without questionnaires included** | **Value with questionnaires included** |
|---|---|---|
| *18 errors or more* | | |
| Sensitivity | 62% | 62% |
| Specificity | 95% | 97% |
| PPV | 75% | 82% |
| NPV | 91% | 91% |
| Likelihood Ratio | 12 | 21 |
| *17 errors or more* | | |
| Sensitivity | 79% | 79% |
| Specificity | 94% | 96% |
| PPV | 77% | 82% |
| NPV | 95% | 95% |
| Likelihood Ratio | 13 | 20 |
| *16 errors or more* | | |
| Sensitivity | 83% | 83% |
| Specificity | 91% | 94% |
| PPV | 71% | 77% |
| NPV | 96% | 96% |
| Likelihood Ratio | 9 | 14 |

In an attempt to optimise the mix of diagnostic efficiency statistics, other cut off points were evaluated. Considering 17 errors or more to be abnormal seemed to result in optimal statistics, with 79% sensitivity, 94% specificity and the highest likelihood ratio of 13 (table 3). In this scenario, the PPV is 77%, meaning that 23% of the children with abnormal TST results will unnecessarily be referred to a speech therapist. With a cut off point below 17, the sensitivity increases, but both the PPV and likelihood ratio will decrease (figure 1A).

The possibility to improve the PPV by taking the school and parent questionnaire into consideration was also evaluated. This was, however, only calculated for the area of the optimal cut off point, i.e. 16, 17 or 18 errors (figure 1B). When children who scored abnormal on the TST, but normal in both questionnaires were not identified as language delayed, the PPV increased by at least 6% (table 3).

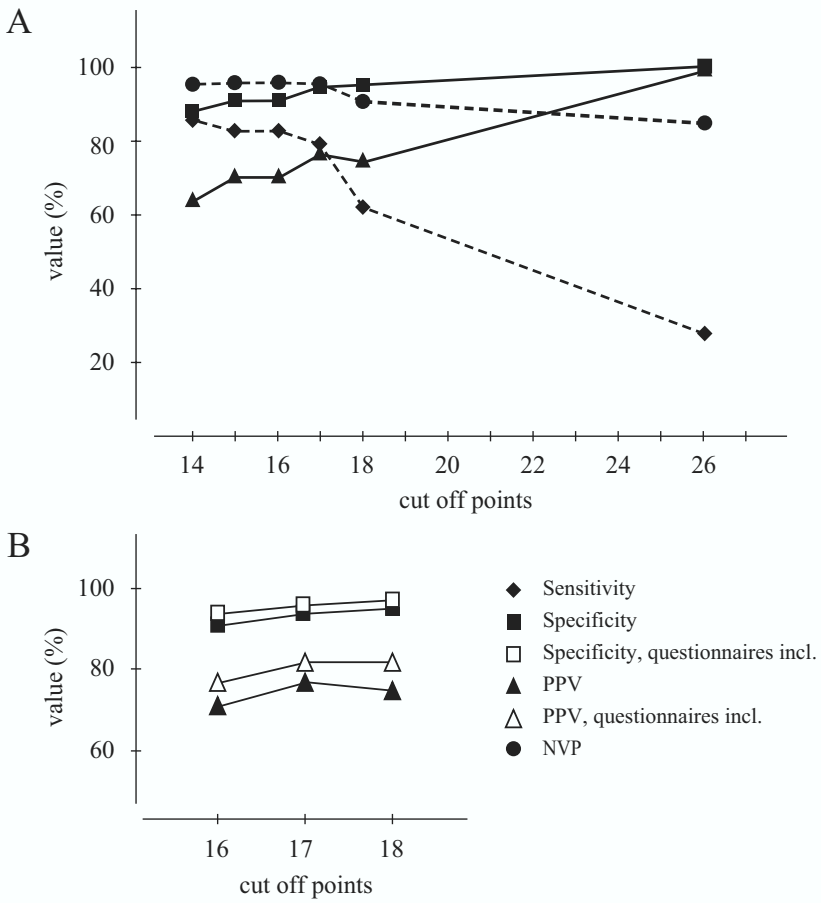



**Figure 1**. *Diagnostic efficiency statistics for different TST cut off points.*

## Discussion

It was earlier explained that within this study population of very preterm infants, the risk for specific language impairments is not increased [15]. This is confirmed by the results of the TST in the present study. These results are, however, in contrast to the results of a large follow up study in the Netherlands in 1983 (the POPS study) [21]. That study concluded that 40% of the children had a language delay. The difference in study populations can explain the differences between the results. In the present study the attendance was 75% and having a (severe) handicap was a reason not to invite children for the assessment. Therefore, the results are not representative for the whole population of preterm children. In addition, children who did not have Dutch as their first language were excluded, because the language test is normalised for monolingual Dutch children.

Since all correlations between the TST and the standardised language tests were significant, it can be concluded that the TST gives a good impression of the language development.

A screening instrument should have a good sensitivity (all children with a language delay must be identified). It is also preferably that it has a good specificity (children without a language delay are identified as such).

The original norm in the TST for a severe language delay (26 faults) seems to be too high, since only 28% of the children with language delays will be referred. Even the norm of 18 errors for a moderate delay can be discussed, since only 62% of the children with language delays will be identified. The optimal cut off point for identifying abnormal development in this population appeared to be 17. The positive predictive value increased when the parent and school questionnaire are included. This means an overall improvement in the quality of the referral. Thus, the strategy of referral by the paediatrician should be based on the TST result, plus the four questions from the school questionnaire and the opinion of the parents generated from the parent questionnaire. The findings, indicating the importance of the combination of the parents' opinion about the language development of their children, mixed with the results of language tests, are consistent with other studies [3,15,22].

According to Klee et al. [22], the positive predictive value of a language screening test should be as high as possible, because over-referral due to false positive results will lead to longer waiting lists. However, the TST is primarily designed to be a screening tool to identify children who should be referred to a speech therapist for further examination. The therapist in turn will decide whether the child needs therapy or not. In addition, a 10-minutes screening instrument such as the TST cannot be compared with a full examination by a speech therapist. Therefore, a practical a strategy as shown in figure 2 is proposed.

We conclude that the TST can be used as a screening instrument by the paediatrician in order to identify speech and language delays in five year old preterm infants when up to 16 faults are accepted for the judgement of a normal language development.

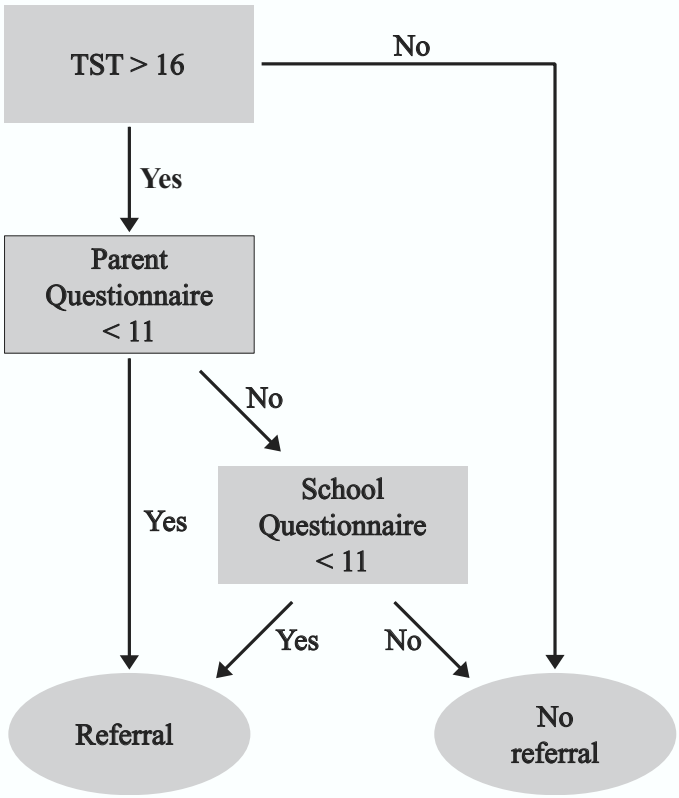



**Figure 2.** *Strategy for referral by the paediatrician.*

## Acknowledgements

The authors thank Lya den Ouden (TNO Prevention and Health, Leiden), Ria Nijhuis-van der Sanden and Boudien van Kessel-Feddema (Radboud University Nijmegen Medical Centre, Nijmegen), Judy Briët, Adri Ilsen and Rinske Breur-Pieterse (Academic Medical Centre, Amsterdam) and Anneloes van Baar (Máxima Medical Centre, Veldhoven) for their participation in the Validation Study Group.

The study was supported by a grant of the Praeventiefonds, project number 28-2756 and ZorgOnderzoek Nederland (ZonMw), project number 10010004-20.

## References


1. Cherkes-Julkowski M. Learning Disability, Attention-Deficit Disorder, and Language Impairments as Outcomes of Prematurity: A Longitudinal Descriptive Study. J Learn Disabil 1998;31:294-306.
2. Largo RH, Molinari L, Kundu S, Lipp A, Duc G. Intellectual outcome, speech and school performance in high risk preterm children with birth weight appropriate for gestational age. Eur J Pediatr 1990;149:845-50.
3. Jennische M, Sedin G. Speech and language skills in children who required neonatal intensive care: evaluation at 6.5 y of age based on interviews with parents. Acta Paediatr 1999;88:975-82.
4. Veen S, Ens Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. The Collaborative Project on Preterm and Small for Gestational Age Infants (POPS) in The Netherlands. Lancet 1991;338:33-6.
5. Brooks Gunn J, Liaw FR, Klebanov PK. Effects of early intervention on cognitive function of low birth weight preterm infants. J Pediatr 1992;120:350-9.
6. Luoma L, Herrgard E, Martikainen A, Ahonen T. Speech and language development of children born at < or = 32 weeks' gestation: a 5-year prospective follow-up study. Dev Med Child Neurol 1998;40:380-7.
7. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986;1:55-7.
8. Tomblin JB, Smith E, Zhang X. Epidemiology of specific language impairment: prenatal and perinatal risk factors. J Commun Disord 1997;30:325-43.
9. Eilers RE, Kimbrough Oller D, Levine S, Basinger D, Lynch MP, Urbano R. The Role of Prematurity and Socioeconomic Status in the Onset of Cononical Babbling in Infants. Infant Behavior and Development 1993;16:297-315.
10. Feldman HM, Janosky JE, Scher MS, Wareham NL. Language abilities following prematurity, periventricular brain injury, and cerebral palsy. J Commun Disord 1994;27:71-90.
11. Schreuder AM, Ens-Dokkum MH, Veen S, Verloove-Vanhorick SP, Brand R, Ruys JH. Language and speech delay at 5 years of age in very preterm and very low birthweight infants I: a follow-up study of a national cohort. Leiden: State University Leiden: 1993.
12. Grunau RV, Kearney SM, Whitfield MF. Language development at 3 years in preterm children of birth weight below 1000 g. Br J Disord Commun 1990;25:173-82.
13. Briscoe J, Gathercole SE, Marlow N. Short-term memory and language outcomes after extreme prematurity at birth. J Speech Lang Hear Res 1998;41:654-66.
14. Duncan N, Schneider P, Robertson CMT. Language Abilities in Five-Through-Seven-Year-Old Children Born at or Under 28 Weeks Gestational Age. J Med Speech-Lang Pathol 1996;4:71-9.
15. Knuijt S, Sondaar M. De spraak-taalontwikkeling van ex-premature kinderen op de leeftijd van vijf jaar. Tijdschr Kindergeneeskd 2001;2:43-9.
16. Gerritsen FME. V.T.O. Taalscreening 3- tot 6-jarigen. State University Leiden, 1988.
17. Van Eldik MCM, Schlichting JEPT, Lutje Spelberg HC, van der Meulen BF, van der Meulen S. Reynell Test voor Taalbegrip. Nijmegen: Berkhout Nijmegen B.V., 1995.
18. Schlichting JEPT, van Eldik MCM, Lutje Spelberg HC, van der Meulen S, van der Meulen BF. Schlichting Test voor Taalproduktie. Nijmegen: Berkhout Nijmegen B.V., 1995.
19. Statistical Package for the Social Sciences. Chicago Il. USA: 1999.
20. Bouter LM, van Dongen MCIM. Epidemiologisch onderzoek: opzet en interpretatie. 4 ed. Houten: Bohn Stafleu Van Loghum, 2000.

21. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in The Netherlands in 1983. Early Hum Dev 2000;59:175-91.
22. Klee T, Pearce K, Carson DK. Improving the positive predictive value of screening for developmental language disorder. J Speech Lang Hear Res 2000;43:821-33.

Chapter 6

# Lower mortality but higher neonatal morbidity over a decade in very preterm infants

M.J.K. de Kleine, A.L. den Ouden, L.A.A. Kollée, A. Ilsen, A.G. van Wassenaer, R. Brand, S.P. Verloove-Vanhorick.

Submitted.

## Summary

Better perinatal care has led to increased survival of very preterm infants, but may or may not have increased the number of children with cerebral and pulmonary morbidity. We therefore investigated the relation between changes in perinatal care through one decade and short term outcome in very preterm infants below 32 weeks' gestation.

Perinatal risk factors and their separate effects on mortality, intraventricular haemorrhage and bronchopulmonary dysplasia were compared in two surveys of very preterm infants in the Netherlands.

In-hospital mortality decreased between 1983 and 1993 from 63.5% to 38.0% in infants age < 28 weeks' gestation and from 19.5% to 12.0% in infants of 28 - 31 weeks' gestation. The incidence of intraventricular haemorrhage in survivors did not change: 42% and 43% in infants < 28 weeks' gestation and 27% and 21% in infants of 28 - 31 weeks' gestation. The incidence of bronchopulmonary dysplasia increased from 41% to 58% in infants < 28 weeks' gestation and remained similar in infants with gestational of 28 - 31 weeks' gestation (9% and 11%, respectively).

Adverse outcome, defined as in-hospital mortality, intraventricular haemorrhage and/or bronchopulmonary dysplasia, remained similar in infants < 28 weeks' gestation (83% in 1983 and 85% in 1993), but improved in infants of 28 - 31 weeks' gestation (48% in 1983 and 37% in 1993).

In multivariate analysis, congenital malformations, short gestational age, no administration of steroids before birth, low Apgar scores and artificial ventilation were associated with adverse outcome, in 1983 as well in 1993. Male gender and not being born in a perinatal intensive care unit were associated with adverse outcome in 1983, but not in 1993.

We conclude that adverse outcome in infants of 28 - 31 weeks' gestation has improved between 1983 and 1993, but not in those < 28 weeks' gestation, due an increased incidence of bronchopulmonary dysplasia.

## Introduction

Survival of very preterm infants (gestational age < 32 weeks) and very low birth weight infants (birth weight < 1500 grams) has increased substantially in the 1980s and 1990s. The rates vary depending on population investigated and mortality definition used. Survival of very low birth weight infants increased from less than 70% to over 80% and of extremely low birth weight (< 1000 grams) infants from less than 40% to over 50% in North America, Australia and Europe [1-3]. In extremely preterm infants (gestational age < 28 weeks), survival increased from 65% to 72% in Canada, from 49% to 72% in Australia and from 35% to 75% in the Netherlands [4-7].

The increased survival is the result of major changes in obstetric and neonatal strategies and therapies during these two decades. For instance, concentration of perinatal care in tertiary centres in the Netherlands increased between 1983 and 1993 from 61% to 89% of all very preterm infants and/or very low birth weight infants [8]. Obstetricians have made more consequent use of antenatal steroids and advanced ultrasound techniques and a more liberal application of Caesarean section became common practice. The introduction of artificial surfactant in neonatal intensive care since 1990 contributed to a 30% decrease in mortality in infants with respiratory distress syndrome (RDS) [9]. Improved neonatal technologies also included sophisticated artificial ventilation and total parenteral nutrition.

On the other hand, some demographic changes may have had a negative influence. Mean maternal age increased in the Netherlands from 27.9 in 1983 up to 29.8 years in 1995, resulting in increased infertility treatment and increased numbers of multiple births [10-12]. These changes not only increased the number of preterm births, but may also have had a negative effect on the condition of infants at birth [13].

It is an important question whether the decrease in mortality has led to fewer children with developmental disabilities due to decreasing common risk factors or to more children with disabilities due to an increase in other risk factors and survival of younger and more vulnerable babies.

To investigate the relationship between changes in perinatal care and short term outcome in very preterm infants, we compared perinatal risk factors and their effect on neonatal mortality and morbidity at discharge in two geographical surveys of very preterm infants born in 1983 and in 1993.

## Methods

Data of the 1983 population were obtained from the Project On Preterm and Small for gestational age infants, a nationwide follow up study including 94% of all live-born very preterm and/or very low birth weight infants in the Netherlands in 1983 [6,14].

The '1993' population consisted of very preterm infants born in three perinatal care regions in the Netherlands. For logistic reasons the inclusion periods differed slightly: from 1 October 1992 to 15 June 1994 in the region of the Radboud University Nijmegen Medical Centre, Nijmegen, from 15 November 1992 to 1 January 1994 in the region of the Academic Medical Centre, Amsterdam and from 1 January 1993 to 1 January 1995 in the region of the Máxima Medical Centre, Veldhoven. These three tertiary centres cover together approximately 30% of the Dutch neonatal intensive care. To avoid selection bias due to non-referral, data of all very preterm infants that were born in the area served by these three tertiary centres and that were not referred were retrieved from The Netherlands Perinatal Register and were included in the study [8].

From both populations only infants with a gestational age below 32 weeks were included. Perinatal data were prospectively collected during admission. Antenatal factors included cultural background (Dutch versus non-Dutch), maternal age, gender, congenital malformations (major and minor), multiple birth, gestational age, birth weight and small for gestational age (< 10$^{th}$ centile) [15]. As intrapartum factors we used level of care at birth, the administration of antenatal steroids, mode of delivery (vaginally versus Caesarean section) and Apgar scores at 5 minutes. As postnatal factors we used respiratory problems (need of artificial ventilation, use of surfactant and bronchopulmonary dysplasia), cerebral problems (intraventricular haemorrhage and seizures) and infections (sepsis and necrotising enterocolitis). Mortality was divided in first day, first week, neonatal, postneonatal, in-hospital and first year mortality. Intraventricular haemorrhage, diagnosed during admission, and bronchopulmonary dysplasia at 28 days of age was used to define neonatal morbidity, because these factors are strongly related with neurodevelopmental disabilities. Adverse outcome was defined as in-hospital death, intraventricular haemorrhage and/or bronchopulmonary dysplasia.

Definitions of all factors and outcome measures were similar in both cohorts: Intraventricular haemorrhage was defined according to the grading system by Papile et al.: grade I, subependymal haemorrhage; grade II, intraventricular haemorrhage, grade III, intraventricular haemorrhage with ventricular dilatation and grade IV, parenchymal haemorrhage [16]. Routine and systematic ultrasound of the brain was performed in six out of eight neonatal intensive care units in 1983 and in all centres in 1993 [17]. Ultrasound was performed soon after admission; further examinations were performed

at least twice in the first week of life and once weekly thereafter. In 1983, the incidence of intraventricular haemorrhages in those six centres that performed ultrasound routinely and systematically was 37% versus 29% in the other two centres (p = 0.042). Therefore, analyses of intraventricular haemorrhage and adverse outcome in 1983 were made for those six neonatal intensive care units only. Bronchopulmonary dysplasia in both cohorts was defined according to the definition of Bancalari (supplemental oxygen for more than 28 days), as this was the only available definition of bronchopulmonary dysplasia in 1983 [18].

When appropriate, outcome data were stratified for infants born after less than 28 weeks (< $28^0/_7$) of gestation and infants born between 28 and 32 weeks of gestation (gestational age 28 - 31 weeks; i.e. $28^0/_7$ up to and including $31^6/_7$ weeks).

*Statistical analyses*

Differences between means of continuous perinatal risk factors were compared with the Student *t* test and differences in proportions, displayed as 2 x 2 tables, with Fisher's exact test or chi square test; p values < 0.05 were considered statistically significant.

