Additional Information	NOTICE: This is the author's version of a work that was accepted for publication in Infant Behavior and Development. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Infant Behavior and Development, Vol. 36, Issue 4, (2013). doi: http://doi.org/10.1016/j.infbeh.2013.07.003

Title
Trajectories and predictors of developmental skills in healthy twins up to 24 months of age

Short title: Trajectories and predictors of twin development

Authors
Cassandra Nan[1,2], MSc, Jan Piek[2], PhD, Claire Warner[1], MSc, Diane Mellers[3], Ruth Elisabeth Krone[1], MD, Timothy Barrett[4], PhD and Maurice Zeegers[1,5], PhD

[1]Unit of Urologic and Genetic Epidemiology, Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, United Kingdom
[2]School of Psychology and Speech Pathology, Curtin University, Perth, Western Australia, Australia
[3]Birmingham Women's Hospital, Birmingham, United Kingdom
[4]School of Clinical and Experimental Medicine, University of Birmingham, United Kingdom
[5]Department of Complex Genetics, Cluster of Genetics and Cell Biology, Nutrition and Toxicology Research Institute Maastricht, Maastricht University, the Netherlands

Corresponding author
Cassandra Nan
Unit of Urologic and Genetic Epidemiology, Department of Public Health and Epidemiology, University of Birmingham, Birmingham, B15 2TT, United Kingdom
E-mail: cxn833@bham.ac.uk, Phone: +31 (0)6 42831566

Abstract: 233, **Body:** 3,821, **Tables:** 2, **Figures:** 1, **References:** 64

Keywords: twins; early development; developmental trajectories; longitudinal; perinatal factors

Abbreviations
ASQ-3 = Ages and Stages Questionnaires-3
BiRTHS = Birmingham Registry for Twins and Heritability Studies

Abstract

Background: Low birth weight and low 5-minute Apgar scores have been associated with developmental delay, while older maternal age is a protective factor. Little is known about trajectories and predictors of developmental skills in infant twins, who are generally born with lower birth weights, lower Apgar scores and to older mothers.

Methods: Developmental skills were assessed at 3, 6, 9, 12, 18 and 24 months using the Ages and Stages Questionnaires in 152 twins from the Birmingham Registry for Twin and Heritability Studies. Multilevel spline and linear regression models (adjusted for gestational age, gender, maternal age) were used to estimate developmental trajectories and the associations between birth weight, maternal age and Apgar scores on developmental skills.

Results: Twins performed worse than singletons on communication, gross motor, fine motor, problem solving and personal-social skills ($p<0.001$). Twins caught up around 6 months (score within -1 standard deviation of norm), except on gross motor skills, which did not catch up until after the age of 12 months. A one-year increase in maternal age was significantly associated with decreases in gross motor and personal-social z-scores of up to -0.09, whereas one unit increases in Apgar score increased z-scores up to 0.90 ($p<0.01$).

Conclusions: Healthy twins should be considered at a higher risk for developmental delay. Whether these results are comparable to preterm singletons, or whether there are twin-specific issues involved, should be further investigated in a study that uses a matched singleton control group.

1 Introduction

Developmental skills can be roughly categorised as cognitive, communicative, social and emotional, and psychomotor development. Children are expected to have reached developmental milestones in each of these categories within a certain age window(Anderson, Northam, Hendy, & Wrennal, 2005). Developmental delays have been related to a variety of biological, economic and lifestyle factors. Firstly, children need to have access to resources that allow them to have the opportunity to learn necessary skills; these resources may be lacking in families of lower socioeconomic status (SES). Indeed, low SES has been found to be a risk factor for developmental and behavioural problems(Bradley & Corwyn, 2002; Laucht, Esser, & Schmidt, 1997; Pike, Iervolino, Eley, Price, & Plomin, 2006). Further, biological factors that have been related to delays in cognitive skills are mostly ante- and perinatal, such as lower birth weight(Datar & Jacknowitz, 2009; Jeng, Yau, Liao, Chen, & Chen, 2000), shorter gestations(Hediger, Overpeck, Ruan, & Troendle, 2002; Sullivan & Margaret, 2003; Thun-Hohenstein, Largo, Molinari, Kundu, & Duc, 1991) and Apgar scores lower than 7(de Moura, et al., 2010; Odd, Rasmussen, Gunnell, Lewis, & Whitelaw, 2008). Maternal age could be a mediator between perinatal outcomes and developmental skills in the following ways: very young maternal age (<15 years)(Mousiolis, Baroutis, Sindos, Costalos, & Antsaklis, 2013) has been associated with an increased risk for preterm birth, while older maternal age (>40 years) has been related to shorter gestations and low birth weight(Lung, Shu, Chiang, & Lin, 2009). Lung et al. (2009) (Lung, et al., 2009)also reported that older mothers were more likely to opt for assisted reproductive treatment, particularly in vitro fertilisation (IVF), which has been related to preterm birth and low birth weight(Hvidtjorn, et al., 2009). In a recent review on the developmental consequences of IVF, Hart and

