Stability of neurodevelopmental trajectories in moderately late and early preterm children born 15 years apart

van Dokkum, Nienke H; Lepe, Alexander; van Buuren, Stef; Reijneveld, Sijmen A; de Kroon, Marlou L A

Published in:
Pediatric Research

DOI:
10.1038/s41390-024-03188-2

European Society for Paediatric Research
Society for Pediatric Research
www.nature.com/pr

POPULATION STUDY ARTICLE

Check for updates

Stability of neurodevelopmental trajectories in moderately late and early preterm children born 15 years apart

Nienke H. van Dokkum[1,2]✉, Alexander Lepe[2], Stef van Buuren[3,4], Sijmen A. Reijneveld[2] and Marlou L. A. de Kroon[2,5]

BACKGROUND: Neurodevelopmental trajectories of preterm children may have changed due to changes in care and in society. We aimed to compare neurodevelopmental trajectories in early and moderately late preterm children, measured using the Developmental (D)-score, in two cohorts born 15 years apart.

METHODS: We included early preterm and moderately late preterm children from two Dutch cohorts (LOLLIPOP, 2002–2003 and ePREM, 2016–2017). ePREM counterparts were matched to LOLLIPOP participants by gestational age and sex. D-score trajectories were summarized by a multilevel model with random intercepts and random slopes, and multigroup analyses were used to test if the intercepts and slopes differed across cohorts.

RESULTS: We included 1686 preterm children (1071 moderately late preterm, 615 early preterm) from LOLLIPOP, and matched these with 1686 ePREM counterparts. The neurodevelopmental trajectories of the two cohorts were mostly similar. For early preterm children, we found no statistically significant differences. For moderately late preterm children, both the intercept (43.0 vs. 42.3, $p < 0.001$) and slope (23.5 vs. 23.9, $p = 0.002$) showed some, but only clinically minor, differences.

CONCLUSION: Developmental trajectories, measured using the D-score, in the first four years of life are comparable and stable across a period of 15 years for both early and moderately late preterm children.

Pediatric Research; https://doi.org/10.1038/s41390-024-03188-2

IMPACT:

- Neurodevelopmental trajectories are similar for early and moderately late preterm children born 15 years apart and thus seem quite stable in time.
- The validated Developmental score visualizes these trajectories based on developmental milestone attainment
- Because of its stability over time, the Developmental score trajectory may aid clinicians in neurodevelopmental assessment of preterm children as this simplifies monitoring and interpretation, similar to a growth chart.

INTRODUCTION

Worldwide, approximately every one in ten babies is born preterm, i.e., before 37 weeks of gestation.[1] Over the last decades, the incidence of preterm birth has increased.[1] Survival rates of preterm infants, particularly the youngest preterm infants at 22–25 weeks, are also rising.[2] With more preterm infants surviving, the focus shifted towards neurodevelopmental impairments following preterm birth. These youngest preterm infants are at higher risk for cerebral palsy, developmental delays, neurosensory impairment, emotional and behavioral problems, and attention problems.[2–6] moderately late preterm infants, born between 32 and 36 weeks of gestation, also exhibit neurodevelopmental problems, albeit less severe compared to the youngest preterm infants.[7–11] This non-optimal early life neurodevelopment may lead to problems in school participation.

Early life monitoring of neurodevelopment is of high importance because studies have shown that early interventions may improve neurodevelopmental outcomes in preterm children.[12–14] Monitoring of neurodevelopment in early life is mostly done by using tests or screening instruments that rely on developmental milestone assessment, either by direct observation by healthcare professionals or by parental reports. Examples of such tests are the Denver Developmental Screener, a widely used developmental test performed by professionals,[15] and the Ages and Stages Questionnaire, a globally used parental questionnaire.[16] The underlying developmental milestones are, on average, achieved in a consistent and ordinal pattern during the first few years of life.[17] However, the attainment of one developmental milestone by itself does not provide information on over-time neurodevelopmental trajectories. As preterm children may 'grow into deficit' over time,[18,19] the

[1]Department of Pediatrics, Division of Neonatology, Beatrix Children's Hospital, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Groningen 9713 GZ, the Netherlands. [2]Department of Health Sciences, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Groningen 9713 GZ, the Netherlands. [3]Netherlands Organization for Applied Scientific Research TNO, Schipholweg 77, Leiden 2316 ZL, The Netherlands. [4]Department of Methodology and Statistics, Utrecht University, Padualaan 14, Utrecht 3584 CH, The Netherlands. [5]Department of Public Health and Primary Care, Environment and Health, Youth Health Care, University of Leuven, KU Leuven, Leuven, Belgium. ✉email: n.h.van.dokkum@umcg.nl

Received: 22 June 2023 Revised: 24 February 2024 Accepted: 24 March 2024
Published online: 16 April 2024

deviation of their neurodevelopmental trajectory compared to the standard, i.e., similar to a growth chart, may integrate developmental milestone assessment over time and visually aid clinicians in interpretation.