For multivariable analyses, perinatal risk factors were chronologically ordered, grouping together antenatal, perinatal and postnatal factors respectively. The impact of any risk factor was quantified by its Odds Ratio for the outcome considered. In equation: Log (Odds (outcome))= $\alpha + \beta 1R1 + \beta 2R2 + \beta 3R3 + \ldots\ldots + \beta(n)R(n)$ in which the perinatal risk factors are denotated by R1, R2, R3, ….. R(n), and the corresponding log (Odds Ratio) as $\beta 1, \beta 2, \beta 3, \ldots\ldots \beta(n)$. Adjusted OR's were computed by adjusting only for risk factors in the same group or a chronologically earlier one to avoid biases due to 'adjustment for consequences rather than causes'. Odds Ratios were estimated using a backwards stepwise procedure in which risk factors from an earlier chronological period were retained and risk factors from the same chronological period were removed if they were not statistically significant. However, risk factors from an earlier chronological period that were significant in that period were not removed while fitting the model for a later period, even if they became not significant then; in doing so we retained overall model consistency and had a proper interpretation of our OR's and p values. All these procedures are standard modelling procedures for regression analyses.

The outcome measure studied was 'adverse outcome', including in-hospital death, intraventricular haemorrhage and bronchopulmonary dysplasia. The modelling procedure was first performed separately for each of the two populations to obtain the OR's for all perinatal risk factors in each period. The outcome parameters death, intraventricular haemorrhage and bronchopulmonary dysplasia were not used as risk

factors in the modelling procedure for adverse outcome.

Because the magnitude of the effect (i.e. the OR) of each of the significant perinatal factors may have changed over time, we also analysed the significant perinatal risk factors in a model including year of birth (1983 or 1993). This model, including all perinatal risk factors and the interaction of these factors with the year of birth, was fitted to the combined populations. We could then calculate, using 'adverse outcome' as the dependent variable, the odds ratio for each clinical risk factor as well as the odds ratio quantifying the influence of year of birth on such a risk using the interaction terms. Significance of an interaction term corresponded to a significant change in the magnitude of that risk factor between 1983 and 1993.

## Results

*Perinatal data*

The 1983 population consisted of 1007 (252 < 28 weeks' gestation and 755 of 28 - 31 weeks' gestation and the 1993 population of 724 (150 < 28 weeks' gestation and 574 of 28 - 31 weeks' gestation) infants below 32 completed weeks' gestation (table 1). In 1983, all live-born infants had been included in the study, whether they were admitted to a paediatric or neonatal department or not; 24 infants (2.4%) had a gestational age below 25 completed weeks. In 1993 only the infants admitted to a neonatal department were included; 4 infants (0.6%) had a gestational age below 25 completed weeks. Therefore, mean gestational age was slightly higher in the 1993 population. Mean birth weight, gender and the prevalence of congenital malformations did not change. In ten years, mean maternal age increased by three years and the incidence of multiple birth increased from 26% to 34%.

Perinatal care changed considerably over time. Thanks to a policy of predelivery referral, the percentage of infants born in a tertiary perinatal centre increased from 42% to 76%. This significantly reduced the need for postnatal transport while at the same time the proportion of infants treated in a tertiary perinatal centre increased. In 1993, 94% of all infants were referred. Only 5 of the 40 non-referred infants had a gestational age below 28 weeks. The administration of antenatal steroids increased from 17% to 42%. The Caesarean section rate increased in infants < 28 weeks' gestation as well as in infants of 28 - 31 weeks' gestation. The proportion of small for gestational age infants increased only in the very preterm infants. These changes in perinatal care were associated with better Apgar scores in the infants < 28 weeks' gestation, but not in the infants of 28 - 31 weeks' gestation.

**Table 1:** Perinatal data in two cohorts (1983 and 1993) of live-born, very preterm infants with gestational age < 32 weeks, stratified by gestational age categories (< 28 and 28 - 31 weeks).

| | 1983 N=1007 | 1993 N=724 | p |
|---|---|---|---|
| *Gestational age < 28 wks* | *252* | *150* | |
| *Gestational age 28 - 32 wks* | *755* | *574* | |
| **Antenatal factors** | | | |
| Non-Dutch speaking parents[a] | 3.9% | 10.1% | < 0.01 |
| Maternal age (yrs; mean ± S.D.) | 27.3 ± 4.8 | 30.4 ± 4.9 | < 0.01 |
| Males | 54.9% | 56.9% | 0.43 |
| Multiple birth | 25.9% | 33.7% | < 0.01 |
| Congenital malformations | 9.5% | 11.4% | 0.20 |
| Gestational age (wks; mean ± S.D.) | 29.1 ± 2.0 | 29.4 ± 1.8 | < 0.01 |
| Birth weight (g; mean ± S.D.) | 1247 ± 349 | 1238 ± 363 | 0.58 |
| **Perinatal factors** | | | |
| Transport | | | |
| Inborn | 41.7% | 76.2% | < 0.01 |
| Postpartum transport | 32.8% | 8.2% | < 0.01 |
| No NICU admission | 25.5% | 5,5% | < 0.01 |
| *Gestational age < 28 wks* | *21.8% (N=55)* | *3.3% (N=5)* | *< 0.01* |
| *Gestational age 28 - 32 wks* | *26.8% (N=202)* | *6.1% (N=35)* | *< 0.01* |
| Antenatal steroids | 17.2% | 41.8% | < 0.01 |
| Small for gestational age (< P10) | 17.0% | 22.2% | < 0.01 |
| *Gestational age < 28 wks* | *13.1%* | *21.3%* | *0.02* |
| *Gestational age 28 - 32 wks* | *38.1%* | *49.7%* | *< 0.01* |
| Apgar 5′< 7 | 25.4% | 21.3% | 0.05 |
| *Gestational age < 28 wks* | *47.5%* | *30.7%* | *< 0.01* |
| *Gestational age 28 - 32 wks* | *18.2%* | *18.8%* | *0.41* |
| **Neonatal factors** | | | |
| Artificial ventilation | 52.7% | 59.1% | 0.01 |
| Duration of ventilation in days[b] | 5 (2-9) | 4 (2-10) | 0.23 |
| Surfactant[b] | - | 42.2% | - |
| Pneumothorax | 12.1% | 5.2% | < 0.01 |
| Bronchopulmonary dysplasia | 11.5% | 17.5% | < 0.01 |
| Intraventricular haemorrhage[c] | 36% | 27% | < 0.01 |
| Seizures | 6.6% | 5.4% | 0.43 |
| Sepsis | 10.9% | 17.3% | < 0.01 |
| Necrotising enterocolitis | 5.3% | 3.5% | 0.08 |
| IC days (median and quartiles)[d] | 26 (6.0-51) | 17 (7-39) | < 0.01 |

*a In 1983 documented in N=648 parents of surviving infants at the age of 2.*
*b Ventilated infants only. Median and (25 and 75$^{th}$ centile). Mann Whitney U test.*
*c For infants with systematic and routine ultrasound examination (1983: N=488, 1993: N=648).*
*d Median and (25 and 75$^{th}$ centile). Mann Whitney U test.*

**Table 2:** Mortality up to 1 year of age in two cohorts (1983 and 1993) of live-born, very preterm infants with gestational age < 32 weeks.

| | 1983 N=1007 | | | | 1993 N=724 | | | |
|---|---|---|---|---|---|---|---|---|
| | **N** | **Cum N** | **%** | **Cum %** | **N** | **Cum N** | **%** | **Cum %** |
| First day mortality* | 136 | | 13.5 | | 48 | | 6.6 | |
| 1 - 7 days | 108 | | 10.7 | | 46 | | 6.4 | |
| *First week mortality** | | *244* | | *24.2* | | *94* | | *13.0* |
| 8 - 28 days | 38 | | 3.8 | | 17 | | 2.3 | |
| *Neonatal mortality ( - 28 days)** | | *282* | | *28.0* | | *111* | | *15.3* |
| In-hospital mortality (after 28 days) | 25 | | | | 15 | | | |
| *In-hospital mortality (all)** | | *307* | | *30.5* | | *126* | | *17.4* |
| Post-discharge mortality > 28 days and < 1 year | 18 | | 1.8 | | 3 | | 0.4 | |
| *Mortality 0 - 1 year** | | *325* | | *32.3* | | *129* | | *17.8* |

* *p < 0.001*

**Table 3:** Total number of patients and in-hospital mortality in two cohorts (1983 and 1993) of live-born, very preterm infants with gestational age < 32 weeks, stratified by gestational age categories (< 28 and 28 - 31 weeks).

| | 1983 | | | 1993 | | |
|---|---|---|---|---|---|---|
| | **Total** | **In-hospital mortality** | | **Total** | **In-hospital mortality** | |
| | **N** | **N** | **%** | **N** | **N** | **%** |
| | 1007 | 307 | 30.5% | 724 | 126 | 17.4% |
| Gestational age < 28 wks* | 252 | 160 | 63.5% | 150 | 57 | 8.0% |
| Gestational age 28 - 31 wks* | 755 | 147 | 19.5% | 574 | 69 | 12.0% |

* *p < 0.001*

More infants were ventilated. The introduction of surfactant in the 1990s was associated with a lower incidence of pneumothorax, but not of shorter duration of artificial ventilation; the occurrence of bronchopulmonary dysplasia even increased. The incidence of sepsis increased and that of necrotising enterocolitis remained the same.

*Mortality*

Neonatal mortality decreased from 28.0% to 15.3%, in-hospital mortality from 30.5% to 17.4% and one-year mortality from 32.3% to 17.8% (table 2). Mortality decreased in each time period, but especially on the first day of life. Mortality in 1993 was approximately 40% lower than in 1983 in each gestational age week between 25 and 32 (table 3).

*Neonatal morbidity in infants discharged alive from hospital*

Between 1983 and 1993, there were no major changes in the incidences of intraventricular haemorrhages in survivors (table 4). The incidence of the severe grade 3 and 4 haemorrhages remained similar; grade 2 haemorrhages decreased in infants with gestational ages 28 - 31 weeks. The incidence of intraventricular haemorrhages was considerably higher in infants < 28 weeks' gestation than in infants of 28 - 31 weeks' gestation.

Bronchopulmonary dysplasia was significantly more common in 1993, due to a significant increase in the extremely preterm infants (table 4). In 1993, half of the infants < 28 weeks' gestation with bronchopulmonary dysplasia were treated with steroids, a possible risk factor for behavioural disturbances; 9% were still on oxygen at 38 weeks postmenstrual age (data not shown in the tables).

*Adverse outcome*

Adverse outcome, defined as in-hospital mortality, intraventricular haemorrhage or bronchopulmonary dysplasia decreased significantly in the infants of 28 - 31 weeks' gestation, but not in the infants < 28 weeks' gestation (table 5).

In the multivariable analyses, congenital malformations, short gestational age, no administration of steroids before birth, low Apgar scores and artificial ventilation were all associated with adverse outcome, in 1983 as well in 1993 (table 6). The magnitudes of these effects did not change. Male gender and being outborn were associated with adverse outcome in 1983 but not in 1993. Higher birth weight was, even after correction

**Table 4:** Intraventricular haemorrhage and bronchoplumonary dysplasia in infants discharged alive from hospital in two cohorts (1983 and 1993) of live-born, very preterm infants with gestational age < 32 weeks, stratified by gestational age categories (< 28 and 28 - 31 weeks). Data on intraventricular haemorrhages in 1983 are restricted to the six neonatal intensive care units that performed ultrasound of the brain routinely and systematically.

| | **1983** | **1993** | **p** |
|---|---|---|---|
| *All gestational ages < 32 wks* | *N=334* | *N=598* | |
| No haemorrhage | 71% | 75% | 0.141 |
| Grade 1 | 13% | 16% | 0.150 |
| Grade 2 | 9% | 3% | < 0.001 |
| Grade 3 | 6% | 5% | 0.763 |
| Grade 4 | 2% | 1% | 0.181 |
| *Gestational age < 28 wks* | *N=52* | *N=93* | |
| No haemorrhage | 58% | 57% | 1.0 |
| Grade 1 | 17% | 24% | 0.407 |
| Grade 2 | 13% | 10% | 0.582 |
| Grade 3 | 8% | 8% | 1.0 |
| Grade 4 | 4% | 2% | 0.618 |
| *Gestational age 28 - 31 wks* | *N=282* | *N=505* | |
| No haemorrhage | 73% | 79% | 0.080 |
| Grade 1 | 12% | 15% | 0.236 |
| Grade 2 | 9% | 1% | < 0.001 |
| Grade 3 | 5% | 5% | 0.734 |
| Grade 4 | 1% | < 1% | 0.195 |
| *All gestational ages < 32 wks* | *N=700* | *N=598* | |
| Bronchopulmonary dysplasia | 14% | 18% | 0.026 |
| Steroids for BPD | -- | 5% | |
| *Gestational age < 28 wks* | *N=92* | *N=93* | |
| Bronchopulmonary dysplasia | 41% | 58% | 0.028 |
| Steroids for BPD | -- | 27% | |
| *Gestational age 28 - 31 wks* | *N=608* | *N=505* | |
| Bronchopulmonary dysplasia | 9% | 11% | 0.424 |
| Steroids for BPD | -- | 1% | |

for gestational age, associated with better outcome in 1993, but not in 1983. Necrotising enterocolitis was, when corrected for other perinatal factors, in 1993 associated with adverse outcome.

Only the change of the influence of gender was statistically significant. This means that the relative disadvantage of male gender in 1983 has disappeared in 1993.

**Table 5:** Total number and adverse outcome (in-hospital mortality, intraventricular haemorrhage and/or bronchopulmonary dysplasia) in two cohorts (1983 and 1993) of live-born, very preterm infants with gestational age < 32 weeks, stratified by gestational age categories (< 28 and 28 - 31 weeks). In 1983 restricted to the infants treated in one of the six neonatal intensive care units that performed ultrasound of the brain routinely and systematically.

| | 1983 | | | 1993 | | | p |
|---|---|---|---|---|---|---|---|
| | Total | Adverse outcome | | Total | Adverse outcome | | |
| | N | N | % | N | N | % | |
| All gestational ages < 32 wks | 488 | 277 | 56.8% | 724 | 340 | 47.0% | 0.001 |
| Gestational age < 28 wks | 127 | 105 | 82.7% | 150 | 127 | 84.7% | 0.744 |
| Gestational age 28 - 31 wks | 361 | 172 | 47.6% | 574 | 213 | 37.1% | 0.002 |

**Table 6:** Multivariate logistic regression of antenatal, perinatal and postnatal risk factors on adverse outcome (in-hospital mortality, intraventricular haemorrhage and/or bronchopulmonary dysplasia) in infants < 32 weeks' gestational age, for the 1983 and the 1993 population separately. Antenatal risk factors adjusted for each other; perinatal risk factors adjusted for each other and the antenatal ones; postnatal risk factors adjusted for all others. The outcome parameters death, intraventricular haemorrhage and bronchopulmonary dysplasia are not used as risk factors in the modelling procedure for adverse outcome.
Non-significant factors are retained in the model to allow comparison of the coefficient.
In 1983 restricted to the infants treated in one of the six neonatal intensive care units that performed ultrasound of the brain routinely and systematically.

| | 1983 | | | 1993 | | |
|---|---|---|---|---|---|---|
| | **OR** | **95% C.I.** | **p** | **OR** | **95% C.I.** | **p** |
| Antenatal | | | | | | |
| Male gender* | 1.83 | 1.21 - 2.75 | < 0.01 | 1.52 | 0.82 - 1.61 | 0.41 |
| Congenital malformations | 2.67 | 1.32 - 5.42 | < 0.01 | 1.75 | 1.05 - 2.94 | 0.03 |
| Gestational age (wks) | 0.59 | 0.51 - 0.68 | < 0.01 | 0.60 | 0.53 - 0.68 | < 0.01 |
| Birth weight (100 g) | 0.97 | 0.91 - 1.04 | 0.43 | 0.92 | 0.87 - 0.97 | < 0.01 |
| Perinatal | | | | | | |
| Ante partum steroids | 0.56 | 0.30 - 1.01 | 0.08 | 0.68 | 0.47 - 0.99 | 0.04 |
| Inborn | 0.47 | 0.30 - 0.73 | < 0.01 | 0.78 | 0.50 - 1.20 | 0.78 |
| Low Apgar score | 2.35 | 1.37 - 4.01 | < 0.01 | 1.96 | 1.27 - 3.02 | < 0.01 |
| Postnatal | | | | | | |
| Artificial ventilation (group) | 3.11 | 2.20 - 4.40 | < 0.01 | 2.29 | 1.75 - 2.98 | < 0.01 |
| Necrotising enterocolitis | 1.17 | 0.42 - 3.23 | 0.76 | 5.45 | 1.61 -18.39 | < 0.01 |

** Only the magnitude of the odds ratios of male gender on adverse outcome changed significantly between 1983 and 1993.*

## Discussion

The 1983 Project On Preterm and Small for gestational age population was a virtually complete population of all live-born infants in the Netherlands with gestational ages below 32 weeks and/or birth weights below 1500 grams. When corrected for the variability in perinatal risk factors, neonatal mortality was similar in the individual Dutch neonatal intensive care centres in 1983 [19,20]. We selected from this population the infants with gestational ages below 32 weeks to compare them with the 1993 population consisting of all infants from three health care regions that cover approximately 30% of the Netherlands. Therefore we were able to compare similar geographically defined cohorts ten years apart.

The years between 1983 and 1993 showed major demographic changes in the Netherlands. Mean maternal age increased 1.9 years, as did the proportion of non-Dutch women [10,11]. Both factors may have increased the obstetrical risk and the chance of preterm birth [12,21-24]. In fact a national study showed a proportional increase in live-births with a gestational age of less than 32 weeks or a birth weight of less than 1500 grams from 0.84% in 1983 to 1.13% in 1995 [22]. In our study-group the increase in mean maternal age was 3.2 years; mean maternal age was in 1983 lower in the mothers of very preterm infants than in the general population, while this difference had disappeared in 1993. Although in the literature higher maternal age is related to an increase in adverse outcome, we could not establish a relation between perinatal death or adverse outcome and maternal age or ethnicity in our study population of very preterm infants.

Between 1983 and 1993, almost complete regionalisation and centralisation of perinatal care in the Netherlands took place. A national study showed that in 1993 only 11% of all infants with gestational ages < 32 weeks or birth weight below 1500 grams remained in their local hospital. All others were referred to a tertiary perinatal centre, either before or after birth [8]. The not-referred infants had more advanced gestational ages and higher birth weights. In our 1993 study population, we documented from data of The Netherlands Perinatal Register that only 3% of the infants < 28 weeks' gestation and 6% of the infants of 28 - 32 weeks' gestation were not referred to a neonatal intensive care unit. Official guidelines from the Dutch Society for Obstetrics and Gynaecology and the Dutch Paediatric Association advocate antenatal referral between 24 and 32 weeks' of gestation. An Internet site with available intensive care beds in each of the ten Dutch tertiary perinatal centres, short distances and a professional service for IC-transportation of neonates enable optimal use of the restricted capacity of the Dutch intensive care beds.

The policy on treatment of extremely preterm infants ('micropremies') in the

Netherlands is rather restricted [25]. Both in 1983 and in 1993, no infants with gestational ages < 25 weeks' gestation survived the neonatal period. This reflects the attitude of the majority of Dutch neonatologists to withhold resuscitation in infants below 25 weeks' gestation because of the limited chances of intact survival [25-28]. In such cases, compassionate care in a conjoined effort of parents and doctors is given in the delivery room and not in the neonatal intensive care unit. In the Leiden Follow up Project on Prematurity, the survival of infants of 24 weeks' gestation was 40% in 1996-97, due to a more aggressive treatment than in other parts of the Netherlands [29,30]. Because of the high incidence of adverse outcome the authors suggested that < 25 weeks' gestation 'one should carefully weigh whether or not to aggressively resuscitate and treat these extremely preterm infants' [29].

Obstetrical care in pregnancies of more than 25 weeks' gestation has become more active to aim at the delivery of an infant in optimal condition. Referral before birth and the use of ante partum steroids are offered to all women who are in very preterm labour, unless labour progresses very rapidly or an emergency Caesarean section has to be performed because of foetal distress [8]. Caesarean section has become more generally accepted even at these low gestational ages. These trends were confirmed in our study. Though the protective effect of Caesarean section, steroids and antenatal transport did not change over time, the fact that they were more readily available may have been an important factor in the better survival.

Richardson compared the influence of scores for neonatal acute physiology (SNAP) and perinatal risk factors on mortality in two populations, born in 1989-90 and 1994-95 respectively. He suggested that one third of the improvement in mortality could be attributed to a better condition on admission and two thirds to a more effective neonatal intensive care [1]. In our study the changes in obstetrical care were only partly reflected by the number of infants with low five minutes Apgar scores and not at all by the proportion of infants that needed artificial ventilation. Although the Apgar score is an imperfect parameter of obstetric action, it reflects the condition of the child immediately after birth and the lack of substantial improvement in the infants, especially in infants above 28 weeks' gestation, is worrisome.