Norman (2013)(Hart & Norman, 2013) found no evidence that children conceived through IVF were at higher risk for developmental delays. However, they reported several studies that suggested any differences between IVF- and naturally conceived children could be due to factors related to maternal age and preterm birth.

Not all domains are equally affected by perinatal and maternal factors. For instance, maternal age has been more often related to cognitive development and behaviour(D. M. Fergusson & Lynskey, 1993; Ketterlinus, Henderson, & Lamb, 1991), while motor skills seem more sensitive to perinatal morbidity(Sullivan & Margaret, 2003). Gender differences in behavioural and cognitive development may be observed as early as infancy, but become more apparent from childhood onwards(Alexander & Wilcox, 2012).

Even though most children seem to catch up on developmental skills in early (Gasson & Piek, 2003; van Haastert, de Vries, Helders, & Jongmans, 2006) and middle childhood (Thorpe, 2006), a major consequence of developmental delay is the higher use of special school services(Sullivan & Margaret, 2003) and higher risk of subsequent emotional and behavioural problems(Emerson & Einfeld, 2010). Furthermore, children with general developmental delay (delays in at least 2 developmental domains) show less activity participation compared to their peers, which can be partly explained by impaired social and motor skills(Leung, Chan, Chung, & Pang, 2010). Psychomotor skills have also been argued to change the way an infant interacts with its environment, which could be important for the development of communicative skills(Iverson, 2010). Moreover, fine motor skills in infancy are strong predictors of later intellectual functioning(Thun-Hohenstein, et al., 1991) and general academic achievements at school age(Sullivan & Margaret, 2003). Furthermore, earlier attainment of gross motor milestones have been associated with better executive performance at school age(Piek, Dawson, Smith, & Gasson, 2008) and learning to stand at an earlier age has been associated with better cognitive categorisation in adulthood(Murray, et al., 2006), which in turn has been associated with reasoning ability(Salthouse, 2005) and specific areas of language development(Gopnik & Meltzhoff, 1987).

Tools to assess these developmental milestones include the Bayley Scales of Infant Development(Bayley, 2006), Wechsler Intelligence Scales for Children(Wechsler, 2003) and Griffiths Mental Development Scales(McLean, McCormick, & Baird, 1991). There are also developmental screening tools, which give a general overview of a child's cognitive abilities and do not take as much time to administer, as they are often questionnaires that can be completed by parents. Examples of such screening tools are the Ages and Stages Questionnaire(Squires, Bricker, & Potter, 2009) and the Parents Evaluations of Developmental Status(Glascoe, 2010). Generally, normative groups for developmental assessments consist of healthy singletons. Sutcliffe and Derom (2006)(Sutcliffe & Derom, 2006) addressed the question whether twins can be directly compared to singletons. Even when otherwise physically healthy, twins are likely to be born at earlier gestations (average of 37 weeks), with lower average birth weights and have older mothers. All of these factors have been directly or indirectly related to developmental delay and later cognitive skills as outlined above. Twins have been previously found to have lower intelligence quotients (IQ) than their singleton peers up to adolescence (Cooke, 2010; Drillien, 1969; Ronalds, De Stavola, & Leon, 2005). Furthermore, it has been suggested that development may vary between types of twins, whereby dizygotic twins attain gross motor milestones at an earlier age than monozygotic twins(Brouwer, van Beijsterveldt, Bartels, Hudziak, & Boomsma, 2006).

The majority of previous studies have mainly focused on risk factors for developmental delay after the first 2 years. However, to our knowledge, no previous research has been conducted into the development in the first 2 years of healthy twins, who were not born extremely small, small-for-gestational-age or very premature. Therefore, this study serves to investigate the association of perinatal outcomes (birth weight and Apgar score) and maternal age with developmental skills in the first 2 years of life in healthy infant twins. Additionally, we used a developmental screening tool to compare the development of healthy twins between the ages of 3 and 24 months with the standardised scores based on normally developing singletons.