The Developmental (D-)score has been developed as such an integrative measure.[20] This score constitutes a continuous interval-scale measure that is based on developmental milestones assessment. It leads to a developmental chart that is similar to a growth chart, and that facilitates monitoring of neurodevelopmental trajectories over time. The D-score has been extensively validated in several settings with varying developmental milestone assessment instruments as the basis, including the Dutch Developmental Instrument (DDI) or so-called 'Van Wiechen Onderzoek' that is used by preventive child healthcare in the Netherlands to assess developmental milestones for all Dutch children.[17] Additionally, the D-score can be used to compare cohorts of children born in different periods.

The currently available evidence mostly compares cohorts of extremely preterm children at specific points in time and, therefore, lacks trajectories of neurodevelopment, as well as the inclusion of moderately late preterm children. Assessment of neurodevelopmental trajectories of preterm children in consecutive birth cohorts may help to determine whether neurodevelopmental outcomes of preterm-born children have changed. Therefore, the aim of this study was to compare neurodevelopmental trajectories in early and moderately late preterm children, measured using the Developmental (D)-score, in two cohorts born 15 years apart.

METHODS

Setting and population

This study used data from two Dutch cohorts that include both preterm children born before 32 weeks gestation (early preterm children) as well as preterm children born between 32- and 36 weeks' gestation (moderately late preterm children), born 15 years apart.

The first cohort is the Longitudinal Preterm Outcome Project (LOLLIPOP). This community-based cohort consists of children born in 2002–2003. A total of 45,446 files were checked in preventive child healthcare centers across the Netherlands, and every child with a gestational age below 36 weeks was selected. From five neonatal intensive care units, preterm children born before 32 weeks, born in 2003, and alive upon discharge, were additionally approached to participate to enrich the cohort. Data on neurodevelopment was collected from preventive child healthcare centers at the final visit at 4 years of age. Children with major congenital malformations and syndromes were excluded, and eventually 1145 preterm children born moderately late preterm and 698 preterm children born early preterm participated in LOLLIPOP. The LOLLIPOP study was approved by the Institutional Review Board of the University Medical Center Groningen (METc 2005/130), and written informed consent was provided by all parents. From the LOLLIPOP cohort, we extracted the participants with data on developmental milestones. A full description of the LOLLIPOP cohort was previously provided by Kerstjens et al.[11]

The second cohort is the 'e-health preterm' cohort (ePREM). In this cohort, two preventive child healthcare centers in the Netherlands, one localized in the North and one in the southwest of the country, participated in the development and implementation of two e-health tools for preterm children. The e-health tools focused on implementing knowledge on preterm growth and development, to improve quality of care. Growth and neurodevelopmental data were obtained from a group of preterm infants before implementation and after implementation of the e-health tools. For the current study, data on developmental milestone attainment of the group of preterm infants assessed before the implementation of the e-health tools was used. For this cohort, parental informed consent for the use of registered data was obtained at the start of routine care.

Both the LOLLIPOP and ePREM cohorts were based on samples from the community-based clinical practice of preventive child healthcare. In the Netherlands, preventive child healthcare is attended by over 95% of children and their parents. Consultations are regularly scheduled throughout the first 4 years of life and entail growth and neurodevelopmental monitoring, providing parenting advice, and providing vaccinations according to the national program. Neurodevelopment is monitored using the DDI.[17] This instrument consists of 75 developmental milestones that are grouped together so that 90% of typically developing full-term infants attain them at the target age at which these are assessed. Several consultations over the first 4 years take place, in which a selection of these 75 milestones is assessed. On average, in each domain, two to five milestones are assessed. These developmental milestones are grouped into three domains: (1) fine motor, personal social, and adaptation, (2) gross motor, and (3) communication. For most of these milestones, clinicians are expected to observe the skills of the infants during a consultation. For several milestones in the communication domain, however, the parental report is considered to suffice if a child does not show the skill during the consultation. In clinical practice, all preterm children are invited to consultations at calendar ages, but prematurity is considered in the interpretation of attainment ages of the milestones. In our study, we chose to both use chronological and corrected ages throughout the full trajectory depicted (i.e., from birth to 4 years of age); corrected ages account for the difference between term age and gestational age.