Mortality decreased by 40% in infants of all gestational ages and total adverse outcome improved in the infants of 28 - 32 weeks' gestation.

Intraventricular and periventricular haemorrhages, as detected by ultrasound, are strongly correlated with poor long term outcome, although this association is stronger for periventricular haemorrhage and leukomalacia than for intraventricular haemorrhage [31]. We expected an increased incidence of lower grades of intraventricular haemorrhage due to technically improved ultrasound equipment in the 1990s, but the slight increase was not statically significant. The unaltered incidence of intraventricular

haemorrhage as percentage of the survivors may be reassuring at the first sight, but the decreased mortality achieved by modern perinatal intensive care will be accompanied by an increased incidence of disabled children related to the number of live-born very preterm infants due to cerebral haemorrhage as younger and more vulnerable babies survive.

Between 1983 and 1993 the use of antenatal steroids became standard therapy and artificial surfactant became available for infants with respiratory distress. Meta-analysis of randomised clinical trials has showed a 30% reduction in mortality by the use of surfactant [9]. It was hypothesised that also the incidence of respiratory distress, the frequency and mean duration of artificial ventilation and the incidence of bronchopulmonary dysplasia would decrease, but these main indicators of pulmonary disease did not improve in our study. The Leiden Follow up Project on Prematurity reported, in the Leiden area, an unchanged incidence of RDS, an increased incidence and duration of artificial ventilation and an increased incidence of bronchopulmonary dysplasia between 1983 and 1996 - 1997 [30]. However, in the 1983 population, 24.2% of the ventilated patients died during the first week of life versus 13.0% in 1993 in our population. The use of surfactant may have played a role in decreasing the mortality in the first days of life, but it did not decrease the mean duration of artificial ventilation in the surviving infants.

Long term ventilation and chronic neonatal lung disease have been reported as risk factors for long term adverse outcome before [32] and after the introduction of surfactant [33-35]. Although not all studies addressing developmental outcome for infants with bronchopulmonary dysplasia are conclusive, many report poorer developmental outcome and increased frequencies of neurological problems, particularly cerebral palsy. In 910 children with birth weights between 500 to 999 grams, Schmidt et al. found an odds ratio for adverse outcome, defined as the presence of cerebral palsy, cognitive delay, severe hearing loss or blindness and corrected for perinatal confounders, of 2.4 for bronchopulmonary dysplasia (defined as oxygen dependency at 36 weeks postmenstrual age). This figure was in the same order of magnitude as the odds ratio of 3.7 for brain injury and 3.1 for severe ROP [33]. In a cohort of 217 children born at 24 - 28 weeks' gestation, Grégoire et al. found neurodevelopmental delay in those children who were oxygen dependent at 36 weeks postmenstrual age, compared to controls (mean developmental quotient 90.7 (S.D. 19.3) versus 97.4 (S.D. 15.0)), but not in those who were oxygen dependent at 28 days postnatal age [34]. In a population of 122 very low birth weight infants and 123 controls at three years of age, Singer et al. calculated the effect of bronchopulmonary dysplasia as a change of -12 in the standard score of the Psychomotor Index of the Bayley Scales of Development and confirmed their finding at eight years of age [35,36]. They defined bronchopulmonary dysplasia as

oxygen dependence at 28 days postnatal age. They found no difference in MDI and PDI scores when they compared these children with children based on the 36-weeks postmenstrual age definition within their sample. The Leiden Follow up Project on Prematurity documented in preterm infants < 32 weeks' gestation at two years of age a MDI score of 97.3 (S.D. 24.8) and a PDI score of 95.8 (S.D. 21.7). In multivariate analysis, birth weight and postnatal treatment with steroids were the only independent predictors of delayed psychomotor development; ethnicity, maternal age at birth and male gender were predictive of delayed mental development as well. The effects of postnatal steroids on development after correction for these significant confounders were a 9.3 (S.E. 6.6) point lower MDI score and 12.7 (S.E. 6.1) points lower PDI score [37].

To compare the data of the 1983 cohort with the 1993, we could only use the classical 28-days postnatal age definition of Bancalari as marker for adverse outcome. The incidence of bronchopulmonary dysplasia in infants of 28 - 31 weeks' gestation that were discharged alive from hospital increased from 41% in 1983 to 58% in 1993. Half of them received prolonged postnatal administration of steroids, that were not used this way in 1983.

Gestational age, ante partum steroids, condition at birth and artificial ventilation were the main determinants of adverse outcome both in 1983 and 1993. The literature shows that the incidence of preterm birth among white infants has increased during the last decades in spite of efforts to reduce the incidence of preterm birth, presumably because of the increased maternal age and assisted reproduction [21,38,39]. The relative influence of these factors did not change over time, although the incidences did.

The disadvantage of male gender disappeared [40]. This finding has not been reported before [41]. We speculate that the effect of the male disadvantage has been decreased by more intensive treatment and that nurture has beaten nature [42].

## Conclusion

During one decade, the mortality of very preterm infants in the Netherlands decreased by 40%. Even when corrected for known perinatal risk factors, the odds on dying in 1993 were half of those in 1983. Socioeconomic circumstances in this decade led to a higher maternal age and thereby to an increase in multiple births and prematurity.

We hypothesise that the whole package of better obstetric and neonatal care, including prenatal transfer, earlier Caesarean section, systematic administration of steroids and specially the administration of surfactant, has accomplished this increased survival.

However, the better survival was not accompanied by a decreased incidence of

intraventricular haemorrhage, as percentage of the survivors, and the incidence of bronchopulmonary dysplasia in the infants born after < 28 weeks'gestation even increased. We therefore do not share the optimistic view that better perinatal care will be accompanied by better long term outcome in surviving children, but fear that it will be accompanied by an increasing incidence of disabled children because ever younger and more vulnerable babies survive.

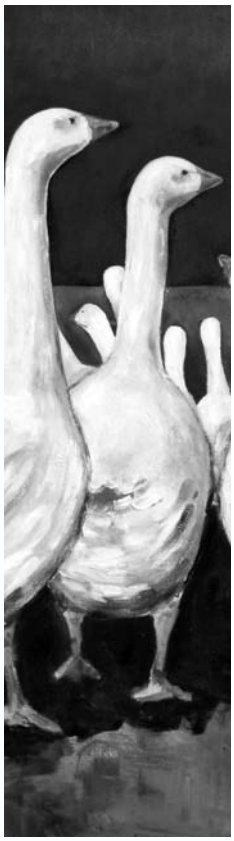
## References


1. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care [see comments]. Pediatrics 1998;102:893-9.
2. International Neonatal Network Scottish Neonatal Consultants Nurses Collaborative Study Group. Risk adjusted and population based studies of the outcome for high risk infants in Scotland and Australia. Arch Dis Child Fetal Neonatal Ed 2000;82:F118-23.
3. The Victorian Infant Collaborative Study Group. Economic outcome for intensive care of infants of birthweight 500-999 g born in Victoria in the post surfactant era. J Paediatr Child Health 1997;33:202-8.
4. Battin M, Ling EW, Whitfield MF, Mackinnon M, Effer SB. Has the outcome for extremely low gestational age (ELGA) infants improved following recent advances in neonatal intensive care? Am J Perinatol 1998;15:469-7.
5. The Victorian Infant Collaborative Study Group. Improved outcome into the 1990s for infants weighing 500-999 g at birth. The Victorian Infant Collaborative Study Group. Arch Dis Child Fetal Neonatal Ed 1997;77:F91-4.
6. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986;1:55-7.
7. Health Council of the Netherlands. Intensive care at birth. 2000. The Hague, Health Council of the Netherlands; publication no. 2000/08.
8. Kollée LA, den Ouden AL, Drewes JG, Brouwers HA, Verwey RA, S.P. [Increase in perinatal referral to regional centers of premature birth in The Netherlands: comparison 1983 and 1993]. Ned Tijdschr Geneeskd 1998;142:131-4.
9. Soll RF. Synthetic surfactant for respiratory distress syndrome in preterm infants (Cochrane Review). In: The Cochrane Library, 1, 2002.Oxford: Update Software. 2002. Oxford, The Cochrane Library.
10. Centraal Bureau voor de Statistiek. Maandbericht gezondheid 85/7. The Hague:, Sdu /Centraal Bureau voor de Statistiek 2002.
11. Centraal Bureau voor de Statistiek. Statistisch Jaarboek 1995. The Hague: Sdu/Centraal Bureau voor de Statistiek. 2002:36.
12. Seoud MA, Nassar AH, Usta IM, Melhem Z, Kazma A, Khalil AM. Impact of advanced maternal age on pregnancy outcome. Am J Perinatol 2002;19:1-8.
13. Anthony S, Buitendijk SE, Dorrepaal CA, Lindner K, Braat DD, den Ouden AL. Congenital malformations in 4224 children conceived after IVF. Hum Reprod 2002;17:2089-95.
14. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birhtweight infants at five years of age. The collaboration project on preterm and small for gestational age infants (POPS) in the Netherlands. Lancet 1991;338:33-6.
15. Kloosterman GL, Huidekoper BL. Over intra-uteriene groei en de intra-uteriene groeicurve. Maandschr voor Kindergeneesk 1969;37:209-25.
16. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr. 1978;92:529-34.

17. van de Bor M, Ens-Dokkum M, Schreuder AM, Veen S, Brand R, Verloove-Vanhorick SP. Outcome of periventricular-intraventricular haemorrhage at five years of age. Dev Med Child Neurol 1993;35:33-41.
18. Bancalari E, Abdenour GE, Feller R, Gannon J. Bronchopulmonay dysplasia: clinical presentation. J Pediatr 1979;95:819-23.
19. Brand R, van Hemel OJ, Elferink-Stinkens PM, Verloove-Vanhorick SP. Comparing mortality and morbidity in hospitals: theory and practice of quality assessment in peer review. Methods Inf Med 1994;33:196-204.
20. Verloove-Vanhorick SP, Brand R. [Inter-hospital comparison of perinatal intensive care in academic hospitals]. Ned Tijdschr Geneeskd 1989;133:1883-6.
21. Abel EL, Kruger M, Burd L. Effects of maternal and paternal age on Caucasian and Native American preterm births and birth weights. Am J Perinatol 2002;19:49-54.
22. Anthony S, den Ouden L, Brand R, Verloove-Vanhorick SP, Bennebroek Gravenhorst J. Changes in perinatal care and survival of preterm and extremely preterm infants in the Netherlands between 1983 and 1995. Eur J Obstet Gynecol Reprod Biol 2003.
23. Buitendijk SE. Children after in vitro fertilization. An overview of the literature. Int J Technol Assess Health Care 1999;15:52-65.
24. Evers JL. Female subfertility. Lancet 2002;360:151-9.
25. Sheldon T. Dutch doctors change policy on treating preterm babies. BMJ 2001;322:1383.
26. De Leeuw R, Cuttini M, Nadai M, Berbik I, Hansen G, Kucinskas A et al. Treatment choices for extremely preterm infants: an international perspective. J Pediatr 2000;137:608-16.
27. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. N Engl J Med 2000;343:378-84.
28. Tin W, Wariyar U, Hey E. Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. BMJ 1997;314:107-11.
29. Rijken M, Stoelhorst GM, Martens SE, van Zwieten PH, Brand R, Wit JM, Veen S. Mortality and neurologic, mental, and psychomotor development at 2 years in infants born less than 27 weeks' gestation: the Leiden follow-up project on prematurity. Pediatrics 2003;112:351-8.
30. Stoelhorst GMSJ. Development, Quality of Life and Behavior at 2 years of age in very preterm infants. The Leiden Follow-Up Project on Prematurity 1996-1997. Thesis. Leiden University 2003.
31. Vermeulen GM, Bruinse HW, de Vries LS. Perinatal risk factors for adverse neurodevelopmental outcome after spontaneous preterm birth. Eur J Obstet Gynecol Reprod Biol 2001;99:207-12.
32. Saigal S, O'Brodovich H. Long-term outcome of preterm infants with respiratory disease. Clin Perinatol 1987;14:635-50.
33 Schmidt B, Asztalos EV, Roberts RS, Robertson CMT, Sauve RS, Whitfield MF. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months. JAMA 2003;289:1124-9.
34. Grégoire M-C, Lefebvre F, Glorieux J. Health and developmental outcome at 18 months in very preterm infants with bronchopulmonary dysplasia. Pediatrics 1988;1010:856-60.
35. Singer L, Yamashita T, Lilien L, Collin M, Baley J. A longitudinal study of developmental outcome of infants with bronchopulmonary dysplasia and very low birth weight. Pediatrics 1997;1000:987-93.
36. Short EJ, Klein NK, Lewis BA, Fulton S, Eisengart S, Kercsmar C, Baley J, Singer LT. Cognitive and academic consequences of bronchopulmonary dysplasia and very low birth weight: 8-year old outcomes. Pediatrics 2003;112:e359-66.
37. Stoelhorst GM, Rijken M, Martens SE, van Zwieten PH, Feenstra J, Zwinderman AH, Wit JM, Veen S; Leiden Follow-Up Project on Prematurity. Developmental outcome at 18 and 24 months of age in very preterm children: a cohort study from 1996 to 1997. Early Hum Dev 2003;72:83-95.

38. Branum AM, Schoendorf KC. Changing patterns of low birthweight and preterm birth in the United States, 1981-98. Paediatr Perinat Epidemiol 2002;16:8-15.
39 Alexander GR, Slay M. Prematurity at birth: trends, racial disparities, and epidemiology. Ment Retard Dev Disabil Res Rev 2002;8:215-20.
40. Naeye RL, Burt LS, Wright DL, Blanc WA, Tatter D. Neonatal mortality, the male disadvantage. Pediatrics 1971;48:902-6
41. Stevenson DK, Verter J, Fanaroff AA, Oh W, Ehrenkranz RA, Shankaran S, Donovan EF, Wright LL, Lemons JA, Tyson JE, Korones SB, Bauer CR, Stoll BJ, Papile LA. Sex differences in outcomes of very low birthweight infants: the newborn male disadvantage Arch Dis Child. Fetal Neonatal Ed 2000;83: F182-5.
42. Eisenberg L. Experience, brain, and behavior: The importance of a head start. Pediatrics 1999;103:1031-5.

Chapter 7



# Outcome of perinatal care for very preterm infants at five years of age; a comparison between 1983 and 1993

M.J.K. de Kleine, A.L. den Ouden, L.A.A. Kollée, A.L. van Baar,
M.W.G. Nijhuis-van der Sanden, A. Ilsen, R. Brand, S.P. Verloove-Vanhorick.

Submitted.

## Summary

Perinatal mortality in very preterm infants has decreased up to fifty percent during the last decades. Studies of changes of long term outcome are inconclusive. Therefore we studied the visual, auditory, neuromotor, cognitive and behavioural development of two geographically defined populations of very preterm infants, born in 1983 (N=1007) and in 1993 (N=724), and analysed the relationship between perinatal risk factors and outcome.

The incidence of disabling cerebral palsy increased from 6.3% to 12.3% in surviving children. Impaired vision and strabismus decreased significantly, presumable by continuous monitoring of $pO_2$. Hearing problems, the need for special education and the incidence of behavioural problems did not change over time.

The frequency of children that performed optimally in every domain increased from 30.6% in 1983 to 41.1% in 1993.

Cerebral palsy was associated with male gender in 1983 and with low Apgar score, intraventricular haemorrhage and postnatal steroids for bronchopulmonary dysplasia in 1993.

The intensiveness of neonatal treatment has increased, leading to the survival of more children but at the cost of more children with cerebral damage. Modern perinatal care is not longer limited by the devastating effects of pulmonary problems as it was in the past, but fails in safeguarding cerebral integrity in very preterm infants.

## Introduction

Perinatal mortality in very preterm and very low birth weight infants has decreased up to 50% during the last decades. A major part of this improvement is due to the use of steroids prior to preterm birth and the introduction of artificial surfactant in the early 1990s [1-3].

These treatments may also prevent cerebral damage causing developmental disabilities. However, cerebral palsy rates have been reported to increase as well as to decrease in very preterm infants [4-7]. Changes in other developmental disability rates have hardly been reported and are often inconclusive [8-10].

We used the opportunity presented by data from two geographically defined populations of very preterm infants, born in 1983 and in 1993, to compare long term visual, auditory, neuromotor, cognitive and behavioural developmental outcomes over time and to analyse possible changes in the relationship between perinatal risk factors and such outcomes.

## Methods

*Study populations*

The 1983 cohort included 94% of all very preterm (< 32 weeks' gestation) and/or very low birth weight (< 1500 grams) infants born in the Netherlands in that year in a nationwide follow up study, the Project On Preterm and Small for gestational age infants (POPS) [11-13].

The 1993 cohort included very preterm and very low birth weight infants born between October 1992 and December 1994 in the regions of three perinatal intensive care centres covering approximately one third of the Dutch neonatal intensive care service. For logistic reasons, the inclusion periods varied slightly between the regions: the Radboud University Nijmegen Medical Centre, Nijmegen 20 months, the Academic Medical Centre, Amsterdam 14 months and the Máxima Medical Centre, Veldhoven 24 months. Data of 30 very preterm infants born in these regions but not referred to a tertiary perinatal centre were retrieved from the Netherlands Perinatal Register [14].

From both cohorts we selected live-born infants with a gestational age of less than 32 weeks, 1007 in 1983 and 724 in 1993.

*Perinatal data*

Perinatal data were collected during admission. Antenatal factors included maternal age, gender, congenital malformations (structural and non-structural as well as minor and major), multiple birth, gestational age and birth weight. Perinatal factors were administration of antenatal steroids, mode of delivery (vaginally versus Caesarean section), Apgar scores at 5 minutes, place of birth (tertiary centre or local hospital) and transport after birth. Postnatal factors were respiratory problems (need of artificial ventilation, use of surfactant or the diagnosis of bronchopulmonary dysplasia), cerebral problems (intraventricular haemorrhage (grade 1 to 4) or seizures) and infections (sepsis or necrotising enterocolitis). We used the grading system by Papile et al. for intraventricular haemorrhage: grade I, subependymal haemorrhage; grade II, intraventricular haemorrhage; grade III, intraventricular haemorrhage with ventricular dilatation; and grade IV, parenchymal haemorrhage [15]. Routine and systematic ultrasound of the brain was performed in six out of eight neonatal intensive care units in 1983 and in all centres in 1993 [16]. Ultrasound was performed soon after admission; examinations were performed at least twice in the first week of life and once weekly thereafter. In 1983, the incidence of intraventricular haemorrhages in those six centres that performed ultrasound routinely and systematically was 37% versus 29% in the other two centres ($p = 0.042$). Therefore, analyses of intraventricular haemorrhages and their effect on outcome were restricted to those six neonatal intensive care units. Bronchopulmonary dysplasia in both cohorts was defined according to the definition of Bancalari (supplemental oxygen for more than 28 days), as this was the only available definition of bronchopulmonary dysplasia in 1983 [17].

*Data collection at five years of age*

In the 1983 cohort, three paediatricians assessed the overall outcome at age five during a home visit in 648 out of 678 survivors (96%) [12].

In the 1993 cohort, overall outcome was available for 511 out of 592 survivors (86%). The majority (N=406; 79%) was assessed in the tertiary centre following a protocol similar to the 1983 cohort. The outcome of 85 children (17%) was based on a routine assessment by their attending paediatrician in the local hospital. Another 20 children (4%) did not attend the follow up assessment because of previously established handicaps.

To estimate the possible influence of the higher non-response rate in 1993 on outcome perinatal factors were compared between responders and non-responders in both cohorts.

Vision was assessed by physical examination of the movements of the eyes and the cover test and by history taken from the parents about vision with and without glasses, strabismus and ophthalmologic treatment. Normal vision without glasses was assigned 'optimal'; all other conditions were assigned non-optimal. Strabismus was categorised in yes or no.

Permanent hearing problems were assessed by history, taken from the parents, about hearing with and without hearing aids.

Neurological function was assessed according to Touwen. This assessment consists of 40 items that cover tone and posture, muscle power, reflexes, coordination and balance, and involuntary movements [18]. The children were assigned to three categories: normal (no neurological abnormalities), neurological dysfunction (neurological abnormalities without influence on normal posture or movement or non-disabling cerebral palsy) or disabling cerebral palsy (neurological abnormalities with impairments in sitting, standing, walking or the use of hands).