2. Materials and methods

2.1 Participants

Participants were part of the Birmingham Registry for Twin and Heritability Studies (BiRTHS)(Krone, et al., 2006), a multiple birth registry, which recruits eligible families from three large hospitals in Birmingham, United Kingdom. Mothers of twins in England are under the care of consultants in hospitals. We recruited these mothers, between 12 and 28 weeks of gestation, in the period of 4th August 2008-15th July 2011. Following the World Health Organisation standards for a viable live birth, we included twins who were born after 22 weeks and weighed more than 500 grams at birth. Additionally, we excluded twins who had (suspected) congenital defects that could influence physical growth and mental development.

We approached 365 families, of which 170 agreed to participate in the 2-year study. By 15 July 2011, 11 twin pairs had not yet reached the age of 3 months and were not eligible for follow-up. Of the remaining 159 families, 40 mothers did not complete the antenatal questionnaire (explained below), and we did not receive any follow-up questionnaires for 36 twin pairs. Finally, 152 infants (27 monozygotic and 48 dizygotic pairs, and 2 dizygotic twins without co-twin), for whom delivery details were also available, were included in this study.

The number of completed questionnaires at each follow-up can be found in Table 1. Due to the relatively late introduction of the ASQ-3 into the study (March 2009), 11 twin pairs were missed for the 3- or 6-month questionnaires. We were unable to complete the 24-month follow-up for all twins, because 50% of children had not yet reached the age of 18 months by the end of the study period. Response rate at each follow-up was approximately 50% and we found no maternal or obstetric differences between responders and non-responders. Overall, 22% completed only one follow-up, 12% had two completed follow-ups, 28% completed three follow-ups, 18% completed four follow-ups, 17% completed five follow-ups and 3% completed all follow-ups.

2.2 Measures

Upon joining, parents completed an antenatal questionnaire about demographic, lifestyle, employment status and social support details. Maternal medical history and delivery details were extracted from maternal medical records. Developmental skills were assessed with the ASQ-3.

The ASQ-3 System is a series of validated and standardized questionnaires that measure developmental milestones of children at regular intervals between 2 and 60 months, and is meant as a simple method for initial screening for any developmental delays (Klamer, Lando, Pinborg, & Greisen, 2005; Limbos & Joyce, 2011). The questionnaires consist of simply phrased questions based on developmental milestones at each follow-up age in the following subscales: communication, gross motor, fine motor, problem solving and personal social skills. Each subscale consists of six questions, which parents answer with ‘yes’ (score=10), ‘sometimes’ (score=5) or ‘not yet’ (score=0), resulting in a range of scores between 0 and 60. These scores correspond with three areas in which a child can be categorised: ‘normal development’ (above -1 standard deviation (SD)), ‘cause for concern and further monitoring needed’ (between 1 and 2 SD below the mean) and ‘referral for further developmental assessment needed’ (>2SD below the mean).

Normed scores are based on healthy American singleton children born at 40 weeks gestation. There are currently no norm scores for the United Kingdom. Based on an American sample of 15,138 children, sensitivity and specificity for all questionnaires in the current study were between 70% and 100%. Furthermore, there was a 84% overall agreement between the ASQ-3 and the Bayley Scales of Infant Development(Squires, et al., 2009). Finally, although the sensitivity of the ASQ-3 in very preterm populations has been previously questioned, Schonhaut et al. (2013) reported that sensitivity of the ASQ-3 in extremely premature infants was 86%(Schonhaut, Armijo, Schonstedt, Alvarez, & Cordero, 2013).

2.3 Procedure

Maternal medical records provided data on mode of conception, gender, Apgar scores and birth anthropometry. Maternal age, employment status and ethnic background were extracted from the antenatal parental questionnaire. Similar to a large UK-based follow-up study(Smyth, Spark, Armstrong, & Duley, 2009; The Magpie Trial Follow Up Study Management & The Magpie Trial Follow Up Study Collaborative, 2004; Yu, et al., 2007), we assessed developmental skills at 3, 6, 9, 12, 18 and 24 months with the ASQ-3(Squires, et al., 2009). The questionnaires were completed within 2 weeks of each chronological follow-up age, e.g.: the 12-month questionnaire was completed between 11 month 2 weeks and 12 months 2 weeks.