Data handling and statistical analyses

ePREM counterparts were matched with replacement to the LOLLIPOP participants, based on gestational age and sex, on a one-to-one ratio. We chose to do so because the LOLLIPOP cohort oversampled preterm children, especially early preterm children, and the general population of the ePREM cohort did not include enough early preterm children to match on a one-to-two ratio, even though it was much larger.

D-scores are calculated across all milestones combined, without distinguishing separate domains. The D-scores were calculated as follows. First, missing data that were reliably obtained by cross-checking similar developmental milestones based on level of difficulty were added to the sample. For example, if a child was able to build a tower of six cubes and data for building a tower of three cubes were missing, we could reliably assume that this child would be able to also build a tower of three cubes and assess the missing milestone to be attained. Similarly, if a child was not able to build a tower of three cubes and data for building a tower of six cubes were missing, we could reliably assume that this child would not be able to build a tower of six cubes and assess the missing milestone to be not attained. This method was based on expert consensus within the research team. While the general recommendation is to calculate the D-score from only the realized responses, we decided to apply this method here to increase the stability and robustness of the D-score estimates in light of the very high (90 percent) passing probability in the standard test administration used in the Dutch preventive child healthcare. Using this method, we were able to reliably obtain 126 of 616 missing values on a total of 27,117 observations in 3372 children.

Next, the D-score was calculated using the D-score package version 1.4.0 and the "gcdg" key (https://cran.r-project.org/web/packages/dscore/index.html) in R version 4.0.2.[21] Before analyzing the data, age was log-transformed to approximate a linear relationship with the D-score. Then, in both cohorts, D-score trajectories were estimated using a multilevel model with both random intercepts (i.e., the initial level at ~9 months of age due to the log transformation) and random slopes (i.e., change per one unit increase in age in the log scale) using M-Plus version 8.3,[22] which summarizes each individual trajectory by a subject-specific intercept and slope. Due to the log transformation of age, which included the addition of the constant 0.27, the intercept is the initial level at ~9 months of age, i.e., at the centers of the slopes. For all models, we used full information maximum likelihood estimation with robust standard errors to account for missing data, which we assumed to be missing at random. To compare the developmental trajectories across eras, we tested if the intercepts and slopes differed across LOLLIPOP and ePREM using multigroup analyses. This comparison was made across the complete preterm group from both cohorts, as well as separately for moderately late and early preterm children. Finally, intercepts and slopes were estimated after adjusting for maternal educational level.

We repeated all analyses with a D-score corrected for gestational age. We chose to include the corrected-age-based D-score analyses, because correcting for gestational age is common practice in many countries when assessing preterm neurodevelopment, but it is not common practice in Dutch preventive child healthcare.

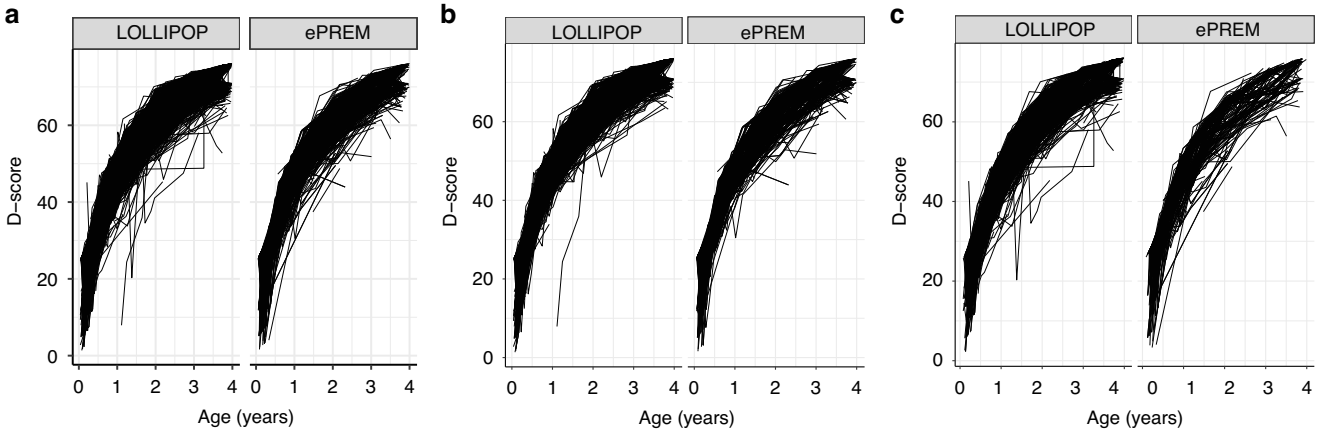

Fig. 1 Distribution of calendar-age-based D-scores. Panel **a** shows the complete sample of preterm children, panel **b** the moderately late preterm children, and panel **c** the early preterm children. Panels are split by cohorts, i.e., LOLLIPOP and ePREM.