School performance was assessed by history, taken from the parents, and categorised in mainstream or special education.

Problem behaviour was assessed by a shortlist of the 15 problem items from the Child Behavior CheckList (CBCL) that discriminated most between normal children and children referred to mental health services in the general population and categorised as optimal or non-optimal [19].

*Statistical analyses*

Univariate analyses of differences in perinatal risk factors included the Student *t* test, Fisher's exact test or chi square test; p values $< 0.05$ were considered statistically significant.

Multivariate analyses included a logistic regression model using backwards stepwise procedures in which risk factors from an earlier chronological period were retained and risk factors from the same chronological period were removed if they were not significant. This procedure was first performed separately for both populations to obtain the OR's for the risk factors in each of them. The risk factors found were subsequently introduced in a model including an interaction term for year of birth (1983 or 1993) to calculate possible changes in the effect of each of the significant perinatal factors over time.

## Results

*Perinatal data*

In the source populations of 1007 (1983 cohort) and 724 (1993 cohort) live-born infants mean gestational age was 29.1 and 29.4 weeks and mean birth weight 1247 and 1238 grams respectively. Neonatal mortality decreased from 28.0% to 15.3% and first year mortality from 32.3% to 17.8%. Between the age of one and five years another 0.4% and 0.5% of the children died.

Perinatal data of the 678 (1983 cohort) and 592 (1993 cohort) survivors at five years of age are presented in table 1. The higher frequency of multiple birth, antenatal referral and Caesarean section reflect the changes in obstetrical practice. Despite the increase of antenatal steroid administration and the introduction of surfactant the need for artificial ventilation increased slightly and the duration did not decrease. The incidence of pneumothorax decreased, that of bronchopulmonary dysplasia increased. The frequency of intraventricular haemorrhage was similar in both cohorts, although the distribution in severity differed (Pearson chi square < 0.001).

*Assessment at five years of age*

The incidence of disabling cerebral palsy almost doubled from 6.3% to 12.3% of the assessed children (table 2). Impaired vision and strabismus decreased significantly. Hearing problems, the need for special education and the incidence of behavioural problems did not change over time. The frequency of children that performed optimally in every domain increased from 30.6% in 1983 to 41.1% in 1993 (table 2).

In the multiple logistic regression analyses, cerebral palsy was associated with male gender and seizures in 1983 and with low Apgar score, intraventricular haemorrhage and postnatal steroids for bronchopulmonary dysplasia in 1993 (table 3). The influence of male gender changed significantly over time, implying that the male disadvantage found in 1983 had disappeared in 1993.

Perinatal characteristics of the responders were different in 1983 and 1993. In 1983 they had received significantly more often antenatal steroids (20.9% versus 3.3%; $p = 0.017$). The responders in 1993 were less mature and smaller than the non-responders (29.6 versus 30.3 weeks; $p < 0.01$ and 1270 versus 1420 grams; $p < 0.001$). They had also received more often antenatal steroids (47.8% versus 26.6%; $p < 0.01$), had more intraventricular haemorrhages (27.2% versus 8.6%; $p < 0.001$) and more artificial ventilation (54.4% versus 35.8%; $p < 0.01$).

**Table 1:** Perinatal data in two cohorts (1983 and 1993) of live-born, very preterm infants, born after < 32 weeks gestational age and discharged home alive.

| | **1983 N=678** | **1993 N=592** | **p** |
|---|---|---|---|
| Gestational age (wks; mean ± S.D.) | 29.8 ± 1.5 | 29.7 ± 1.6 | 0.297 |
| Birth weight (g; mean ± S.D.) | 1330 ± 315 | 1291 ± 336 | 0.032 |
| Percentage small for gestational age (< P10) | 16.2% | 19.9% | 0.092 |
| Male gender | 53.8% | 54.4% | 0.866 |
| Congenital malformations | 7.8% | 9.6% | 0.272 |
| Multiple birth | 23.0% | 33.8% | < 0.001 |
| Antenatal steroid administration | 20.1% | 45.0% | < 0.001 |
| Caesarean section | 35.7% | 42.4% | 0.015 |
| Apgarscore 5′< 7 | 13.8% | 15.5% | 0.422 |
| Transport | | | |
| Inborn | 42.0% | 76.0% | < 0.001 |
| Transported postpartum to NICU | 32.4% | 19.1% | < 0.001 |
| No NICU admission | 25.6% | 4.9% | < 0.001 |
| Artificial ventilation | 44.7% | 51.8% | 0.011 |
| Days on the ventilator (mean ± S.D.) | 8.5 ± 9.6 | 8.2 ± 8.3 | 0.716 |
| Surfactant (ventilated infants only) | - | 39.3% | - |
| Pneumothorax | 7.8% | 3.0% | < 0.001 |
| Bronchopulmonary Dysplasia (BPD) | 12.8% | 17.7% | 0.018 |
| Steroids for BPD | - | 4.9% | |
| Intraventricular haemorrhage* | | | |
| No haemorrhage | 230 (78.6%) | 446 (75.3%) | 0.158# ‡ |
| Grade 1 | 40 (12.3%) | 96 (16.2%) | 0.320† |
| Grade 2 | 29 (9.0%) | 16 (2.7%) | < 0.001† |
| Grade 3 | 19 (5.9%) | 30 (5.1%) | 0.534† |
| Grade 4 | 6 (1.9%) | 4 (0.7%) | 0.101† |
| Seizures | 3.7% | 2.0% | 0.094 |
| Sepsis | 10.0% | 17.2% | < 0.001 |
| NEC | 5.6% | 2.9% | 0.019 |

* *Data on intraventricular haermorrhages in 1983 are restricted to infants from those neonatal intensive care units that performed ultrasound of the brain routinely and systematically (N=324)*

\# *No haemorrhage versus all haemorrhages*

† *Particular grade versus no haemorrhage*

‡ *Difference in distribution p < 0.001*

**Table 2:** Outcome at five years of age in two cohorts (1983 and 1993) of live-born, very preterm infants, born after < 32 weeks' gestation; univariate analysis.

| | 1983 | | | 1993 | | | p |
|---|---|---|---|---|---|---|---|
| | Total N | Outcome N and % | | Total N | Outcome N and % | | |
| Survivors | 678 | | | 592 | | | |
| number assessed | 648 | | 95.6% | 511 | | 86.3% | |
| *Neuromotor development (Touwen)* | 648 | | | 491 | | | |
| Normal | | 419 | 64.7% | | 322 | 65.4% | |
| Gross motor problems, MND or non-disabling cerebral palsy | | 188 | 29.0% | | 109 | 22.3% | |
| Disabling cerebral palsy | | 41 | 6.3% | | 60 | 12.3% | < 0.001 |
| *Visual domain* | | | | | | | |
| Strabismus | 648 | 213 | 32.9% | 359 | 45 | 12.5% | < 0.001 |
| Impaired vision | 648 | 105 | 16.2% | 506 | 60 | 11.9% | 0.042 |
| *Auditory domain* | | | | | | | |
| Hearing problems | 569 | 52 | 9.1% | 417 | 37 | 8.9% | 0.911 |
| *School* | | | | | | | |
| Special school | 646 | 77 | 11.9% | 438 | 43 | 9.8% | 0.255 |
| *Behaviour* | | | | | | | |
| Non-optimal | 646 | 184 | 28.5% | 365 | 117 | 32.1% | 0.252 |
| *All domains optimal* | *598* | *180* | *30.6%* | *379* | *157* | *41.4%* | *0.001* |

**Table 3:** Multivariate logistic regression model of statistical significant influences of antenatal, perinatal and postnatal risk factors on disabling cerebral palsy of two cohorts (1983 and 1993) of very preterm infants, born after < 32 weeks' gestation. Non-significant factors are retained in the model to allow comparison of the coefficient.

| | 1983 | | | 1993 | | |
|---|---|---|---|---|---|---|
| | **OR** | **95% C.I.** | **p** | **OR** | **95% C.I.** | **p** |
| *Antenatal* | | | | | | |
| Male gender | 5.94 | [1.68; 21.0] | 0.006 | 1.47 | [0.84; 2.57] | 0.174 |
| *Perinatal* | | | | | | |
| Apgar score 5' < 7 | 1.75 | [0.58; 5.31] | 0.320 | 2.60 | [1.42; 4.77] | 0.002 |
| *Postnatal* | | | | | | |
| Intraventricular haemorrhage | 1.29 | [0.81; 2.06] | 0.291 | 1.61 | [1.23; 2.12] | 0.001 |
| Seizures | 12.27 | [3.33; 45.19] | < 0.001 | 1.69 | [0.36; 7.81] | 0.516 |
| Steroids for BPD | - | - | - | 3.93 | [1.52; 10.15] | 0.005 |

** The change of OR's of male versus female gender on cerebral palsy is statistically significant.*

## Discussion

In these two comparable cohorts of very preterm infants, born ten years apart, we found a significant increase in disabling cerebral palsy at age five years and a decrease in visual problems and strabismus. Hearing, school and behaviour outcomes were similar. The percentage of children without any problem increased. The administration of postnatal steroids had an adverse effect on the incidence of disabling cerebral palsy.

Although it concerned two geographically defined cohorts, the comparability may be questioned. First, the 1983 cohort was nationwide, while the 1993 cohort concerned approximately one third of the Netherlands. However, since regional differences in neonatal health care in the Netherlands are small, selection bias is unlikely [20,21]. Second, different paediatricians assessed the children both in 1983 and 1993. However, these paediatricians were trained in the same way. Moreover, Paneth et al. showed that the inter-observer agreement of the assessment of disabling cerebral palsy is excellent, especially when a rigorously standardised test, like the Touwen examination, is used [22]. Third, the attrition rate in 1993 was much larger than in 1983. Twenty infants did not

attend the follow up assessment, because of previously assessed severe handicaps. The analysis of perinatal characteristics of responders and non-responders showed a negative selection bias in 1993. Since the attrition rate in 1983 was very small, bias in this cohort was negligible. In 1993 on the other hand, bias might be considerable. Therefore, we extrapolated the incidences of developmental outcomes in survivors in 1993 assuming that non-responders had normal outcome. Even in such a most favourable outcome scenario, the cerebral palsy rate (10.0%) was significantly higher than in the 1983 cohort ($p < 0.01$).

The aetiology of cerebral palsy is complicated and many factors may be involved, such as preconceptional factors, thyroid hormone and growth factor deficiencies, perinatal infections with hyperthermia and production of cytokines, hypoxic-ischemic episodes and the use of steroids [23-26]. Data from the present study do not allow conclusions in this respect.

A possible explanation is that the intensiveness of neonatal treatment has increased, leading to the survival of more children but at the cost of more children with cerebral damage. This is previously suggested in a comparison of the extremely preterm participants of the 1983 cohort with a contemporary cohort from New Jersey [27]. This would also explain the fact that further analysis of our data showed that the cerebral palsy rate in survivors less than 28 weeks increased relatively more than the cerebral palsy rate in 28 weeks and over. This has been predicted by Paneth 28 and is in accordance with the findings in some other studies [4-6,29].

The decrease of visual problems and strabismus may be a result of the same increased intensiveness of care and the introduction of continuous monitoring of transcutaneous $pO_2$ and oxygen saturation between the 1980s and the 1990s.

These changes in perinatal management strategies, however, did not influence outcome in hearing, behaviour and the need for special education.

The percentage of children in special education (11.9% and 9.8%) was substantially higher than the 1.3% of five year old children in the general population. In the general population the participation in special education remained stable for all age groups born between 1983 and 1993 [30]. In the Netherlands, children at age five are in a transition between Kindergarten and elementary school. Hille et al. documented previously that school problems in the 1983 cohort are considerably underestimated at age five years and that problem behaviour is a risk factor for later school failure [31]. The incidence of problem behaviour in 1993 is still twice that of a control population. This problem behaviour will probably hamper future academic achievements in the 1993 cohort as much as is did in the 1983 cohort.

In conclusion, by comparing the long term results of very preterm infants, born ten years apart, we found that the decrease in mortality was accompanied by an increase of

disabling cerebral palsy. Although the incidence of learning and behavioural problems was similar over time, due to the survival of more babies, the absolute number of children with these problems also increased. The introduction of continuous monitoring of transcutaneous $pO_2$ and oxygen saturation probably has resulted in a significant decrease in visual problems, particularly strabismus. Modern perinatal care is not longer limited by the devastating effects of pulmonary problems as it was in the past, but fails in safeguarding cerebral integrity in very preterm infants. Ideally, the benefits of modern care should outweigh the disadvantages. Therefore, neonatologists should invent better strategies to protect the developing brain in the perinatal period. As long as this has not been achieved they should urge society to improve the number and quality of provisions to care for handicapped survivors of modern perinatal care.

## Acknowledgements.

The authors thank Ina Kloosterboer-Boerrigter (TNO Prevention and Health, Leiden), Boudien van Kessel-Feddema, Meta Sondaar and Simone Knuijt (Radboud University Nijmegen Medical Centre, Nijmegen), Aleid van Wassenaer, Judy Briët, Rinske Breur-Pieterse and Karin Koldewijn (Academic Medical Centre, Amsterdam), Jan Buijs and Marion Wit (Máxima Medical Centre, Veldhoven) for their participation in the Validation Study Group.

The children in the POPS cohort have been assessed by Sylvia Veen, Martina Ens-Dokkum and Anneke Schreuder.

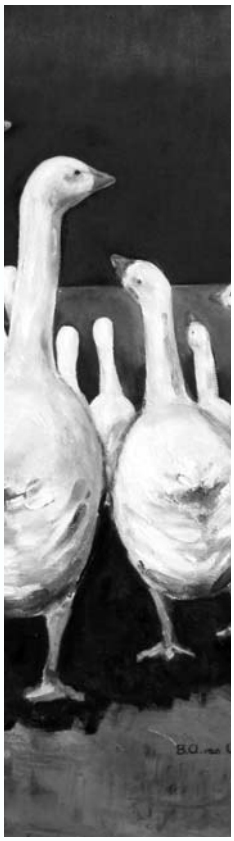
The study was supported by a grant of the Praeventiefonds, project number 28-2756 and ZorgOnderzoek Nederland (ZonMw), project number 10010004-20.

## References


1. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care. Pediatrics 1998;102:893-9.
2. Crowley P. Prophylactic corticosteroids for preterm birth. Cochrane Database Syst Rev 2000;CD000065.
3. Soll RF. Synthetic surfactant for respiratory distress syndrome in preterm infants (Cochrane Review). In: The Cochrane Library, 1, 2002.Oxford: Update Software. 2002. Oxford, The Cochrane Library.
4. Blair E. Trends in cerebral palsy. Indian Journal of Pediatrics. 2001;68:433-8.
5. Hagberg B, Hagberg G, Beckung E, Uvebrant P. Changing panorama of cerebral palsy in Sweden. VIII. Prevalence and origin in the birth year period 1991-94. Acta Paediatr 2001;90:271-7.
6. Pharoah PO, Platt MJ, Cooke T. The changing epidemiology of cerebral palsy. Arch Dis Child Fetal Neonatal Ed 1996;75:F169-73.
7. Cooke RW. Trends in incidence of cranial ultrasound lesions and cerebral palsy in very low birthweight infants 1982-93. Arch Dis Child Fetal Neonatal Ed 1999;80:F115-7.
8. Battin M, Ling EW, Whitfield MF, Mackinnon M, Effer SB. Has the outcome for extremely low gestational age (ELGA) infants improved following recent advances in neonatal intensive care? Am J Perinatol 1998;15:469-77.
9. Tin W, Wariyar U, Hey E. Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. BMJ 1997;314:107-11.
10. Hack M,.Fanaroff AA. Outcomes of children of extremely low birthweight and gestational age in the 1990's. Early Hum Dev 1999;53:193-218.
11. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986;1:55-7.
12. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birhtweight infants at five years of age. The collaboration project on preterm and small for gestational age infants (POPS) in the Netherlands. Lancet 1991;338:33-6.
13. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in the Netherlands in 1983. Early Hum Dev 2000;59:175-91.
14. Kollée LA, den Ouden AL, Drewes JG, Brouwers HA, Verwey RA, S.P. [Increase in perinatal referral to regional centers of premature birth in The Netherlands: comparison 1983 and 1993]. Nederl Tijdschr Geneesk 1998;142:131-4.
15. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr 1978; 92:529-34.
16. van de Bor M, Ens-Dokkum M, Schreuder AM, Veen S, Brand R, Verloove-Vanhorick SP. Outcome of periventricular-intraventricular haemorrhage at five years of age. Dev Med Child Neurol 1993; 35:33-41.
17. Bancalari E, Abdenour GE, Feller R, Gannon J. Bronchopulmonay dysplasia: clinical presentation. J Pediatr 1979;95:819-23.
18. Touwen BC. Examination of the child with minor neurological dysfunction. Clinics Dev Med 1989;11:353-363.

19. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991.
20. Brand R, van Hemel OJ, Elferink-Stinkens PM, Verloove-Vanhorick SP. Comparing mortality and morbidity in hospitals: theory and practice of quality assessment in peer review. Methods of Information in Medicine 1994;33:196-204.
21. Verloove-Vanhorick SP, Brand R. [Inter-hospital comparison of perinatal intensive care in academic hospitals]. Nederl Tijdschr Geneesk 1989;133:1883-6.
22. Paneth N, Qiu H, Rosenbaum P, Saigal S, Bishai S, Jetton J et al. Reliability of classification of cerebral palsy in low-birthweight children in four countries. Dev Med Child Neurol 2003;45:628-33.
23. den Ouden AL, Kok JH, Verkerk PH, Brand R, Verloove-Vanhorick SP. The relation between neonatal thyroxine levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996;39:142-5.
24. Inder TE, Wells SJ, Mogridge NB, Spencer C, Volpe JJ. Defining the nature of the cerebral abnormalities in the premature infant: a qualitative magnetic resonance imaging study. J Pediatr 2003;143:171-9.
25. Volpe JJ. Neurobiology of periventricular leukomalacia in the premature infant. Pediatr Res 2001;50:553-62.
26. Murphy BP, Inder TE, Huppi PS, Warfield S, Zientara GP, Kikinis R et al. Impaired cerebral cortical gray matter growth after treatment with dexamethasone for neonatal chronic lung disease. Pediatrics 2001;107:217-21.
27. Lorenz JM, Paneth N, Jetton JR, den Ouden L, Tyson JE. Comparison of management strategies for extreme prematurity in New Jersey and the Netherlands: outcomes and resource expenditure. Pediatrics 2001;108:1269-74.
28. Paneth N, Kiely JL, Stein Z, Susser M. Cerebral palsy and newborn care. III: Estimated prevalence rates of cerebral palsy under differing rates of mortality and impairment of low-birthweight infants. Dev Med Child Neurol 1981;23:801-6.
29. Stanley FJ. Cerebral palsy trends. Implications for perinatal care. Acta Ob Gyn Scand 1994;73:5-9.
30. Centraal Bureau voor de Statistiek. [Tabellen deelname speciaal onderwijs] Tables of participation in special education. http://statline.cbs.nl/StatWeb. 2003. Voorburg, Centraal Bureau voor de Statistiek.
31. Hille ET, den Ouden AL, Bauer L, van den Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: perinatal risk factors and predictors at five years of age. Collaborative Project on Preterm and Small for Gestational Age (POPS) Infants in The Netherlands. J Pediatr 1994;125:426-34.

Chapter 8



# General discussion and conclusions

Neonatal intensive care started in the 1960s when technology made long term artificial ventilation of preterm newborn infants with hyaline membrane disease possible. After forty years, neonatology has gone on in years, but the follow up is still in its infancy. We have a good understanding of the causal relation between perinatal risk factors and mortality and short term morbidity, but the relation between perinatal risk factors and long term outcome has only been partly clarified.

The first objective of this thesis was to develop and validate an assessment tool that could help paediatricians to identify, before six years of age, which very preterm and/or low birth weight survivors of neonatal intensive care have developmental disturbances that may interfere with normal education and normal life. We hypothesised that a paediatrician could identify, within one hour and with the help of a structured and standardised assessment, 80% of the very preterm and very low birth weight infants with developmental disturbances in the somatic, motor, language and behavioural domains, but would have considerable difficulties in the cognitive domain.