2.4 Statistical analyses

2.4.1 Descriptive data

All analyses were performed using STATA 11(StataCorp, 2009). Raw ASQ-3 scores were converted into z-scores using the norm sample mean and SD for each of the categories described in section 2.3. The z-scores were used in further analyses. T-tests were performed to investigate ASQ-3 z-score differences between twins and the singleton norm sample. We also used t-tests to investigate developmental differences by type of twin, birth order, mode of conception, gender, maternal employment status (as proxy for socioeconomic status) and ethnic background.

2.4.2 Determinants of developmental skills

Cross-sectional multilevel linear regression analyses were performed to investigate the associations between birth weight, maternal age and Apgar scores on the ASQ-3 at each follow-up age. Although the twins were considered individuals in these analyses, their relatedness was taken into account by including this in a nested level. We adjusted for prematurity by including gestational age at birth as a confounder. Gender was also included in the regression model as a confounder. Ethnicity and mode of conception did not improve the regression models and were therefore not included. Results were considered statistically significant at α=0.01 due to multiple testing.

2.4.3 Developmental trajectories

Exploratory analyses showed nonlinear patterns for developmental skills over time. Therefore, multilevel spline regression models(Gutierrez, 2008) were used in which repeated measures within each individual and the relatedness of twins were included as random effects in a nested level. Gestational age at birth, gender and maternal age (based on the results from the analyses in section 2.4.2) were included as confounders. Predicted values from these models were graphically displayed for each ASQ-3 subscale.

3 Results

3.1 Participants

Overall, 53% of the children were male. A third of the children were conceived through fertility treatment, of which 26% were through IVF. The median gestational age at birth was 37 weeks (range: 26-39 weeks). Mean birth weight was 2.3 kilograms (kg; SD=0.55), ranging from 0.94 to 3.5kg. Two infants weighed <1kg at birth and 11 infants weighed between 1 and 1.5kg. The median Apgar score at 5 minutes was 9 (range: 4-10). Mothers' mean age at delivery was 31.9 years (standard deviation [SD]=5.51, range: 20-41 years), 87% of parents were Caucasian and 88% were in paid employment.

3.2 Descriptive data

There were no differences in developmental skills between mono- and dizygotic twins or within twin pairs. We also did not find any significant differences when stratified by mode of conception, gender, birth weight, Apgar score, maternal age, employment status or ethnicity ($p>0.01$). Also, we did not find a significant correlation between maternal age and IVF. We did, however, find significant mean z-score differences between the twin sample and singleton norm on gross motor, problem solving and personal-social subscales at 3, 6, 9 and 12 months ($p<0.001$) (Table 1). Z-scores on the communication and fine motor subscales were significantly lower than the singleton norm at 3, 6 and 9 months ($p<0.001$). Furthermore, twins' performance on fine motor and personal-social skills at 3 months, and gross motor skills at 3 and 6 months, fell below the -1SD cut-off score ($p<0.001$). Problem solving skills at 3 months were significantly worse than the -2SD cut-off score (t=-2.43, $p=0.01$).

[INSERT TABLE 1]

3.3 Determinants of developmental skills

Table 2 describes the effects of birth weight, Apgar score and maternal age on developmental skills. Birth weight had no significant effect on developmental skills at $p=0.01$, although there was a positive trend for higher birth weight and 6-month personal-social skills ($p=0.04$).

[INSERT TABLE 2]

A point increase in Apgar score was associated with a 0.81 increase in gross motor skills at 18 months ($p<0.001$) and a 0.90 increase at 24 months ($p<0.01$). Similar trends were found for fine motor and problem solving at 24 months ($p=0.02$). Older maternal age, on the other hand, seemed to have a predominantly negative effect on developmental skills. A one-year increase in maternal age was associated with decreases on communication z-scores of 0.08 at 9 months ($p<0.01$), 0.06 at 12 months ($p<0.001$), 0.05 at 18 months ($p=0.04$) and 0.06 at 24 months ($p=0.01$). Similarly, older maternal age decreased problem solving by 0.09 at 9 months ($p=0.01$) and by 0.07 at 24 months ($p=0.01$). Personal-social skill z-scores decreased by 0.07 at 9 months (*p*<0.01) for each year's increase in maternal age. There were negative trends for maternal age and 18-month communication, 12-month gross motor, 9- and 12-month fine motor, and 3- and 12-month personal-social skills (*p*=0.04).