RESULTS

Participant characteristics

In total, 1686 preterm infants with valid milestone data in LOLLIPOP were matched to 1686 ePREM counterparts. Of these, 1071 were born moderately late preterm, and 615 were born early preterm. In total, 54.7% were males, and 45.4% were females. After matching, cohorts differed regarding maternal educational level ($p < 0.001$). In LOLLIPOP 1047 (72.5%) were of medium/low educational level, whereas this regarded 743 (54.3%) in ePREM.

D-score trajectories in preterm children

In Fig. 1a–c, we present the D-score trajectories for the complete sample of preterm children of both LOLLIPOP and ePREM, as well as separately for moderately late preterm children, and early preterm children, respectively. Visually, the trajectories of the LOLLIPOP and ePREM cohorts were similar. For the ePREM cohort, the cloud of data was slightly less dense, because of fewer observations per child. In Table 1 we present the results from the multilevel model. Overall, the results confirm that the trajectories of LOLLIPOP and ePREM cohorts were similar (Fig. 2). In the complete sample of preterm children, the intercept differed slightly, but statistically significantly, between the LOLLIPOP and ePREM cohort (mean 42.0 vs. 41.7, $p = 0.002$), whereas the slope did not differ between the cohorts. Split by gestational age group, for moderately late preterm children, both the intercept (mean 43.0 vs. 42.3, $p < 0.001$) and the slope (mean 23.5 vs. 23.9, $p = 0.002$) differed between LOLLIPOP and ePREM.

In contrast, for early preterm children, neither the intercept (mean 40.2 vs. 40.1 in LOLLIPOP and ePREM, respectively) nor the slope (mean 25.3 vs. 25.5 in LOLLIPOP and ePREM, respectively) differed between the cohorts. Finally, children in the LOLLIPOP cohort had a smaller random slope variance, indicating that the curves of children in LOLLIPOP were more similar in terms of slope than the curves of children in ePREM. The covariance between the random slope and the random intercept was negative, so children starting with a lower D-score increased in D-score more rapidly than their peers who started with a higher D-score. Adjustment of all models for maternal educational level did not change findings (data not shown).

Adjustment for gestational age

Adjustment for gestational age made intercepts of early and moderately late preterm children more similar. For moderately late preterm children, these were mean 46.3 for LOLLIPOP and 45.5 for ePREM, and for early preterm children these were mean 46.1 for LOLLIPOP and 45.8 for ePREM. Slopes also became similar for both moderately late preterm children (mean 20.9 for both LOLLIPOP and ePREM cohorts) and early preterm children (mean 20.9 for LOLLIPOP and 20.5 for ePREM). Interpretation of the findings did not change for early preterm children, but for moderately late preterm children, the slope was not significantly different between the two cohorts anymore ($p = 0.43$). These data are shown in Table 2.

DISCUSSION

This study aimed to compare neurodevelopmental trajectories, measured using the D-score in two cohorts of early and moderately late preterm infants born 15 years apart. We found no differences in the intercept and slope of neurodevelopmental trajectories in early preterm children. For moderately late preterm children, we found clinically minor differences, although statistically significant, in both the intercept and slope between the two cohorts. These differences were not statistically significant after adjustment for gestational age.