The second objective was to investigate the relation between changes in perinatal care and short term and long term outcome, mortality as well as morbidity, of very preterm infants in the Netherlands between 1983 and 1993. We hypothesised that the increased survival of very preterm infants would be accompanied by improved later outcome, i.e. that, related to the number of live-born preterm infants, the incidence of children without impairments would increase and the incidence of children with impairments would not increase

This implied that, related to the surviving children, we expected an increased incidence of children without impairments and a decreased incidence of handicaps.

## First objective

Because recognition of developmental impairments in survivors of neonatal intensive care increases with advancing age and these impairments may lead to school failure, the need for accurate developmental assessment is obvious [1-7].

In 1996, a national working party on neonatal follow up presented a programme for standardised follow up from birth to eight years of age that would offer an equal follow up programme to every individual survivor of neonatal intensive care and yield comparable follow up data for all [8].

The working party calculated the costs of such a follow up programme of survivors equivalent to the cost of one day of intensive care treatment [25], which amounted to £1000 in the United Kingdom and $1277 (at 32 weeks' gestation) to 2203 (at 25 weeks' gestation) in the USA [9,10].

Some studies indicate that early low cost intervention can improve cognition and behaviour [11], but cost-benefit studies in which the costs of screening preterm infants for developmental problems (other than retinopathy of prematurity and deafness) have been compared with the benefits of intervention are lacking [9].

Because of shortage of budget and staff, most neonatal departments are forced to restrict follow up programmes to the bare minimum. Therefore Field et al., for instance, evaluated two simple methods of collecting data in preterm infants born at ≤ 32 weeks of gestation, one by a questionnaire to the parents, the other by collecting clinical information from hospital admissions, routine health surveillance and outpatient letters [12]. The average cost per case recruited was £37 for the questionnaire and £61 for the community review. Compared to the costs of intensive care treatment, these figures are trivial, but both methods focused only on major health problems at the age of two years, failed to collect outcome data for at least 95% of the children as was the objective, and did not ascertain the reliability of the gathered data.

To include mild developmental problems that may interfere with learning, the Validation Study Group* developed a comprehensive programme for assessment at five years of age to enable paediatricians to identify, before school age, which children have developmental disturbances that may interfere with the acquisition of daily skills and normal life. In addition, such a comprehensive assessment would offer these paediatricians a standardised tool to evaluate the perinatal care.

*The Paediatrician's assessment*

The assessment consisted of a parental questionnaire and a paediatric evaluation. The Validation Study Group intended a three-step procedure, consecutively consisting of a questionnaire, a comprehensive Paediatrician's assessment and detailed standardised and domain-specific assessment batteries. Each step aimed at identifying children with definite developmental impairments in one or more domains, children with possible impairments and children without impairments. The definitely impaired children should be referred for further domain-specific analysis and treatment, the optimal children labelled optimal and discharged from further paediatric control. The children with suspected or possible impairments should enter the next step.

The first step, the questionnaire, addressed the following domains: socioeconomic status, general health, visual and auditory function, motor development, cognitive development, language, behaviour and school outcome. Several parts had been designed for the assessment of the very preterm and very low birth weight infants born in 1983, the POPS study [13]. The questionnaire included a shortlist of 15 items of the Child Behavior Checklist.

The questionnaire gave information on the individual child and saved time during the next step, the Paediatrician's assessment. However, it failed as a simple tool for identification of five year old children with normal development, who would need no further paediatric assessment to evaluate their outcome.

In the second step, the paediatricians used the information from the questionnaire, assessed neurological function by the Touwen examination, neuromotor and cognitive development by the Denver Development Screening Test (DDST) and language development by the Dutch Language Screening Test [14-16]. This Paediatrician's assessment took one hour on average, 40 to 50 minutes in normal children to sometimes much longer than 60 minutes in children with multiple impairments. On the basis of the results of all tests, paediatricians gave an overall judgement of health, neuromotor development, cognition, language and behaviour. In summarising their findings, they gave a final judgement: 'optimal' (all domains normal), 'at risk' (further examination necessary) or 'abnormal' (treatment necessary or already treated for impairment in one or more domains).

The result of the Paediatrician's assessment was validated in 566 very preterm and low birth weight infants against a battery of standardised tests for neuromotor and cognitive development and behaviour that could be regarded as the third step. This included the Movement Assessment Battery for Children (movement ABC) by a child physiotherapist [27], the revised Amsterdam child intelligence test (IQ test) by a developmental psychologist [26] and the full Child Behavior CheckList (CBCL) filled in by the parents [19,20].

We hypothesised that paediatricians could identify, with the help of the comprehensive assessment, at least 80% of the very preterm and/or low birth weight infants with developmental disturbances in the somatic, motor, cognitive, language and behavioural domains that may interfere with normal education and normal life.

The study showed that the sensitivity of this Paediatrician's assessment, tested against a battery of standardised tests, was low. Only 48% of the children with abnormal results on one or more of the standard tests were identified and 52% were not detected. The majority of the missed cases had either motor problems or behaviour problems or a combination, a minority failed on the IQ test.

When tested by the standardised tests, the movement ABC, the IQ test and the CBCL, 56%, 78% and 77% performed optimally. The combined results showed optimal performance on all three tests in 167 children (43.3%), on two tests in 134 (34.7%), on one test in 61 (15.8%) and on none of the tests in 24 (6.2%) children (table 1). The definition of optimal on the various tests was above P15 for the movement ABC test, above -1 standard deviation for the IQ test and above P10 for the CBCL. This means that in a normal population 0.85 x 0.84 x 0.90 = 64% will perform optimally by this

**Table 1:** 386 not-severely handicapped, very preterm and low birth weight survivors of neonatal intensive care at five years of age that were classified optimal or non-optimal by respectively the movement ABC, IQ test and/or CBCL.

| | N | % |
|---|---|---|
| Movement ABC optimal, IQ optimal, CBCL optimal | 167 | 43.3% |
| Movement ABC optimal, IQ optimal, CBCL non-optimal | 42 | 10.9% |
| Movement ABC optimal, IQ non-optimal, CBCL optimal | 8 | 2.1% |
| Movement ABC optimal, IQ non-optimal, CBCL non-optimal | 4 | 1.0% |
| | | |
| Movement ABC non-optimal, IQ optimal, CBCL optimal | 84 | 21.8% |
| Movement ABC non-optimal, IQ optimal, CBCL non-optimal | 17 | 4.4% |
| Movement ABC non-optimal, IQ non-optimal, CBCL optimal | 40 | 10.4% |
| Movement ABC non-optimal, IQ non-optimal, CBCL non-optimal | 24 | 6.2% |
| | 386 | 100% |

definition when the results of the tests are independent of each other, almost twice the number in our very preterm population.

In a population with such a high prevalence of developmental problems, it is necessary to detect as many children with developmental problems as possible; some over-referral is acceptable. Therefore, tests should aim at high negative predicting values and likelihood ratios of negative tests close to zero. The likelihood ratio of the negative test in the present study indicated that 59% of the children with a negative test were impaired.

We therefore conclude that follow up studies that do not include detailed and standardised tests for several domains will underestimate developmental problems in survivors of neonatal intensive care and may be the reason for incomparability of follow up results and a seemingly increase of developmental problems at a later age. There is no simple way to identify children in need of extra help or to enable a true evaluation of neonatal intensive care.

Because motor disturbances were both the most frequent and the most frequently missed by the paediatrician, we further analysed the relations between the paediatric judgement of motor performance, the Touwen examination, the DDST and the movement ABC (Chapter 3). Subsequently, we also analysed the relations between the paediatric judgement, the DDST, the Language Screening Test and the IQ test (Chapter 4) and between the Dutch Language Screening Test and a formal language assessment (Chapter 5).

*Motor development*

In the individual child, functional motor capacity is not an isolated event, but emerges from physical abilities and conditions and executes through the cooperative effort of many brain structures and processes. When we study overt goal-directed motor behaviour in children, we are looking at the result of a long chain of processes that produce and coordinate the observed movements. Therefore, motor performance problems may be related to cognitive, behavioural, visual-perceptual and health factors [21,22]. The integrity of the central nervous system is an important condition for the development of perceptual-motor skills in preterm infants and many studies have focused on the relationship between cerebral anatomical abnormalities and developmental outcome [23,24]. The main purpose of a screening assessment is not to detect these relations, but to determine if the quality and quantity of the motor performance is normal or abnormal for age. When an abnormal score is found, a more detailed multidisciplinary assessment is necessary to unravel which developmental aspects are the most determining factors.

To detect developmental motor problems, most paediatricians use neurological examinations to test the integrity of the nervous system. We therefore compared the neurological examination according to Touwen with the movement ABC, that indicates motor function in daily life [17]. It is most widely used to diagnose developmental motor coordination disorder in several circumstances [21,24,25], as well as in follow up studies of premature infants [5,26-36].

With the movement ABC, we found in our population of very preterm infants without severe handicaps clinically important motor disorders in 22.0% and borderline disturbances in 21.5%. This is four times and about twice the figure of the normal population [31]. Several authors found clinically important motor disorders in 20% to 30% of all very preterm infants and in 50% to 55% of them when borderline motor problems were included [5,27,28,31,34,37]. Our slightly lower incidence can be explained by the exclusion of the 3% most severely handicapped children in our study.

Comparison of the examination of neuromotor development by a paediatrician with the movement ABC by a child physiotherapist showed that paediatricians, even when specifically trained for the occasion, are not very capable in identifying children with developmental motor disorders. Paediatricians may be biased by their knowledge of the impressive perinatal problems and the relief that the development looks better than expected. They may also be biased because they seldom assess normal children and they therefore easily accept minor motor impairments. The paediatric examination may also be subjective because there is no rigid structure for assessing or counting the results. We could improve the sensitivity of the Touwen examination from 0.19 to 0.62

when the algorithm was calculated by computer.

Comparison of the Touwen examination, the DDST and the movement ABC show that these assessments test different aspects of motor development. The Touwen examination aims at the neurological development, the DDST at general development and the movement ABC at functional performance of motor tasks.

The movement ABC indicates motor functioning in daily life, which means that it tests the functions children need at school every day. Marlow et al. described in a longitudinal study that poor performance at the movement ABC at six years of age is predictive for later school problems and indicates that the longitudinal development may be threatened [26,38]. Diamond too documented a close interrelation between motor development and cognitive development [39]. We found that 64% of the children with non-optimal movement ABC scores already had school problems. We suppose that in these children concentration and attention problems are present and learning capacity is decreased.

As the movement ABC is widely accepted and is relatively cheap to perform in terms of time and costs, we recommend using movement ABC for the assessment of motor function.

On the other hand, we documented 13 children with optimal movement ABC and unmistakable cerebral palsy. Apparently these children had learned to function well in spite of some constraints of their neuromotor system. Therefore they met the accuracy and velocity requirements of the movement ABC. To detect these children a thorough neurological assessment, such as the Touwen examination, should be added to the assessment of these children. These children do not cause a problem in clinical decision making, since they might be referred for further assessment. When they get older, only longitudinal follow up is necessary to detect if they learn new motor tasks as skilfully as their normal peers when the task demands increase. For the evaluation of perinatal care however, the detection of cerebral palsy as a marker of cerebral damage is essential.

If both purposes, early identification of children who need extra help and evaluation of perinatal care, have to be met, systematic follow up assessment should include both the movement ABC and a thorough, neurological assessment like the Touwen examination.

*Cognitive development*

The objective was to evaluate if a standardised paediatric assessment enables identification of prematurely born children at five years of age who should be referred to a child psychologist for a standardised cognitive assessment. Therefore, we compared a paediatric judgement, based on the parental questionnaire, the overall impression during examination, the Denver Developmental Screening Test (DDST) and the Dutch Language Screening Tests (Taal Screening Test, TST) in 368 very preterm and/or very low birth weight infants, not yet diagnosed with cognitive impairments, with the results of a standardised intelligence test, the Revised Amsterdam KinderIntelligentie Test (RAKIT).

The sensitivity of the overall judgement of cognition by the paediatrician (sensitivity 0.34) and of the DDST subscales (sensitivity 0.18 - 0.36) was low and underestimated the real need for formal cognitive testing. The complete DDST (sensitivity 0.72) and school performance (sensitivity 0.79) performed best. Paediatricians often use the development of language as an indicator of cognition, but the sensitivity of the Language Screening Test was disappointingly low (0.55). We documented in chapter 5 that the performance of the Language Screening Test for the detection of speech and language problems may be improved by lowering the cut off point from 18 to 17 as the lower limit of abnormal language development [15].

School performance obviously is a good indicator of cognition. The decision to refer a prematurely born child for further assessment may be based on the information on school performance. Such information can be obtained easily and without further constraint and costs. It may therefore be used as a screening that reduces the costs of a full developmental assessment by a psychologist for all preterm born children. However, when the aim is to remediate or even prevent school problems, referral of the child after the learning problems have emerged is too late. Since 37% of the children in this study already have school assistance or grade retention, a preschool cognitive assessment is to be recommended. We therefore recommende not to wait until cognitive deficits have led to problems at school, but to test all these children with a thorough, formal assessment of their IQ with up to date tests before school entry.

We recommend that school problems in very preterm born children are taken serious and should lead to further formal assessment. It is even better not to wait until cognitive deficits have led to problems at school, but to test all very preterm born children at five years of age with a thorough, formal assessment of their IQ with up to date tests before they enter elementary school at six years of age.

We used these data also to describe a population of very preterm and very low birth weight infants at five years of age. In this '1993' cohort we measured IQ in 454 of 566 eligible children. Average IQ was 96.6 (S.D. 16.4); 78% of the children had IQ's in the normal range, 16% between -1 and -2 standard deviations, and 6% below -2 standard deviations. The patients of the thyroxine supplementation study, excluded in chapter 4, were also included. We excluded 21 children with known severe handicap. If we assume that these 21 children had IQ's below -2 standard deviations, 75% of all children had measured optimal IQ's above minus one, 15% between -1 and -2 standard deviations and 10% below -2 standard deviations. Studies comparing same aged controls usually find average IQ's 7 - 15 points (0.5 to 1 standard deviations) lower for all very low birth weight infants and 7 - 10 points lower for very low birth weight infants without major neurological impairment. Most studies find severe cognitive impairments (< -2 standard deviations) in 10% to 25% of all very low birth weight infants, compared to the expected 2.3% in a normal population [3,40,41].

Several authors showed that non-responders have less favourable outcome than responders [42-44]. The response percentage in our population was 80% of the eligible children. The perinatal risk factors did not differ between the assessed and non-assessed children, with the exception of the lower incidence of multiple births in the non-assessed group. Therefore, our study may slightly overestimate the actual average IQ in very preterm and very low birth weight infants because of selection bias and non-response, but is otherwise in line with most data from the literature.

Measurements of IQ, however, may overestimate the true average IQ because of the Flynn effect [45]. Flynn has found that the IQ increases approximately 0.3 points per year. The RAKIT was standardised in 1984. With a yearly increase of 0.3 points, mean IQ in a contemporary control group may well be 5 points higher than the norm we used. Moreover, intelligence tests such as the RAKIT may overestimate true intellectual capacities, because, in contrast to everyday situations, the tasks that the children have to perform during these tests are structured, one-dimensional and taken in a one-to-one situation without disturbances from outside the test environment.

*Language*

Speech and language are important areas of child development. Therefore, a language screening test (the Dutch 'Taal Screening Test', TST) was included in the Paediatrician's assessment. To validate the TST in this group of preterm infants, the results of the TST performed by the paediatrician were compared with the results of the standardised language tests performed by a professional speech therapist. The main goal was to examine the diagnostic efficiency and the usefulness of the TST to identify

preterm infants with a language delay accurately. Based on the total number of errors, children were categorised as 'normal' (≤ 17 errors), 'moderately delayed' (18 - 25 errors) or 'severely delayed' (≥ 26 errors) [20]. When moderately and severely delayed were combined (≥ 18 errors), the sensitivity of the TST was 62% and the specificity 91%. The sensitivity could be increased to 79% and the negative predictive value to 95% when 17 errors or more were considered as delayed.

The Dutch Language Screening Test (Taal Screening Test, TST) is an excellent instrument to identify speech and language delay in five-year-old very preterm and/or low birth weight infants, provided the cut off point is lowered from 18 to 17 errors. We recommend the incorporation of this screening instrument in the systematic follow up assessments at five years of age.

## Second objective

*Perinatal changes and short- and long term outcome*

We compared two geographically defined populations ten years apart. Socioeconomic circumstances in this decade led to a higher maternal age and thereby to an increase in multiple births and preterm birth, partly by assisted reproduction. In spite of these unfavourable factors that caused an increase in frequency of preterm birth, we found a 40% reduction of mortality in almost every week of gestation. Richardson has suggested that one third of the improvement in mortality might be attributed to a better condition on admission and two thirds to a more effective neonatal intensive care [47]. Our findings suggest also that not one specific factor, but the improvement of many aspects of obstetric and neonatal care, including prenatal transfer, earlier Caesarean section, systematic antenatal administration of steroids and especially the administration of surfactant, has accomplished this improved survival.

The increased survival, however, was not accompanied by a decreased incidence of intraventricular haemorrhage as percentage of the survivors, and the incidence of BPD even increased. So, although more children survived healthy, the improved perinatal care did not lead to fewer impaired children. This has also been documented by the Leiden Follow up Project on Prematurity [48].

When we studied long term outcome in the two populations, we found an increase in disabling cerebral palsy at five years of age. Not only the incidence of disabling cerebral palsy, related to the number of live-born infants, increased, but also did the incidence in survivors. This outcome is alarming. Hagberg found an increasing

incidence of cerebral palsy in the Swedish population during the 1970s and suggested a relation with the improved survival of preterm infants [49]. The decade thereafter, the incidence of cerebral palsy stabilised or even decreased [50]. It looks as if the history of an increased survival in combination with an increased incidence of handicaps in preterm infants in the 1970s has repeated itself in very preterm and extremely preterm infants in the late 1980s and early 1990s [51]. Further analyses of our data in infants below 28 weeks' gestation and infants of 28 - 31 weeks' gestation show this effect to be present in both subpopulations, but more distinctly in the subpopulation of infants below 28 weeks' gestation.

A difference in cognition could not be established, because in the 1983 cohort no formal cognitive test was applied. The difference in school results, although in favour of the 1993 cohort, was not statistically significant. In both cohorts, the percentage of children in special education (11.9% and 9.8%) was substantially higher than in the general population.

The incidence of problem behaviour in both cohorts was twice that of a control population and did not decrease over time.

Because the incidence of neuromotor, cognitive and behavioural problems in survivors did not decrease in the decade studied, while mortality dropped to half of what it was before, the incidence of children with impairments and school problems, related to the number of live-born very preterm infants, almost doubled as will the need for provisions to treat these children in future.

The incidence of visual problems and strabismus in survivors of neonatal intensive care decreased. We assume that this is due to the introduction of continuous monitoring of transcutaneous $pO_2$ and oxygen saturation between the 1980s and the 1990s. Several authors reported the close relation between visual problems and neurodevelopment outcome [22,52,53]. Ongoing attention to careful monitoring is therefore important as has been stressed repeatedly by Cats and colleagues [54,55].

## Conclusion

Simultaneous with the improved survival of very preterm and low birth weight infants, the awareness increased that not all survive without sequelae. Mild developmental disturbances that interfere with the acquisition of everyday skills and normal learning appeared to be very frequent and were not detected until at school age. The expectation was that such developmental disabilities were temporary side effects and would diminish with increasing perinatal knowledge and experience. The data in this thesis show that impairments, disabilities and handicaps remain invariably high,

with the exception of the decreased incidence of visual problems. The increased incidence of cerebral palsy of children born in 1993, compared to 1983, is worrisome.

Factors that are responsible for a decreased mortality such as antenatal steroids, artificial ventilation and surfactant, do not decrease these developmental problems in the surviving children. The use of dexamethasone in the neonatal period is a serious extra risk factor for developmental problems.

Severe developmental problems are mostly diagnosed at the age of two years. Diagnosis of mild motor problems, learning disabilities and problem behaviour however, is mostly delayed until school age. When with advancing age the demands imposed by society increase, the problem rate increases even further. Singularly, these deficits may be subtle, but combined, without compensation in other domains, they may pose a lifelong burden. Therefore, long term longitudinal follow up of these children is necessary.

Paediatricians who assess very preterm and low birth weight infants with assessments that include structured elements from the Touwen examination, Denver Development Screening Test, Dutch Language Screening Test and a shortlist from the Child behaviour Checklist overlook developmental motor coordination disorders and learning and behavioural problems. As the movement ABC is an accurate and widely used instrument to detect these motor coordination disorders, the movement ABC should be included in the follow up of very preterm and low birth weight infants at the age of five. To detect all neurological problems a neurological assessment is also necessary. Cognitive problems are indicated by the appearance of learning difficulties at school. As the presence of learning difficulties in very preterm and low birth weight infants is an alarm-signal for cognitive delay, referral and formal cognitive assessment by a child psychologist is indicated. Repeating the school year may not be a good solution. To identify cognitive delays that may interfere with learning before school age, formal cognitive assessment before six years of age should be incorporated in a systematic follow up program.

As over one half of all very preterm and low birth weight infants show these developmental problems, routine assessment of all very preterm and low birth weight infants before school age is necessary. Standardised and normalised test instruments that cover all developmental domains must have a rigid structure for managing the test process and calculating the results General judgement by paediatricians in a routine outpatient clinic, even when these paediatricians are well trained, is insufficient.

Follow up studies that do not include detailed standardised tests for several domains will underestimate developmental problems in survivors of neonatal intensive care. There is no simple way to identify children in need of extra help or to enable a true evaluation of neonatal intensive care.

This thesis deals with problems in very preterm and low birth weight infants. It also documented, however, that modern intensive care saves the lives of many children 'born to soon or born too small' [56] and that many of them grow up without any problem at all. Although modern neonatal intensive care will not restore complete integrity of all preterm born infants, it certainly is worthwhile.

## Acknowledgement

* The Validation Study Goup consisted of A. Lya den Ouden (TNO Prevention and Health, Leiden and Health Care Inspectorate, Ministry of Health, The Hague), Ina Kloosterboer-Boerrigter (TNO Prevention and Health, Leiden), Anneloes van Baar (University of Tilburg), Louis A.A. Kollée, Maria W.G. Nijhuis-van der Sanden, Boudien J.M. van Kessel-Feddema, Meta Sondaar and Simone Knuijt (Radboud University Nijmegen Medical Centre, Nijmegen), Adri Ilsen, Aleid G. van Wassenaer, Judy M. Briët, Rinske Breur-Pieterse and Karin Koldewijn (Academic Medical Centre, Amsterdam), Martin J.K. de Kleine, Jan Buijs and Marion J. Wit (Máxima Medical Centre, Veldhoven).

Ronald Brand, Department of Medical Statistics, Leiden University Medical Centre advised on the statistical methods and analyses and S. Pauline Verloove-Vanhorick, TNO Prevention and Health, Leiden, and Department of Child Health, Leiden University Medical Centre supervised the study.

The study was supported by a grant of the Dutch Health Organisations Praeventiefonds, project number 28-2756, and ZorgOnderzoek Nederland (ZonMw), project number 1010004-20.

## References


1. Hille ET, den Ouden AL, Bauer L, van den Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: perinatal risk factors and predictors at five years of age. Collaborative Project on Preterm and Small for Gestational Age (POPS) Infants in The Netherlands. J Pediatr 1994;125:426-34.
2. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in the Netherlands in 1983. Early Hum Dev 2000;59:175-91.
3. Wolke D. Psychological development of prematurely born children. Arch Dis Child 1998;78:567-70.
4. Horwood LJ, Mogridge N, Darlow BA. Cognitive, educational, and behavioural outcomes at 7 to 8 years in a national very low birthweight cohort. Arch Dis Child Fetal Neonatal Ed 1998;79:F12-F20.
5. Botting N, Powls A, Cooke RW, Marlow N. Cognitive and educational outcome of very-low-birthweight children in early adolescence. Dev Med Child Neurol 1998;40:652-60.
6. McGrath MM, Sullivan MC, Lester BM, Oh W. Longitudinal neurologic follow-up in neonatal intensive care unit survivors with various neonatal morbidities. Pediatrics 2000;106:1397-405.
7. Weindrich D, Jennen-Steinmetz C, Laucht M, Schmidt MH. Late sequelae of low birthweight: mediators of poor school performance at 11 years. Dev Med Child Neurol 2003;45:463-9.
8. Werkgroep Landelijke Neonatale Follow-up van de Sectie Neonatologie Nederlandse Vereniging voor Kindergeneeskunde, 1996.
9. Centre d'expertise d'Institut national de la santé et de recherche médicinale (INSERM). Approche économique du dépistage et de la prise en charge In: Déficiences ou handicaps d'origine périnatale. Dépistage et prise en charge. Paris: INSERM 2004, Chapter 12:169-177.
10. Gilbert WM, Nesbitt TS, Danielsen B. The cost of prematurity: quantification by gestational age and birth weight. Obstet Gynecol 2003;102:488-92.
11. Feldman R, Eidelman AI. Intervention programs for premature infants. How and do they affect development? Clin Perinatol 1998;25:613-26.
12. Field D, Draper ES, Gompels MJ, Green C, Johnson A, Shortland D, Blair M, Manktelow B, Lamming CR, Law CL. Measuring later health status of high risk infants: randomised comparison of two simple methods of data collection. BMJ 2001;323:1-5.
13. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birhtweight infants at five years of age. The collaboration project on preterm and small for gestational age infants (POPS) in the Netherlands. Lancet 1991;338:33-6.
14. Touwen BC. Examination of the child with minor neurological dysfunction. Clinics in Developmental Medicine 1989;11:353-63.
15. Cools ATM, Hermans JMA. Dos-handleiding. Denver Ontwikkeling Screeningstest. Lisse: Swets & Zeitlinger, 1979.
16. Gerritsen FME. VTO Taalscreening 3- tot 6-jarigen: de ontwikkeling van een taalscreeningsinstrument voor gebruik in de Jeugdgezondheidszorg. Amsterdam/Lisse: Swets & Zeitlinger, 1988.
17. Henderson SE, Sugden DA. The Movement Assessment Battery for Children. San Antonio: Psychological Corporation, 1992.
18. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. Intelligentie-meting bij kinderen: empirische en methodologische verantwoording van de geReviseerde Amsterdamse Kinder Intelligentie Test. Lisse: Swets & Zeitlinger, 1987.
19. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991.

20. Verhulst FC, van-der-Ende J, Koot HM. Manual for the Child BehaviorChecklist/4-18 (Dutch version). Rotterdam: Department of Childpsychiatry, Sophia Children's Hospital/Academic Hospital Rotterdam/Erasmus Universiteit Rotterdam, 1996.
21. Gillberg C, Kadesjo B. Why bother about clumsiness? The implications of having developmental coordination disorder (DCD). Neural Plast 2003;10:59-68.
22. Scherjon S, Briet J, Oosting H, Kok J. The discrepancy between maturation of visual-evoked potentials and cognitive outcome at five years in very preterm infants with and without hemodynamic signs of fetal brain-sparing. Pediatrics 2000;105:385-91.
23. Jongmans M, Henderson S, de Vries L, Dubowitz L. Duration of periventricular densities in preterm infants and neurological outcome at 6 years of age. Arch Dis Child 1993;69:9-13.
24. Skranes JS, Vik T, Nilsen G, Smevik O, Andersson HW, Brubakk AM. Cerebral magnetic resonance imaging and mental and motor function of very low birth weight children at six years of age. Neuropediatrics 1997;28:149-54.
24. Geuze RH, Jongmans M, Schoemaker M, Smits-Engelsman B. Developmental coordination disorder. Hum Mov Sci 2001;20:1-5.
25. Henderson SE, Henderson L. Toward an understanding of developmental coordination disorder: terminological and diagnostic issues. Neural Plast 2003;10:1-13.
26. Marlow N, Roberts L, Cooke R. Outcome at 8 years for children with birth weights of 1250 g or less. Arch Dis Child 1993;68:286-90.
27. Powls A, Botting N, Cooke RW, Marlow N. Motor impairment in children 12 to 13 years old with a birthweight of less than 1250 g. Arch Dis Child Fetal Neonatal Ed 1995;73:F62-6.
28. Powls A, Botting N, Cooke RW, Marlow N. Handedness in very-low-birthweight (VLBW) children at 12 years of age: relation to perinatal and outcome variables. Dev Med Child Neurol 1996;38:594-602.
29. Katz KS, Dubowitz LM, Henderson S, Jongmans M, Kay GG, Nolte CA et al. Effect of cerebral lesions on continuous performance test responses of school age children born prematurely. J Pediatr Psychol 1996;21:841-55.
30. Katz-Salamon M, Gerner EM, Jonsson B, Lagercrantz H. Early motor and mental development in very preterm infants with chronic lung disease. Arch Dis Child Fetal Neonatal Ed 2000;83:F1-6.
31. Jongmans M, Mercuri E, de Vries L, Dubowitz L, Henderson SE. Minor neurological signs and perceptual-motor difficulties in prematurely born children. Arch Dis Child Fetal Neonatal Ed 1997;76:F9-14.
32. Torrioli MG, Frisone MF, Bonvini L, Luciano R, Pasca MG, Lepori R et al. Perceptual-motor, visual and cognitive ability in very low birthweight preschool children without neonatal ultrasound abnormalities. Brain Dev 2000;22:163-8.
33. Cooke RW, Foulder-Hughes L. Growth impairment in the very preterm and cognitive and motor performance at 7 years. Arch Dis Child 2003;88:482-7.
34. Erikson C, Allert C, Carlberg EB, Katz-Salamon M. Stability of longitudinal motor development in very low birthweight infants from 5 months to 5.5 years. Acta Paediatr 2003;92:197-203.
35. Foulder-Hughes LA, Cooke RW. Motor, cognitive, and behavioural disorders in children born very preterm. Dev Med Child Neurol 2003;45:97-103.
36. Mercuri E, Barnett A, Rutherford M, Guzzetta A, Haataja L, Cioni G et al. Neonatal cerebral infarction and neuromotor outcome at school age. Pediatrics 2004;113:95-100.
37. Hall A, McLeod A, Counsell C, Thomson L, Mutch L. School attainment, cognitive ability and motor function in a total Scottish very-low-birthweight population at eight years: a controlled study. Dev Med Child Neurol 1995;37:1037-50.
38. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. N Engl J Med 2000;343:378-84.

39. Diamond A. Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. Child Dev 2000;71:44-56.
40. Wolke D, Meyer R. Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. Dev Med Child Neurol 1999;41:94-109.
41. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 2002;288:728-37.
42. Wolke D, Ratschinski G, Ohrt B, Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. Eur J Pediatr 1994;153:906-15.
43. Tin W, Fritz S, Wariyar U, Hey E. Outcome of very preterm birth: children reviewed with ease at 2 years differ from those followed up with difficulty. Arch Dis Child Fetal Neonatal Ed 1998;79:F83-7.
44. Wariyar UK, Richmond S. Morbidity and preterm delivery; importance of 100% follow-up. Lancet 1989;8634:387-8.
45. Flynn JR. Massive IQ gains in 14 nations: what IQ test really measure. Psychological Bulletin 1987;101:171-91.
46. Flynn JR. Searching for Justice: the discovery of IQ gains over time. American Psychologist 1999;54:5-20.
47. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care. Pediatrics 1998;102:893-9.
48. Stoelhorst GMSJ. Development, Quality of Life and Behavior at 2 years of age in very preterm infants. The Leiden Follow-Up Project on Prematurity 1996-1997. Thesis. Leiden University 2003.
49. Hagberg B, Hagberg G, Zetterstrom R. Decreasing perinatal mortality—increase in cerebral palsy morbidity. Acta Paediatr Scand 1989;78:664-70.
50.. Hagberg B, Hagberg G, Beckung E, Uvebrant P. Changing panorama of cerebral palsy in Sweden. VIII. Prevalence and origin in the birth year period 1991-94. Acta Paediatr 2001;90:271-7.
51. Hagberg B. Lessons and indications from three decades of West-Swedish Cerebral Palsy data. Neuropediatrics 2000;31:284-6.
52. Hof-van Duin J, Cioni G, Bertuccelli B, Fazzi B, Romano C, Boldrini A. Visual outcome at 5 years of newborn infants at risk of cerebral visual impairment. Dev Med Child Neurol 1998;40:302-9.
53. Cioni G, Bertuccelli B, Boldrini A, Canapicchi R, Fazzi B, Guzzetta A et al. Correlation between visual function, neurodevelopmental outcome, and magnetic resonance imaging findings in infants with periventricular leucomalacia. Arch Dis Child Fetal Neonatal Ed 2000;82:F134-40.
54. Termote J, Schalij-Delfos NE, Donders AR, Cats BP. The incidence of visually impaired children with retinopathy of prematurity and their concomitant disabilities. J AAPOS. 2003;7:131-6.
55. Schalij-Delfos NE, Termote JU, Cats BP. [Retinopathy in premature infants]. Ned Tijdschr Geneesk 2002;146:977-80.
56. Neligan GA, Kolvin I, Scott DM, Garside RF. Born Too Soon or Born Too Small. Cambridge: Cambridge University Press, 1976.

# Samenvatting

**Samenvatting van het proefschrift**
**‘Follow-up onderzoek van veel te vroeg geboren kinderen op de leeftijd van vijf jaar’**

## Inleiding

Neonatale intensieve zorg is begonnen in de jaren ‘60 toen technologische ontwikkelingen het mogelijk maakten om te vroeg geboren (zwangerschapsduur < 37 weken) kinderen met respiratoir distress syndroom kunstmatig te beademen. Tussen 1960 en 1980 namen daardoor de overlevingskansen van te vroeg geboren kinderen aanzienlijk toe.

De overleving van *veel* te vroeg geboren (zwangerschapsduur < 32 weken) kinderen is in de jaren ’80 en ‘90 aanzienlijk toegenomen. Follow-up van deze kinderen gedurende de eerste twee jaren van het leven laat bij 10 - 20% chronische longziekten en handicaps zien, zoals spasticiteit, ernstige achterstand in psychomotore ontwikkeling en ernstige stoornissen van het gehoor en het gezichtsvermogen. Wij hebben een redelijk begrip van de relaties tussen de problemen rond de geboorte en ernstige stoornissen in de ontwikkeling. Lange-termijn follow-up van veel te vroeg geboren kinderen tot aan de schoolleeftijd laat een nog hogere frequentie van ontwikkelingsstoornissen zien. Deze stoornissen betreffen de visus, het gehoor, de motoriek, de cognitie, het leren op school en het gedrag.

Op het moment van ontslag uit het ziekenhuis is de uitkomst van het individuele kind op lange termijn moeilijk voorspelbaar. Op de leeftijd van twee jaar zijn de grote handicaps meestal wel duidelijk.

De diagnose milde motorische stoornissen, leerproblemen en/of gedragsstoornissen wordt meestal pas tijdens de schoolleeftijd gesteld. Daarom is lange-termijn follow-up van deze kinderen nodig. Dergelijke lange-termijn follow-upprogramma’s dienen drie doelen: (1) het ontdekken van ontwikkelingsstoornissen bij het individuele kind waarvoor interventie nodig is, (2) het evalueren van de perinatale behandeling en (3) het bijdragen aan de wetenschappelijke kennis over de relatie tussen schade in de perinatale periode en een afwijkende ontwikkeling later. Herkenning van deze relaties zal moeten leiden tot primaire preventie.

In Nederland wordt de overgrote meerderheid van de veel te vroeg geboren kinderen behandeld in een van de tien centra voor perinatale intensieve zorg. Na behandeling worden de overlevende kinderen terug verwezen naar hun ‘eigen’ ziekenhuis voor verdere neonatale behandeling en voor nazorg in de eerste levensmaanden. In de eerste levensjaren komen deze kinderen een aantal keren terug in het centrum van hun initiële behandeling voor het vaststellen van de follow-upgegevens. Hoewel de perinatale behandeling in de tien perinatale centra vergelijkbaar is, varieert de postneonatale zorg in ons land en zijn de follow-upprogramma’s niet gestandaardiseerd. Om deze redenen ontwikkelde de Werkgroep Landelijke Neonatale Follow-up van de Nederlandse Vereniging voor Kindergeneeskunde een gestandaardiseerd multidisciplinair follow-

upprogramma, dat zowel nazorg in het eerste levensjaar, als gestandaardiseerde follow-up in de daaropvolgende jaren zou moeten verzorgen voor alle veel te vroeg geboren kinderen in Nederland. De werkgroep berekende de kosten van een dergelijk follow-upprogramma op een bedrag equivalent aan de kosten van één IC-dag.

Een dergelijk programma is kostbaar. Daarom wordt in het eerste deel van dit proefschrift de ontwikkeling en validatie van een geïntegreerd kinderartsenonderzoek beschreven. Met dit onderzoek zouden kinderartsen in staat moeten zijn om, binnen één uur en vóór de leeftijd van zes jaar, vast te stellen bij welke veel te vroeg geboren kinderen ontwikkelingsstoornissen zijn ontstaan die een belemmering zouden kunnen vormen voor een normale scholing en een normaal leven. Wij veronderstelden dat een kinderarts, met behulp van een dergelijk gestructureerd en gestandaardiseerd onderzoek, 80% van de veel te vroeg geboren kinderen met een ontwikkelingsstoornis van de gezondheid, de motoriek, de spraak en taal en/of het gedrag zou kunnen herkennen. Wij veronderstelden ook dat hij meer moeite zou hebben met het onderkennen van stoornissen in de cognitie.

In het tweede deel van dit proefschrift wordt het verband tussen veranderingen in perinatale zorg tussen 1983 en 1993 enerzijds en korte en lange termijn uitkomst, zowel mortaliteit als morbiditeit, anderzijds beschreven. Wij veronderstelden dat de verbeterde neonatale zorg voor veel te vroeg geboren kinderen niet alleen zou hebben geleid tot toegenomen overleving, maar ook tot een verbeterde lange-termijn uitkomst. Dat wil zeggen dat, berekend op het aantal levendgeborenen, een hoger percentage gezonde kinderen in leven zou blijven en het percentage kinderen met een gestoorde ontwikkeling niet zou toenemen.

## Deel I. Ontwikkeling en validatie van een kinderartsenonderzoek

Het onderzoek bestond uit drie opeenvolgende stappen: een vragenlijst voor de ouders, een gestandaardiseerd onderzoek door een kinderarts (samen het kinderartsenonderzoek genoemd) en een batterij van gestandaardiseerd en gevalideerde, domein-specifieke testen voor de ontwikkeling van motoriek, cognitie, spraak en taal en gedrag: de movement Assessment Battery for Children (movement ABC), de gereviseerde Amsterdam KinderIntelligentieTest (RAKIT), in een van de drie deelnemende centra uitgebreide spraaktaaltesten, en de volledige Child Behavior CheckList (CBCL). Elke stap had tot doel om kinderen met een zekere ontwikkelingsstoornis, kinderen met mogelijk een ontwikkelingsstoornis en kinderen zonder een ontwikkelingsstoornis te selecteren. Kinderen met een zekere of mogelijke stoornis zouden de volgende stap moeten ondergaan zo nodig leidend tot

domeingerichte behandeling; kinderen zonder stoornis zouden moeten worden gelabeld als ‘optimaal ‘ en worden ontslagen van verder follow-up.

Het kinderartsenonderzoek werd beschouwd als een diagnostische test in een populatie met een hoog a priori risico op ontwikkelingsstoornissen. Zo’n test zou een hoge sensitiviteit moeten hebben omdat geen kinderen mogen worden gemist. Een hoge specificiteit zou een te lang voortgezette controle moeten voorkomen, maar was van minder groot belang. De waarde van dit onderzoek werd bepaald door het kinderartsenonderzoek te vergelijken met de derde stap, een gouden standaard bestaande uit movement ABC, RAKIT en CBCL.

De vragenlijst ging in op de volgende domeinen: sociaal-economische status, algemene gezondheid, visuele en auditieve functies, motorische ontwikkeling, cognitieve ontwikkeling, spraaktaal en gedrag. De sociale status van het kind werd vastgesteld met behulp van vragen over de samenstelling van het gezin, het aantal broertjes en zusjes en vragen over beroep en scholing van beide ouders. De algemene gezondheid werd vastgesteld met behulp van vragen over de medische consumptie (het aantal bezoeken aan een polikliniek, het aantal ziekenhuisopnamen, het aantal operaties, het aantal diagnostische testen), somatische symptomen op het gebied van KNO, ademhaling, maagdarmsysteem en centrale zenuwstelsel, en met behulp van vragen over het gebruik van medicijnen. De visuele en auditieve functies werden vastgesteld met behulp van vragen over het gehoor met of zonder hulpmiddelen en over de visus met en zonder bril. De cognitieve ontwikkeling werd vastgesteld met behulp van vragen over de schoolprestaties, leerproblemen en de noodzaak van ‘remedial teaching’ of speciaal onderwijs. Gedragsproblemen werden vastgesteld met een korte gedragsvragenlijst, die bestond uit die 15 vragen uit de Child Behavior Checklist (CBCL) waarvan uit voorgaand onderzoek was gebleken dat deze het meest discrimineerden tussen normale kinderen en kinderen die waren verwezen naar psychologische of psychiatrische hulpdiensten in de algemene populatie.