Our results align with most previous findings from comparisons of outcomes at one specific age. For example, in the United Kingdom-based EPICure cohorts born in 1995 and 2006, developmental scores on the Bayley Scales of Infant and Toddler development at 3 years corrected age were similar.[23] In a large Dutch cohort study, similar findings were described.[24] Likewise, in the Australian Victorian Infant Collaborative Study, the prevalence of developmental delay in early preterm children at 2 years of age was not different between cohorts born in 1997 and 2005, even though severe developmental delay was less common in the cohort born in 2005.[25] This Australian cohort was followed-up until school age (8 years), at which age no improvements in neurodevelopment of preterm children were found either.[26] Similarly, in the US cohort from the Neonatal Research Network, followed-up between 1993 and 2018 decreased cerebral palsy prevalence, but stable neurodevelopmental impairment prevalence was found in extremely preterm children.[27,28] In contrast, the French EPIPAGE cohort studies have reported an increase in the prevalence of survival without major neuromotor or sensory disability between cohorts born in 1997 and 2011 at two years of age, with the most improvements in the preterm children born after 24–26 weeks of gestation.[29] The results of EPIPAGE may differ from ours, because in EPIPAGE, neurodevelopment was measured using parental questionnaires, and only major disabilities were included, whereas our data regard professional assessments covering the full range of development. While survival rates of preterm infants increase, especially those at the youngest gestational ages,[30] and neonatal care practices improve,[27] neurodevelopmental trajectories of early preterm children are not improving. This might be due to the higher survival rates, making that infants who would have died a decade ago are now surviving but with poorer

Table 1. Estimates (standard errors) of intercept, slope, and variance in D-scores using chronological ages of two Dutch birth cohorts born 15 years apart—LOLLIPOP and ePREM.

	All preterm infants		Moderately late preterm infants		Early preterm infants	
	LOLLIPOP	**ePREM**	**LOLLIPOP**	**ePREM**	**LOLLIPOP**	**ePREM**
Intercept[a]	42.0 (0.07)	41.7 (0.08)	43.0 (0.07)	42.3 (0.08)	40.2 (0.12)	40.1 (0.19)
Slope[b]	24.1 (0.07)	24.3 (0.12)	23.5 (0.07)	23.9 (0.12)	25.3 (0.13)	25.5 (0.31)
Residual variance	13.5 (0.21)	14.0 (0.36)	12.8 (0.28)	14.0 (0.40)	15.0 (0.34)	14.1 (0.75)
Random intercept variance	6.1 (0.53)	4.8 (0.34)	3.4 (0.46)	3.0 (0.28)	5.9 (0.72)	6.6 (0.79)
Random slope variance	4.3 (0.25)	9.8 (0.88)	2.7 (0.23)	7.1 (0.70)	5.4 (0.52)	18.2 (2.90)
Covariance between random slope and random intercept	−3.8 (0.26)	−2.9 (0.35)	−2.1 (0.19)	−2.0 (0.25)	−3.6 (0.42)	−3.6 (1.03)
P values regarding differences of parameters between LOLLIPOP and ePREM						
Intercept	0.002		<0.001		0.444	
Slope	0.163		0.002		0.588	
Slope and intercept combined	0.008		<0.001		0.718	
Effect sizes regarding differences of parameters between LOLLIPOP and ePREM						
Intercept[c]	0.085		0.189		0.054	
Slope[d]	0.076		0.209		0.063	

LOLLIPOP Longitudinal Preterm Outcome Project (2002-2003), *ePREM* 'e-health preterm project' (2020–2021). Data are presented as estimates (standard errors, SE).

[a]Because of modelling, the intercept is the initial level estimated at ~9 months of age.

[b]The slope is the change per unit increase of age on the log scale.

[c]The effect size for the intercept is interpreted as the size of the mean difference at ~9 months of age.

[d]The effect size for the slope is interpreted as the size of the mean difference between the slopes with respect to the pooled random slope variance, i.e., amongst all of the possible values of the distribution of slopes.

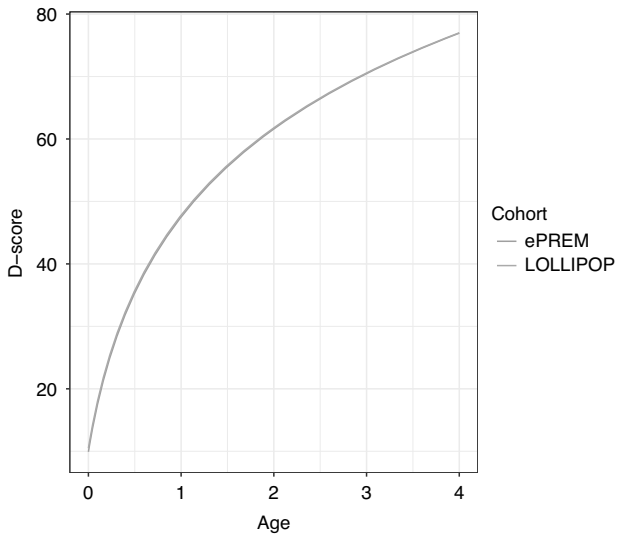

Fig. 2 Developmental trajectories based on the average predicted values from our model using chronological age.

neurodevelopmental outcomes, while infants who would have survived also in the past may have improved regarding neurodevelopmental outcomes. This mechanism would lead to similar averages. Alternatively, a true stagnation may have occurred in the improvement of neurodevelopmental outcomes of preterm children.