Het onderzoek door de kinderarts startte met een controle op de volledigheid van de antwoorden op de enquête en een door de antwoorden gestuurde uitbreiding van de anamnese. Vervolgens werden de lengte, het gewicht, de schedelgroei en de bloeddruk gemeten. Hierna volgde een lichamelijk onderzoek.

Het neurologisch functioneren werd door de kinderarts vastgesteld met behulp van het onderzoek volgens Touwen. Dit onderzoek bestaat uit 40 items die betrekking hebben op de tonus, reflexen, spierkracht, onwillekeurige bewegingen, houding, balans, coördinatie en grove motorische vaardigheden. De kinderen werden geclassificeerd in de volgende drie categorieën: normaal (geen neurologische afwijking), minor neurological dysfunction (MND, afwijkingen zonder invloed op houding of het bewegen) en spasticiteit (cerebral palsy, waaronder werd verstaan: neurologische

afwijkingen met daarbij afwijkende houding of beweging).

De Denver OntwikkelingsSchalen (DOS; in het Amerikaans: Denver Development Screening Test, DDST) werden gebruikt om de neuromotorische en cognitieve functies vast te stellen. De DOS bestaat uit 105 testitems die zijn geclusterd in vier groepen: motorische functie, taal, adaptatie en sociaal gedrag. Elk testitem heeft een afbreekpunt op een leeftijd waarop 90% van de kinderen in staat is om dit item uit te voeren. Per leeftijdscategorie wordt een deel van de 105 testitems gebruikt. Elk van de vier domeinen word gescoord als vertraagd wanneer het kind twee of meer vertragingen ten opzichte van de '90% grens' laat zien, of één vertraging zonder compensatie in hetzelfde domein. Een kind wordt vervolgens geclassificeerd als 'at risk' wanneer één domein is vertraagd en als abnormaal als twee domeinen zijn vertraagd.

De taal werd onderzocht met een bij de leeftijd passende TaalScreeningsTest, ontwikkeld en gevalideerd door Gerritsen. Deze test bestaat uit 39 items, die betrekking hebben op het woordgebruik, begrip, geheugen, productie van taal, gebruik van meervoud en voorzetsels en op de verstaanbaarheid. Een normale score is ≤ 17, 'at risk' is tussen 18 en 25 en abnormaal is ≥ 25, tot een maximum score van 52.

Alle resultaten van het kinderartsenonderzoek werden vervolgens samengevat in een eindbeoordeling voor elk van de volgende vier domeinen: neuromotore ontwikkeling, cognitieve ontwikkeling, taalontwikkeling en gedrag, en in een uiteindelijke beoordeling 'optimaal' (alle domeinen normaal), 'at risk' (verder onderzoek nodig) of 'abnormaal' (behandeling nodig of reeds behandeld in een of meerdere domeinen).

*Validatie instrumenten*

De gecombineerde resultaten van het kinderartsenonderzoek werden gevalideerd tegen een gecombineerde set van gestandaardiseerde testen voor motorische ontwikkeling, cognitie, inclusief taalontwikkeling, en gedrag. Deze set bestond uit de movement Assessment Battery for Children (movement ABC), de gereviseerde Amsterdamse Kinder-IntelligentieTest (RAKIT) en de volledige Child Behavior Check List (CBCL).

Getrainde kinderfysiotherapeuten en kinderpsychologen voerden deze testen uit. Zij waren blind voor de bevindingen van het onderzoek van de kinderarts. De movement ABC duurde 45 minuten en de RAKIT één uur; de CBCL werd thuis ingevuld.

De movement ABC doet een uitspraak over het motorisch functioneren in het dagelijks leven. Hoe hoger de score, hoe slechter de prestatie. Totaalscores ≤ 10,5 (15$^{e}$ percentiel) worden beschouwd als normaal, van 11,0 tot 17,0 (5$^{e}$ percentiel) als 'at risk' en > 17,0 als abnormaal.

De RAKIT is genormeerd voor Nederlandse kinderen tussen vier en zeven jaar.

Kinderen met een score tussen -2 en -1 S.D. worden beschouwd als ‘at risk’, kinderen onder -2 S.D. als ‘abnormaal’.

De CBCL is een gestandaardiseerde maat voor gedrag van het kind, ontwikkeld door Achenbach en voor Nederlandse kinderen genormaliseerd door Verhulst. Totaalscores ≤ 59 worden beschouwd als normaal, van 60 tot en met 63 als overgangsgebied en ≥ 64 als gedragsstoornis.

Wanneer alle drie testen een normaal resultaat gaven, was de uiteindelijke conclusie ‘optimaal’, wanneer één of meer van de drie testen een uitslag ‘at risk’ gaf was de uiteindelijke conclusie ‘at risk’ en wanneer één of meer van de drie testen een uitslag abnormaal gaf was de uiteindelijke conclusie ‘abnormaal’.

*Populatie*

De studiepopulatie bestond uit kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram, behandeld in één van de drie participerende NICU’s en inmiddels vijf jaar oud. Vanwege logistieke redenen varieerden de inclusieperioden enigszins: van 1 oktober 1992 tot 15 juni 1994 in het Universitair Medisch Centrum St. Radboud, Nijmegen, van 15 november 1992 tot 1 januari 1994 in het Academisch Medisch Centrum, Amsterdam en van 1 januari 1993 tot 1 januari 1995 in Máxima Medisch Centrum, Veldhoven. Gedurende deze perioden werden 764 kinderen opgenomen op de drie NICU’s. Van hen stierven 131 (17%) vóór de leeftijd van vijf jaar. Ook werden 46 patiënten met een zwangerschapsduur van minder dan 30 weken en behandeld in het AMC van het vergelijkende onderzoek uitgesloten omdat zij deelnamen aan een andere studie. Eenentwintig kinderen met bekende ernstige spasticiteit, blindheid, ernstige mentale retardatie, chromosomale afwijkingen of een ernstige stofwisselingsstoornis werden uitgesloten omdat van tevoren duidelijk was dat deze kinderen de testen niet zouden kunnen volbrengen. Als gevolg hiervan bleven 566 kinderen beschikbaar voor de studie.

Van de 566 beschikbare kinderen werden 135 (23,9%) niet onderzocht wegens diverse redenen. Taalproblemen speelden een belangrijke rol. De non-respons was 65% bij kinderen van niet-Nederlands sprekende ouders en 19% van wel Nederlands sprekende gezinnen ($p < 0,01$). Van de 431 geteste kinderen volbrachten 395 kinderen alle onderdelen van het kinderartsenonderzoek, de movement ABC, de IQ-test en de CBCL.

De gemiddelde zwangerschapsduur (32,2 (S.D. 2,0) weken) en het gemiddeld geboortegewicht (1276 (S.D. 332) gram) lagen relatief hoog vergeleken met de meeste follow-upstudies uit de jaren ‘90. Dit geeft aan dat door het niet onderzoeken van een aantal kinderen onder de 30 weken uit één van de drie participerende ziekenhuizen een selectiebias was ontstaan in de richting van een relatief gunstige populatie. Er was geen

verschil in perinatale gegevens tussen onderzochte en niet-onderzochte kinderen met uitzondering van meerlingen (36% van de onderzochte versus 21% van de niet-onderzochte kinderen.

*Resultaten van het kinderartsenonderzoek (hoofdstuk 2)*

Bij 412 kinderen was een volledig kinderartsenonderzoek uitgevoerd. De vragenlijst gaf informatie over het individuele kind en bespaarde tijd in de volgende stap. Het kinderartsenonderzoek nam gemiddeld één uur in beslag, variërend van minder dan 50 minuten voor kinderen met een normale ontwikkeling tot meer dan 70 minuten voor kinderen met meerdere ontwikkelingsstoornissen. De meeste kinderen hadden plezier in de verschillende onderdelen van het onderzoek en de ouders waren zeer tevreden met het uitgebreide onderzoek. De kinderartsen hadden het idee dat het onderzoek hen prima in staat stelde de ontwikkeling van deze vijfjarige kinderen te schatten. De kinderartsen classificeerden 275 (67%) kinderen als 'optimaal', 58 (14%) als 'at risk' en 79 (19%) als 'abnormaal'. Van de 137 kinderen die waren geclassificeerd als 'at risk of abnormaal' waren er 82 (60%) van tevoren nog niet als zodanig geïdentificeerd bij eerdere follow-uponderzoeken.

Van de kinderen met afwijkende resultaten op een van de standaardtesten werden door het onderzoek slechts 48% opgemerkt. De meerderheid van hen had motorische problemen, gedragsproblemen of een combinatie van beide; een minderheid had cognitieve problemen.

Van de gestandaardiseerde testen classificeerde de movement ABC 56%, de IQ test 78% en de CBCL 77% als normaal. De gecombineerde resultaten van deze drie standaardonderzoeken gaven aan dat er significant meer kinderen met ontwikkelingsstoornissen waren dan door het kinderartsenonderzoek was vastgesteld. Wanneer de standaardtesten werden gecombineerd scoorden 167 kinderen (43,3%) 'optimaal' op alle testen, 134 (34,7%) op twee testen, 61 (15,8%) op één test en 24 (6,2%) op geen van de testen. De specificiteit (88%) en positief voorspellende waarde (85%) van het kinderartsenonderzoek waren adequaat, maar veel kinderen met een ontwikkelingsstoornis werden niet geïdentificeerd, hetgeen resulteerde in een sensitiviteit van 48% en een negatief voorspellende waarde van 55%.

Wij concludeerden dat het gestructureerde kinderartsenonderzoek niet voldeed als een eenvoudige diagnostische test van vijfjarige kinderen om onderscheid te maken tussen kinderen met een normale ontwikkeling en kinderen die verder moesten worden onderzocht. Follow-upstudies die niet bestaan uit gedetailleerde en gestandaardiseerde testen voor meerdere domeinen onderschatten daarom de ontwikkelingsproblemen bij kinderen geboren na een zwangerschapsduur < 32 weken en/of met een

geboortegewicht < 1500 gram. Deze onderschatting is er mogelijk ook de oorzaak van dat de incidentie van ontwikkelingsstoornissen met het toenemen van de leeftijd schijnbaar toeneemt. Er is blijkbaar geen simpele manier om kinderen die extra hulp nodig hebben te identificeren en er is geen simpele manier om de uitkomst van neonatale intensieve zorg te evalueren.

Omdat motorische stoornissen zowel het meest voorkwamen als het meest werden gemist, analyseerden wij vervolgens de relaties tussen het kinderartsenonderzoek, het onderzoek volgens Touwen, de DOS en de movement ABC verder in hoofdstuk 3. Wij analyseerden ook de relaties tussen het kinderartsenonderzoek, de DOS, de TaalScreeningsTest en de IQ test in hoofdstuk 4 en tussen de TaalScreeningsTest en formele taaltesten in hoofdstuk 5.

*Motorisch ontwikkeling (hoofdstuk 3)*

In het individuele kind is de functionele motorische capaciteit niet een geïsoleerde gebeurtenis, maar een keten van processen die uitgaat van fysieke vaardigheden en mogelijkheden en die wordt gestuurd door vele processen in het centrale zenuwstelsel. Wanneer wij bij kinderen doelgerichte motorische activiteiten bestuderen, kijken wij naar het resultaat van deze keten van processen. Daarom kunnen motorisch stoornissen zowel fysieke als neurologische, visuele, cognitieve of gedragsmatige oorzaken hebben. Het belangrijkste doel van een sceeningsonderzoek is niet om de relaties met deze domeinen te ontdekken, maar om vast te stellen of de kwaliteit en kwantiteit van de motorisch vaardigheid voldoende is voor de leeftijd. Wanneer een afwijkende score wordt gevonden is een meer gedetailleerd, multidisciplinair onderzoek nodig om de samenhang met bovengenoemde factoren te ontrafelen.

De meeste kinderartsen gebruiken een neurologisch onderzoek, dat de integriteit van het neurologische systeem vast stelt, om een motorisch ontwikkelingsprobleem te ontdekken. Wij vergeleken daarom het neurologisch onderzoek volgens Touwen met de movement ABC. De movement ABC geeft een indicatie van het motorisch functioneren in het dagelijkse leven. De movement ABC wordt wereldwijd veel gebruikt om stoornissen in de coördinatie van de motoriek vast te stellen bij kinderen met diverse ontwikkelingsproblemen, maar ook in follow-upstudies van te vroeg geboren kinderen.

Met behulp van de movement ABC vonden we in onze populatie veel te vroeg geboren kinderen zonder zeer ernstige handicaps ernstige motorische stoornissen bij 22.0% en lichte motorische stoornissen bij 21,5%. Dit is respectievelijk vier- en tweemaal de incidentie in een normale populatie. Diverse andere auteurs vonden

ernstige motorisch stoornissen in 20 - 30% van alle veel te vroeg geboren kinderen en 50 - 55% als ook de lichte motorische stoornissen werden meegenomen. De iets lagere incidentie die wij vonden kan worden verklaard doordat wij de 3% meest gehandicapte kinderen in onze studie hebben uitgesloten.

Vergelijking van het neurologisch onderzoek door de kinderarts met de movement ABC liet zien dat kinderartsen, zelfs als zij speciaal zijn getraind, niet goed in staat zijn motorische ontwikkelingsstoornissen te ontdekken. Kinderartsen zijn waarschijnlijk vooringenomen door hun kennis van de indrukwekkende perinatale problemen van hun patiënten en de opluchting dat de ontwikkeling toch beter verloopt dan zij hadden gedacht. Zij zijn misschien ook vooringenomen omdat zij zelden gezonde kinderen onderzoeken en daarom gemakkelijk lichte motorische stoornissen over het hoofd zien. Tenslotte kan het kinderartsenonderzoek subjectief zijn omdat het onderzoek geen rigide structuur kent voor het afnemen, de beschikbare tijd per item en het scoren en optellen van de resultaten. Wij konden de sensitiviteit van het onderzoek verbeteren van 0,19 naar 0,62 door het bestaande algoritme om te zetten in een gecomputeriseerd algoritme.

Vergelijking van het Touwen onderzoek, de DOS en de movement ABC laat zien deze onderzoeken verschillende aspecten van de motorische ontwikkeling testen. Het onderzoek volgens Touwen richt zich op de neurologische ontwikkeling, de DOS op de ontwikkeling in het algemeen en de movement ABC op de functionele motorisch ontwikkeling.

De movement ABC test het motorisch functioneren van alledag, hetgeen betekent dat het de vaardigheden test die kinderen op school nodig hebben. Marlow en medewerkers publiceerden in een longitudinale studie dat een slechte prestatie op de movement ABC op de leeftijd van zes jaar voorspellend was voor latere school problemen en gaven aan dat de lange-termijnontwikkeling bij deze kinderen mogelijk gevaar liep. Ook Diamond toonde aan dat er een nauwe relatie is tussen de motorische en verstandelijke ontwikkeling. Wij vonden dat 64% van de kinderen met een niet-optimale movement ABC al problemen had op school. Wij veronderstellen dat er bij deze kinderen concentratie- en aandachtsproblemen aanwezig zijn en dat de leercapaciteit is verminderd.

Omdat de movement ABC breed is geaccepteerd en relatief goedkoop is uit te voeren, adviseren wij de movement ABC te gebruiken voor het onderzoeken en vastleggen van de motorische functie van kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram op de leeftijd van vijf jaar.

Anderzijds vonden wij ook bij 13 kinderen met een optimale movement ABC een onmiskenbare spastische parese. Deze kinderen hadden blijkbaar geleerd om goed te functioneren, ondanks enkele beperkingen van hun neurologische systeem. Daardoor

konden zij voldoen aan de nauwkeurigheid en de snelheid die de movement ABC vereist. Om deze kinderen te detecteren moet een standaard neurologisch onderzoek deel uit maken van het follow-uponderzoek op de leeftijd van vijf jaar. Deze kinderen vormen in de praktijk geen probleem. Ze kunnen worden verwezen voor verder onderzoek, maar het is alleen nodig om hen te blijven vervolgen om te kijken of zij inderdaad nieuwe motorische taken even vaardig leren als hun leeftijdsgenoten. Voor de evaluatie van perinatale zorg is de detectie van deze spastische parese als marker van perinatale schade wel essentieel.

Voor vroegtijdige onderkenning van kinderen die extra hulp nodig hebben én voor evaluatie van de perinatale zorg moet zowel de movement ABC als een standaard neurologisch onderzoek deel uit maken van de systematische follow-up van kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram.

*Cognitie (Verstandelijke ontwikkeling; hoofdstuk 4)*

Dit deel van het onderzoek had als doel te evalueren of het met een onderzoek door de kinderarts mogelijk is om die te vroeg geboren kinderen te identificeren die op de leeftijd van vijf jaar moeten worden verwezen naar een kinderpsycholoog voor een uitgebreide analyse van de cognitie. Wij vergeleken het oordeel van een kinderarts, gebaseerd op een vragenlijst voor de ouders, de indruk die het kind maakte tijdens het onderzoek, de DOS en de TaalScreeningsTest met de resultaten van de geReviseeerde Amsterdamse KinderIntelligentie test (RAKIT) bij 368 kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram. De sensitiviteit van het oordeel van de kinderarts over de cognitie (sensitiviteit 0,34) en van de DOS onderdelen (sensitiviteit 0,18 - 0,36) was laag en gaf een onderschatting van de werkelijke behoeft aan verwijzing voor verder uitgebreid onderzoek. De complete DOS (sensitiviteit 0,72) en schoolprestaties voldeden nog het best.

Kinderartsen gebruiken vaak de taalontwikkeling als een indicator van de verstandelijke vermogens, maar de sensitiviteit van de TaalScreeningsTest was laag (sensitiviteit 0,55). Uit de in hoofdstuk 5 beschreven studie blijkt dat de voorspellende waarde van de TaalScreeningsTest voor de detectie van spraaktaalstoornissen kan worden verbeterd door de grenswaarde voor het onderscheid tussen normale en niet-normale spraaktaalontwikkeling te verlagen van 18 naar 17. Wij hebben niet geanalyseerd of dat ook geldt voor de beoordeling van de cognitie.

Schoolprestaties zijn blijkbaar een goede indicator voor het verstandelijk functioneren. De beslissing om te vroeg geboren kinderen te verwijzen voor verdere,

uitgebreide evaluatie zou kunnen worden gebaseerd op de schoolprestaties in groep 2 van het basisonderwijs. Deze informatie is gemakkelijk en zonder kosten te verkrijgen. Als echter het doel is om schoolproblemen vroegtijdig te bestrijden of zelfs te voorkomen is het te laat om een kind pas te verwijzen als het al problemen heeft. Omdat 37% van deze kinderen al extra begeleiding of remedial teaching hebben of zijn blijven zitten, raden wij aan om uiterlijk op vijfjarige leeftijd al deze kinderen te testen met een formele intelligentietest.

Schoolproblemen bij kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram moeten serieus worden genomen. Deze kinderen moeten worden verwezen voor een uitgebreide intelligentietest. Gezien de hoge incidentie van leerproblemen is het nog beter om niet te wachten totdat deze zich voordoen, maar een gevalideerde intelligentietest deel uit te laten maken van de systematische follow-up van kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram op vijfjarige leeftijd.

De gegevens uit dit hoofdstuk werden ook gebruikt om het IQ op de leeftijd van vijf jaar te beschrijven van een populatie kinderen geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram. Wij maten het IQ in 454 van de 566 beschikbare kinderen. Het gemiddelde IQ was 96,6 (S.D. 16,4); 78% van de kinderen had een IQ in de normale range, 16% tussen -1 en -2 S.D. en 6% onder -2 S.D.. De kinderen van de thyroxine supplementatie studie, uitgesloten in hoofdstuk 4, werden meegenomen. Wij hadden 21 kinderen van onderzoek uitgesloten vanwege ernstige handicaps. Als we aannemen dat deze 21 een IQ < -2 S.D. hadden, dan had 75% van de kinderen een normaal IQ, 15% tussen -1 en -2 S.D. en 10% onder -2 S.D.. Diverse studies vinden bij kinderen met een geboortegewicht < 1500 gram een gemiddeld IQ dat 7 - 15 punten (0,5 to 1 S.D.) lager is dan van controlekinderen en bij kinderen met een geboortegewicht < 1500 gram zonder grote neurologische stoornissen een gemiddeld IQ dat 7 - 10 punten lager is dan controlekinderen. De meeste studies vinden IQ's < -2 S.D. bij 10 - 20% van alle kinderen met een geboortegewicht < 1500 gram, vergeleken met verwachte frequentie van 2,3% in een normale populatie.

Diverse auteurs hebben laten zien dat non-responders minder gunstige uitkomstresultaten hebben dan responders. De respons in ons onderzoek was 80% van de beschikbare kinderen. De perinatale factoren verschilden niet tussen de onderzochte en niet-onderzochte kinderen, met uitzondering van een lagere incidentie van meerlingen in de niet-onderzochte groep. Door de selectiebias en de non-respons wordt in onze studie het gemiddelde IQ misschien iets te hoog geschat, maar overigens is deze schatting in overeenstemming met de gegevens uit de literatuur.

IQ -metingen geven vaak een iets te hoog IQ weer door het Flynn-effect. Flynn heeft aangetoond dat het IQ ongeveer 0,3 punten per jaar toeneemt. De RAKIT was gestandaardiseerd in 1984. Met een jaarlijkse toename van 0.3 punten zou het gemiddelde IQ in de huidige populatie wel eens 105 kunnen zijn in plaats van 100. Bovendien geven IQ testen bij kinderen met concentratiestoornissen mogelijk een overschatting van de cognitie in de dagelijkse situatie, omdat de taken die de kinderen tijdens de test moeten uitvoeren gestructureerd en ééndimensionaal zijn en worden afgenomen in een één op één situatie met de tester, zonder storingen van buiten af.

*Spraak en taal*

Spraak en taal zijn belangrijke gebieden van de kinderlijke ontwikkeling. Daarom was een taalscreeningstest opgenomen in het kinderartsenonderzoek. Om deze taalscreeningstest te valideren in onze groep te vroeg geboren kinderen werden de resultaten van de taalscreeningstest afgenomen door de kinderarts vergeleken met de resultaten van een gestandaardiseerde volledige spraaktaaltest, afgenomen door een professionele spraaktaaltherapeute. Het doel was te onderzoeken hoe efficiënt en nauwkeurig de taalscreeningstest vaststelt welke te vroeg geboren kinderen een taalachterstand hebben. Op grond van het aantal fouten in de test werden de kinderen geclassificeerd in de groepen ‘normaal’ (≤ 17 fouten), ‘licht vertraagd’ (18 - 25 fouten) of ‘ernstig vertraagd’ (≥ 26 fouten). Wanneer licht en ernstig vertraagd werden samengenomen (≥ 18 fouten), was de sensitiviteit 62% en de specificiteit 95%. De sensitiviteit kon worden verhoogd tot 79% wanneer ≥ 17 fouten als afwijkend werd beschouwd; de specificiteit daalde daarbij tot 94%. De specificiteit kon vervolgens worden verbeterd door enkele vragen uit de vragenlijst in het oordeel te betrekken.

De taalscreeningstest is een uitstekend screeningsinstrument om spraaktaalstoornissen te ontdekken bij vijfjarige kinderen die zijn geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram, onder voorwaarde dat het afkappunt voor niet-normaal wordt verlaagd van ≥ 18 naar ≥ 17 fouten. We adviseren deze screeningstest op te nemen in de follow-up van deze kinderen op de leeftijd van vijf jaar.

## Deel II. Vergelijking van perinatale zorg in 1983 en 1993 en de gevolgen op korte (hoofdstuk 6) en lange (hoofdstuk 7) termijn

Wij vergeleken twee geografisch bepaalde populaties met een tijdsverloop van tien jaar. Door socio-economische veranderingen steeg in deze periode de moederlijke leeftijd. Deze hogere leeftijd leidde, gedeeltelijk via medische stimulering van de vruchtbaarheid, tot een toename van de frequentie meerlingen en vroeggeboorten.
Ondanks deze ongunstige factoren vonden wij een daling van de sterfte in elke zwangerschapsweek van ongeveer 40%. Richardson en medewerkers hebben gesuggereerd dat eenderde van deze verbetering het gevolg is van een betere conditie op het moment van opname op de afdeling neonatologie en tweederde het gevolg van effectievere neonatale zorg. Onze bevindingen suggereren dat niet één specifieke factor, maar veeleer verbeteringen van talloze aspecten van obstetrische en neonatale zorg, inclusief prenataal transport, vroegtijdige sectio Caesarea, systematische toediening van steroïden voor de geboorte en vooral het gebruik van surfactant deze verbetering tot stand heeft gebracht.

De verbeterde overleving ging niet gepaard met een daling van de incidentie (berekend over de overlevenden) van intraventriculaire en periventriculaire bloedingen; de incidentie van bronchopulmonaire dysplasie nam zelfs toe.

Bij het bestuderen van de lange-termijn-uitkomst in deze twee populaties vonden wij bij de overlevenden op de leeftijd van vijf jaar een toename van ernstige cerebral palsy van 6,3% naar 12,3%. Deze bevinding is alarmerend. Hagberg vond in Zweden een toename van de incidentie van cerebral palsy in de jaren ’70 en suggereerde een relatie met de sterk toegenomen overleving van kinderen geboren na een zwangerschapsduur < 37 weken. Het lijkt erop dat de geschiedenis zich nu herhaalt in de groep kinderen geboren na een zwangerschapsduur < 32 weken. In de multivariate analyse was het risico op spasticiteit in 1983 verhoogd bij jongetjes en in 1993 na lage Apgarscores, na intraventriculaire en periventriculaire bloedingen en na gebruik van dexamethason voor bronchopulmonaire dysplasie. Verdere analyse van onze data laat zien dat dit effect vooral, maar niet uitsluitend, optrad in de kinderen geboren na een zwangerschapsduur < 28 weken.

Wij beschikken niet over gegevens over het IQ van kinderen geboren in 1983 en kunnen daarom geen uitspraak doen over verschillen in cognitie. Er waren evenwel geen verschillen in schoolresultaten tussen kinderen geboren in 1983 en 1993. Respectievelijk 11,9% en 9,8% (verschil niet significant) volgden speciaal onderwijs. De incidentie van gedragsproblemen was in beide cohorten veel hoger dan de frequentie in een normale populatie, maar veranderde niet.

Omdat de frequentie van stoornissen in motoriek, cognitie en gedrag in de

overlevende kinderen gelijk bleef en het percentage overlevende kinderen is toegenomen terwijl ook de incidentie van vroeggeboorte is toegenomen, is het aantal kinderen met problemen op het gebied van motoriek, cognitie en gedrag sterk toegenomen. We moeten er rekening mee houden dat de behoefte aan voorzieningen voor deze kinderen zal toenemen.

De incidentie van visuele problemen en scheelzien nam duidelijk af. Wij nemen aan dat dit komt door de introductie van continue, transcutane monitoring van de zuurstofsaturatie in het bloed.

## Conclusie

Tegelijk met de verbetering van de overleving van kinderen geboren na een zeer korte zwangerschapsduur (< 32 weken) en/of met een zeer laag geboortegewicht (< 1500 gram) nam het besef toe dat niet al deze kinderen overleven zonder nadelige gevolgen. De verwachting was dat dergelijke problemen tijdelijke bijwerkingen waren en zouden verbeteren met het toenemen van kennis, kunde en ervaring in de neonatale zorg. De gegevens uit dit proefschrift laten zien dat de frequentie van stoornissen en handicaps onveranderd hoog blijft, met de gunstige uitzondering van visusproblemen.

Factoren die verantwoordelijk zijn voor een afname van de sterfte zoals antenatale steroïden, kunstmatige beademing en surfactant zorgen niet voor een daling van de incidentie van ontwikkelingsproblemen in de kinderen die overleven. Het gebruik van dexamethason na de geboorte is een ernstige extra risicofactor voor het ontstaan van ontwikkelingsstoornissen.

Ernstige ontwikkelingsstoornissen worden meestal al op de leeftijd van twee jaar gediagnosticeerd. Lichte motorische stoornissen, leerproblemen en gedragsproblemen die interfereren met het verwerven van alledaagse vaardigheden en het normale leren. worden vaak pas op de schoolleeftijd onderkend. Wanneer met het toenemen van de leeftijd de eisen die de maatschappij stelt aan het individu toenemen, nemen de problemen nog verder toe. Geïsoleerd kunnen deze problemen nog wel meevallen, maar in combinatie en zonder compensatie in andere domeinen vormen ze een levenslange belasting. Daarom is langdurige follow-up en begeleiding van deze kinderen nodig.

Kinderartsen die kinderen, geboren na een zwangerschapsduur < 32 weken en/of met een geboortegewicht < 1500 gram, onderzoeken met technieken die berusten op het onderzoek volgens Touwen, de Denver OntwikkelingsSchalen (DOS) en een verkorte CBCL zien een aantal kinderen met motorische ontwikkelingsstoornissen, leer- en gedragsstoornissen over het hoofd. Omdat de movement ABC een nauwkeurig en wijd verbreid instrument is om motorische ontwikkelingsstoornissen op te sporen moet de

movement ABC op de leeftijd van vijf jaar opgenomen worden in het follow-uponderzoek van kinderen die zijn geboren na een zwangerschapsduur < 32 weken of met een geboortegewicht < 1500 gram. Om alle neurologische problemen op te sporen is ook een gestandaardiseerd neurologisch onderzoek nodig. Leerproblemen op school zijn in deze groep kinderen een alarmteken van cognitieve problemen en vormen een indicatie voor verwijzing van het betreffende kind naar een psycholoog voor uitgebreid onderzoek. Het schooljaar overdoen is geen goede oplossing. Gezien de hoge incidentie van leerproblemen is het zelfs aan te bevelen om de schoolproblemen voor te zijn en een gestandaardiseerd onderzoek van de cognitie van alle kinderen geboren na een zwangerschapsduur < 32 weken of met een geboortegewicht < 1500 gram op te nemen in het follow-uponderzoek op vijfjarige leeftijd (vóór school groep drie).

Omdat meer dan de helft van alle kinderen geboren na een zwangerschapsduur < 32 weken of met een geboortegewicht < 1500 gram ontwikkelingsproblemen heeft, is routine onderzoek van al deze kinderen noodzakelijk. De te gebruiken testen moeten alle domeinen van de ontwikkeling onderzoeken, moeten een rigide structuur hebben zowel wat betreft het uitvoeren van de test als van het berekenen van de testresultaten. Onderzoek door een kinderarts die zich in een normale poliklinische setting een algemeen oordeel vormt is onvoldoende.

Follow-up studies die niet dergelijke gedetailleerde testen van meerdere domeinen bevatten overschatten de uitkomst van perinatale intensive care. Er is geen simpele manier om kinderen te identificeren die extra zorg nodig hebben en er is geen simpele manier om neonatale intensive care goed te evalueren.

Het in dit proefschrift beschreven onderzoek hield zich bezig met de problemen die vaak ontstaan bij kinderen geboren na een zwangerschapsduur < 32 weken of met een geboortegewicht < 1500 gram. Het liet evenwel ook zien dat moderne intensive care het leven redt van veel kinderen die veel te vroeg of veel te klein zijn geboren en dat velen van hen opgroeien zonder enig probleem. Hoewel moderne neonatale zorg niet in staat is alle te vroeg of te klein geboren kinderen intact in leven te houden is zij zeker de moeite waard.

## Dankbetuiging

De Validatie Studiegroep bestaat uit: dr. A.L. den Ouden en H. Kloosterboer-Boerrigter (TNO Preventie en Gezondheid, Leiden), prof. dr. L.A.A. Kollée, dr. M.W.G. Nijhuis-van der Sanden, M. Sondaar, S. Knuijt en B.J.M. van Kessel-Feddema (Universitair Medisch Centrum St Radboud, Nijmegen), dr. A. Ilsen, dr. A.G. van Wassenaer, R. Breur-Pieterse, dr. J.M. Briët en K. Koldewijn (Academisch Medisch

Centrum, Amsterdam), M.J.K. de Kleine, J. Buijs en M.J. Wit (Máxima Medisch Centrum, Veldhoven) en prof. dr. A.L. van Baar, Universiteit van Tilburg. Dr. R Brand, (Leids Universitair Medisch Centrum) adviseert over de statistiek en prof. dr. S.P. Verloove-Vanhorick (TNO Preventie en Gezondheid, Leiden en Leids Universitair Medisch Centrum ) begeleidt de studie. De studie werd gesteund door een subsidie van het Praeventiefonds, projectnummer 28-2756 en ZorgOnderzoek Nederland (ZonMw), projectnummer 10010004-20.

## Curriculum vitae

Martin Jan Koop de Kleine, geboren 18 september 1948 te Meppel, behaalde in 1966 het eindexamen gymnasium aan het Menso Alting College te Hoogeveen. Hij studeerde geneeskunde aan de Vrije Universiteit te Amsterdam van september 1966 tot december 1973. Van 1974 - 1978 specialiseerde hij zich tot kinderarts en van 1978 - 1980 tot neonatoloog aan het Academisch Ziekenhuis van de Universiteit van Amsterdam.
Hij was van 1980 tot 1989 als kinderarts-neonatoloog en wetenschappelijk hoofdmedewerker werkzaam in het Academisch Ziekenhuis van de Universiteit van Amsterdam, later het Academisch Medisch Centrum.
Vanaf de oprichting in 1989 tot nu toe is hij verbonden aan de afdeling neonatologie van Máxima Medisch Centrum, Veldhoven. Van 2000 tot 2004 was hij wetenschappelijk gastmedewerker van TNO Preventie en Gezondheid, Leiden.
Zijn aandachtsgebieden zijn neonatale epidemiologie & follow-up, pulmonologie en ethiek.

## Nawoord

De Validatiestudie was niet mogelijk geweest zonder de medewerking van zeer vele kinderen en hun ouders. Zelf hebben zij er gedeeltelijk profijt van gehad, maar het onderzoek zal voornamelijk ten goede komen aan de kinderen die na hen werden en worden geboren. Ik wil hen allen heel hartelijk danken voor hun deelname.

Het gereedkomen van dit proefschrift is een goede gelegenheid de velen te bedanken die bij de Validatiestudie waren betrokken:

De deelnemers aan het Validatieproject: Ina Kloosterboer-Boerigter, Lya den Ouden, Anneloes van Baar, Judy Briët, Rinske Breur-Pieterse, Aleid van Wassenaer, Karin Koldewijn, Louis Kollée, Ria Nijhuis, Meta Sondaar, Boudien van Kessel-Feddema, Simone Knuijt, Jan Buijs en Marion Wit. Het Validatieproject speelde zich af op de afdelingen neonatologie, psychologie en fysiotherapie van AMC, UMC St Radboud en Máxima Medisch Centrum en het instituut TNO Preventie en Gezondheid. De inbreng van zoveel kennis en inzet uit verschillende professies maakten het werken aan dit Validatieproject heel stimulerend. Van studiegroep is het uitgegroeid tot een familie met gemeenschappelijke idealen. Voor Anneloes een speciaal woord van dank omdat jij mij hebt overgehaald de bewerking van dit onderzoek op me te nemen.

Astrid Ehrnreich, Lili Jung, Brigit Hamers and Elle van de Bogaart dank ik wil voor de hulp bij het psychologisch onderzoek van de kinderen in de studie.

Tonnie Kelderman (AMC), Corrie van Wolferen (UMC St Radboud) en Nans Wouters (MMC) dank ik voor het leggen van de contacten met de ouders en verzorgen van de dossiers.

Henk en Lars Dinnissen dank ik voor de prachtige lay-out van dit proefschrift.

Collega onderzoekers en projectmedewerkers bij TNO preventie en Gezondheid, Marianne Amelink, Sabine Anthony, Simone Buitendijk, Caroline Dorrepaal, Kathy Herschderfer, Elysée Hille, Karin van der Pal, Ko van Wouwe, Jolanda Anthonissen, Ineke van Kempen, Tineke de Graaf, Tineke de Groot en Ina Kloosterboer-Boerrigter. Het was voor mij een bijzondere ervaring om een tijdlang 'jonge' onderzoeker te zijn in een befaamd onderzoeksinstituut. Ik dank hen hartelijk voor de gastvrijheid, de belangstelling, de praktische adviezen, de tips voor SPSS syntaxen en Reference Manager en de gesprekken bij de koffie wanneer het netwerk weer eens uitviel. Een niet-perfect werkend ICT systeem stimuleert de contacten en verbetert de kwaliteit van de wetenschap. Hans Vogelaar dank ik voor de zorgvuldige wijze waarop hij de data heeft beheerd.

Dr. Ronald Brand. Twintig jaar geleden mocht ik tijdens een cursus statistiek al genieten van zijn enthousiasme en vaardigheid om van statistiek iets logisch te maken. Zonder logica is zelfs statistiek niet eenvoudig. Ik dank hem voor het geduld om het twintig jaar later allemaal nog weer eens uit te leggen.

Collega's neonatologen en kinderartsen van Máxima Medisch Centrum. Jan Buijs en André Koolen, collega's van het eerste uur. Samen met de helaas veel te vroeg overleden Nedzad Beganovic hebben wij de unieke mogelijkheid gekregen om in Nederland een nieuwe neonatale intensive care op te richten. Met aanvankelijk weinig mensen werden relatief veel kinderen behandeld en vanuit verschillende meningen werd aan het 'Veldhovense beleid' vorm gegeven. Achteraf zijn de zwaarste jaren de mooiste geweest. De 'jongere' collega's, Peter Andriessen, Sidarto Bambang Oetomo, Vinus Bok, René van Gent, Anke Hukkelhoven-IJzermans, Luuk Janssens, Jacqueline van der Sluis-Bens, Walter Tjon A Ten, Janneke van Wijk, Marjo Verreussel, Diana Visser, Clarry Waals-van der Wal en Bas Zegers, dank ik voor de dagelijkse samenwerking in één grote vakgroep Kindergeneeskunde en de mogelijkheid om in de afgelopen jaren in de luwte aandacht aan dit proefschrift te kunnen geven. Ondanks de toenemende grootte van de vakgroep werken we in goede harmonie aan ons ideaal: een betere wereld voor zieke kinderen.

Verpleegkundigen van de afdeling neonatologie, die zonder uitzondering zijn geïnteresseerd in hoe het verder gaat met de te vroeg geboren of zieke pasgeborenen die zij zo liefdevol behandelen. Dit proefschrift geeft een aantal antwoorden op hun vragen, zet een aantal kanttekeningen bij het steeds verder naar beneden verschuiven van de ondergrens, maar laat vooral ook zien dat hun werk de moeite waard is.

Je jeugd zet een onuitwisbaar stempel op je. Mijn ouders, beiden al lange tijd geleden overleden, dank ik voor een gelukkige jeugd, voor de eerbied voor God en de medemens en voor het voorbeeld dat de wereld alleen beter wordt als mensen zich beter gedragen. Talenten zijn er om gebruikt te worden en als je zeven jaren hebt gestudeerd, moet je ook wel een boekje kunnen schrijven, vond mijn vader. Het heeft iets langer geduurd.

Mijn kinderen Ruben, Merel en Matthijs. Voor een kinderarts is er geen mooier geschenk mogelijk dan de geboorte van drie gezonde kinderen die in harmonie opgroeien. Elke dag ben ik trots en dankbaar om hun bestaan.

Mijn lieve vrouw Beatrijs ben ik dankbaar voor haar steun in de afgelopen jaren. Zij is het beste dat mij is overkomen:

*Een sterke vrouw, wie zal haar vinden?*
*Zij is meer waard dan edelstenen.*
*Haar man vertrouwt op haar*
*En zal daar rijkelijk bij winnen.*
*Ze brengt hem voorspoed*
*alle dagen van haar leven.*
*Haar kinderen prijzen haar,*
*haar man bejubelt haar:*
*"Er zijn veel sterke vrouwen,*
*maar jij overtreft ze allemaal".*
Spreuken van Salomo, hoofdstuk 31.