For moderately late preterm children, we found statistically significant differences in both the intercept and the slope between the two Dutch cohorts, even though these were clinically minor. As far as we know, this is the first study on trends in neurodevelopmental outcomes in moderately late preterm children in consecutive birth cohorts. On the one hand, the intercept was lower for ePREM than for LOLLIPOP in moderately late preterm children. This may be an artifact due to several small differences in measurement over time. On the other hand, the slope was higher in ePREM than in LOLLIPOP in moderately late preterm children, indicating that the pace of mastering developmental milestones is somewhat higher. This may be explained by a greater emphasis on moderately late preterm children and their outcomes in the last decades,[7] which may have led to changes in the care they received, particularly after discharge from hospital in follow-up programs. However, these differences in slope could also be due to measurement artifact. Both these explanations may exist alongside each other, with differences being only clinically minor.

Strengths and limitations

The major strength of this study is that it is the first study that compares developmental trajectories rather than static measures of development in two large Dutch birth cohort studies, born 15 years apart, with similar developmental assessment tools. We also acknowledge our limitations. First, the two birth cohorts differed regarding maternal educational level. In the LOLLIPOP cohort the educational level was lower, which might influence the D-score trajectories, as socio-economic status has been associated with neurodevelopment.[31] However, adjustment for these differences did not affect the findings. Second, we do not have a control group of fullterm children in ePREM, hampering comparisons between preterm-born children and their full-term-born counterparts in terms of stability of neurodevelopment over time. Finally, the D-score covers all domains of neurodevelopment and does not allow for investigation of differences per developmental domain, which may have disguised differences.

Implications

We found that neurodevelopmental trajectories during ages 0–4 years of early and moderately late preterm-born children hardly

Table 2. Estimates (standard errors) of intercept, slope, and variance in D-scores using corrected ages of two Dutch birth cohorts born 15 years apart —LOLLIPOP and ePREM.

	All preterm infants		Moderately late-preterm infants		Early preterm infants	
	LOLLIPOP	ePREM	LOLLIPOP	ePREM	LOLLIPOP	ePREM
Intercept[a]	45.6 (0.07)	46.2 (0.06)	45.5 (0.08)	46.3 (0.06)	45.8 (0.17)	46.1 (0.10)
Slope[b]	20.8 (0.10)	20.9 (0.06)	20.9 (0.10)	20.9 (0.06)	20.5 (0.26)	20.9 (0.11)
Residual variance	13.7 (0.34)	13.5 (0.22)	13.9 (0.39)	12.9 (0.29)	13.2 (0.70)	14.4 (0.35)
Random intercept variance	3.1 (0.28)	3.2 (0.38)	2.5 (0.28)	2.6 (0.42)	5.0 (0.73)	4.3 (0.69)
Random slope variance	6.6 (0.55)	2.8 (0.17)	4.8 (0.44)	2.1 (0.17)	12.5 (1.82)	3.9 (0.36)
Covariance between random slope and random intercept	−0.8 (0.28)	−1.6 (0.14)	−1.0 (0.21)	−1.5 (0.15)	−0.1 (0.99)	−1.9 (0.29)
P values regarding differences of parameters between LOLLIPOP and ePREM						
Intercept	<0.001		<0.001		0.126	
Slope	0.801		0.428		0.172	
Slope and intercept combined	<0.001		<0.001		0.103	
Effect sizes regarding differences of parameters between LOLLIPOP and ePREM						
Intercept[c]	0.158		0.199		0.071	
Slope[d]	0.015		0.050		0.142	

LOLLIPOP Longitudinal Preterm Outcome Project (2002–2003), *ePREM* 'e-health preterm project' (2020–2021). Data are presented as estimates (standard errors, SE).

[a]Because of modelling, the intercept is the initial level estimated at ~9 months of age.

[b]The slope is the change per unit increase of age on the log-scale.

[c]The effect size for the intercept is interpreted as the size of the mean difference at ~9 months of age.

[d]The effect size for the slope is interpreted as the size of the mean difference between the slopes with respect to the pooled random slope variance, i.e., amongst all of the possible values of the distribution of slopes.

changed over a 15-year period, which indicates that we may need more knowledge on risk selection to better diagnose which infants need interventions, as well as (new) interventions to improve the neurodevelopment of these groups, and further research on the mechanisms relating preterm birth and neurodevelopmental outcomes. Moreover, our findings show that the integration of various developmental assessments in one overarching measure, the D-score, is helpful in mapping developmental trajectories. The wider use of this approach in clinical practice is promising. Additionally, D-scores can be computed from other instruments, like the Bayley or Ages and Stages Questionnaire, thus enhancing the comparability and generalizability of developmental measures. In clinical practice, this may aid pediatricians in the assessment of preterm children, as deviations from the trajectory may be easier to assess and interpret than individual developmental milestone attainment. The implementation of the D-score in clinical practice should, therefore, be considered.

The currently available literature does not entail studies on neurodevelopmental trajectories beyond the first few years of life. As neurodevelopment is a dynamic process, and preterm children may 'grow into deficit' or may 'catch up', such trajectories and their changes between birth cohorts should also be the focus of future studies. Such studies should include both early and moderately late preterm children as groups of interest. Additionally, future studies should also focus on the benefits of early intervention programs in different settings on neurodevelopmental trajectories to support efforts toward improving neurodevelopmental outcomes.

CONCLUSION

In conclusion, neurodevelopmental trajectories, measured using the D-score, in the first 4 years of life are comparable and stable across eras for both early and moderately late preterm children.

Data sharing

De-identified individual participant data will be made available upon reasonable request.

REFERENCES

1. Chawanpaiboon, S. et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob. Health* **7**, e37–e46 (2019).
2. Johnson, S. & Marlow, N. Early and long-term outcome of infants born extremely preterm. *Arch. Dis. Child* **102**, 97–102 (2017).
3. Johnson, S. Cognitive and behavioural outcomes following very preterm birth. *Semin. Fetal Neonatal Med.* **12**, 363–373 (2007).
4. Bos, A. F. & Roze, E. Neurodevelopmental outcome in preterm infants. *Dev. Med. Child Neurol.* **53**, 35–39 (2011).
5. de Kieviet, J. F., Piek, J. P., Aarnoudse-Moens, C. S. & Oosterlaan, J. Motor development in very preterm and very low-birth-weight children from birth to adolescence: a meta-analysis. *JAMA* **302**, 2235–2242 (2009).
6. Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B. & Oosterlaan, J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics* **124**, 717–728 (2009).
7. Johnson, S. et al. Neurodevelopmental outcomes following late and moderate prematurity: a population-based cohort study. *Arch. Dis. Child Fetal Neonatal Ed.* **100**, F301–F308 (2015).
8. de Jong, M., Verhoeven, M. & van Baar, A. L. School outcome, cognitive functioning, and behaviour problems in moderate and late preterm children and adults: a review. *Semin. Fetal Neonatal Med.* **17**, 163–169 (2012).
9. Potijk, M. R., de Winter, A. F., Bos, A. F., Kerstjens, J. M. & Reijneveld, S. A. Higher rates of behavioural and emotional problems at preschool age in children born moderately preterm. *Arch. Dis. Child* **97**, 112–117 (2012).
10. Cserjesi, R. et al. Functioning of 7-year-old children born at 32 to 35 weeks' gestational age. *Pediatrics* **130**, e838–e846 (2012).
11. Kerstjens, J. M. et al. Developmental delay in moderately preterm-born children at school entry. *J. Pediatr.* **159**, 92–98 (2011).
12. Spittle, A. J., Orton, J., Doyle, L. W., Boyd, R. Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants. *Cochrane Database Syst. Rev.* **2015**, CD005495 (2015).

13. Kynø, N. M. et al. Effect of an early intervention programme on development of moderate and late preterm infants at 36 months: a randomized controlled study. *Infant Behav. Dev.* **35**, 916–926 (2012).
14. Van Hus, J. et al. Early intervention leads to long-term developmental improvements in very preterm infants, especially infants with bronchopulmonary dysplasia. *Acta Paediatr.* **105**, 773–781 (2016).
15. Frankenburg, W. K., Dodds, J., Archer, P., Shapiro, H. & Bresnick, B. The denver II: a major revision and restandardization of the denver developmental screening test. *Pediatrics* **89**, 91–97 (1992).
16. Kerstjens, J. M. et al. Support for the global feasibility of the ages and stages questionnaire as developmental screener. *Early Hum. Dev.* **85**, 443–447 (2009).
17. Weber, A. M. et al. The D-score: a metric for interpreting the early development of infants and toddlers across global settings. *BMJ Glob. Health* **4**, e001724 (2019).
18. Hornman, J., de Winter, A. F., Kerstjens, J. M., Bos, A. F. & Reijneveld, S. A. Stability of developmental problems after school entry of moderately late preterm and early preterm-born children. *J. Pediatr.* **187**, 73–79 (2017).
19. van Noort-van der Spek, I. L., Dudink, J., Reiss, I. K. M. & Franken, M. C. J. P. Early speech sound production and its trajectories in very preterm children from 2 to 4 years of age. *J. Speech Lang. Hear Res.* **65**, 1294–1310 (2022).
20. Jacobusse, G., van Buuren, S. & Verkerk, P. H. An interval scale for development of children aged 0-2 years. *Stat. Med.* **25**, 2272–2283 (2006).
21. R Foundation for Statistical Computing. R: a language and environment for statistical computing. [Computer Program], 2020; version 4.0.2.
22. Muthen, L. K. & Muthen, B. Mplus user's guide: Statistical analysis with latent variables, (2017). Available from: https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r6_web.pdf.
23. Moore, T. et al. Neurological and developmental outcome in extremely preterm children born in england in 1995 and 2006: the EPICure studies. *BMJ* **345**, e7961 (2012).
24. van Beek, P. E. et al. Developmental trajectories in very preterm born children up to 8 years: a longitudinal cohort study. *Front Pediatr.* **9**, 672214 (2021).
25. Doyle, L. W., Robert, G. & Anderson, P. J. Outcomes at age 2 years of infants < 28 weeks' gestational age born in victoria in 2005. *J. Pediatr.* **156**, 49–53.e1 (2010).
26. Cheong, J. L. Y. et al. Changing neurodevelopment at 8 years in children born extremely preterm since the 1990s. *Pediatrics* **139**, e20164086 (2017).
27. Stoll, B. J. et al. Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993-2012. *JAMA* **314**, 1039–1051 (2015).
28. Adams-Chapman, I. et al. Neurodevelopmental impairment among extremely preterm infants in the neonatal research network. *Pediatrics* **141**, e20173091 (2018).
29. Pierrat, V. et al. Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in france in 2011: EPIPAGE-2 cohort study. *BMJ* **358**, e3448 (2017).
30. Santhakumaran, S. et al. Survival of very preterm infants admitted to neonatal care in England 2008-2014: Time trends and regional variation. *Arch. Dis. Child Fetal Neonatal Ed.* **103**, F208–F215 (2018).
31. Potijk, M. R., Kerstjens, J. M., Bos, A. F., Reijneveld, S. A. & de Winter, A. F. Developmental delay in moderately preterm-born children with low socioeconomic status: risks multiply. *J. Pediatr.* **163**, 1289–1295 (2013).

ACKNOWLEDGEMENTS

This study was part of a larger study on the growth and development of preterm children. We wish to thank the participating preventive child healthcare PCHC physicians for their contribution to the fieldwork. In particular, we thank Drs. E.M.J. ten Vergert, M. Broer van Dijk, and B. van der Hulst for coordinating the project. We are grateful to the ePREM research group for their contribution to this cohort study. Finally, we thank Dr. J. Almansa for his assistance in the statistical analyses.

AUTHOR CONTRIBUTIONS

Ms. van Dokkum N.H. assisted in the data handling and statistical analyses and drafted the initial manuscript. Mr. Lepe A. performed the data handling, data visualizations, and statistical analyses. Prof van Buuren S. conceptualized and designed the D-score and performed part of the statistical analyses. Prof Reijneveld S.A. conceptualized and designed the LOLLIPOP study and data handling. Prof de Kroon M.L.A. conceptualized and designed the ePREM study and data handling and supervised the execution of the LOLLIPOP and ePREM studies. All authors critically reviewed and revised the manuscript and approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

FUNDING

LOLLIPOP was supported by grants from the research foundation of Beatrix Children's Hospital, the Cornelia Foundation for the Handicapped Child, and the research fund of A. Bulk-Child Preventive Child Health Care, Dutch Brain Foundation, and unrestricted investigator-initiated research grants from FrieslandCampina, Friso Infant Nutrition, and Pfizer Europe. The financiers had no role at any stage of the project, including the decision to submit the manuscript. ePREM was supported by a grant from ZonMw (grant 70-73600-98-519).

COMPETING INTERESTS

The authors declare no competing interests.

CONSENT STATEMENT

The LOLLIPOP study was approved by the Institutional Review Board of the University Medical Center Groningen (METc 2005/130), and written informed consent was provided by all parents.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Nienke H. van Dokkum.