3.4 Developmental trajectories

Figure 1 shows the developmental trajectories based on median ASQ-3 z-scores, adjusted for gender, gestational age at birth and maternal age. There was an overall improvement in communication, gross motor, fine motor, problem solving and personal-social skills between 3 and 24 months.

[INSERT FIGURE 1]

4 Discussion

In this study, we found that healthy twins show signs of developmental delay in communication, motor, problem solving and personal-social skills during the first year of life, even when adjusted for prematurity. These developmental delays remained even when three very preterm twin pairs were excluded (results not shown), and could therefore not be due to the inclusion of children born <30 weeks. In contrast to other studies(Asbury, Almeida, Hibel, Harlaar, & Plomin, 2008; Rutter & Redshaw, 1991; Rutter, Thorpe, Greenwood, Northstone, & Golding, 2003), we did not find twin-singleton differences after 9 months, which implies that twins in our study caught up quicker than expected based on previous reports. The speed in which twins caught up to singletons varied per skill as well as follow-up age. Scores for gross motor skills were one of the last to reach a plateau. This is somewhat in line with previous studies that have reported low stability of motor skills from birth up to school age(Janssen, et al., 2011; Roze, et al., 2010) and that motor skills seem to be worse than cognitive and language skills in preterm infants(Siegel, et al., 1982). Furthermore, motor development delays in very preterm or very low birth weight children seem to last throughout childhood(de Kieviet, Piek, Aarnoudse-Moens, & Oosterlaan, 2009).

We found no gender or within twin-pair differences, which were suggested in previous studies(Asbury, et al., 2008; Berglund, Eriksson, & Westerlund, 2005; Thorpe, 2006). However, as described in the introduction, gender differences are limited in infancy and become more apparent later in life. It could be similar for differences within twin pairs, when unique environmental influences for each co-twin become larger.

Finally, the lack of significant correlation between maternal age and IVF did not agree with Lung et al.(Lung, et al., 2009), who postulated that older mothers were more likely to undergo IVF treatment. About 41% of all women who underwent IVF treatment in 2011 in the UK were in the age range of 18-34 years, compared with 63% up to the age of 34 in our sample("Fertility treatment in 2011: Trends & Figures," 2013). It is possible that we did not find a correlation between maternal age and IVF treatment, because there were fewer older mothers (>34 years) who underwent fertility treatment.

4.1 Determinants of developmental skills

Our findings from the regression analyses suggest that maternal age mostly influences developmental skills in infancy, while birth weight and 5-minute Apgar scores were mostly not related to developmental skills. Overall, maternal age seemed to have a negative effect on developmental skills within the first 24 months. Particularly communication, problem solving and personal-social skills at 9 and 12 months seemed negatively affected by older maternal age. Older maternal age has been indirectly related to delayed motor and social development(Sherlock, Synnes, & Koehoorn, 2008), as well as cognitive development at school age(Harvey, 1999; Waldfogel, Han, & Brooks-Gunn, 2002) in previous studies. It is possible that older mothers are more likely to return to work within the first year after birth(Sherlock, et al., 2008) and spend relatively less time with their children. Maternal age as a mediator for factors, such as mother-child interaction(E. Fergusson, Maughan, & Golding, 2008), educational level and maternal occupational commitments(Sherlock, et al., 2008), could explain the contradictory findings in the current study. We were unable to further investigate the relationship between developmental skills and the age at which mothers returned to work, as we did not have any details on maternal leave duration.

Higher Apgar scores were advantageous for gross motor skills at 18 and 24 months, which is later than in previous studies(Gasson & Piek, 2003). It is possible that the effect of birth outcomes, such as Apgar score, on infant development is not large enough to detect. Measurements beyond infancy would be needed to investigate whether an association can be observed when the required motor skills become more complex later in childhood and developmental delays become more apparent and obstruct other areas of development(Goyen & Lui, 2002). Compared with maternal age, Apgar scores did not seem to have much influence on developmental skills. This could be because there is simply no effect to be found, or that there might be other perinatal factors that mediate the influence of Apgar scores(Erdemoglu, Mungan, Tapisiz, Ustunyurt, & Caglar, 2003; Rossi, Mullin, & Chmait, 2011). It is also possible that our study did not include enough children with low 5-minute Apgar scores to observe an effect.

Also contrary to previous studies(Datar & Jacknowitz, 2009; Jeng, et al., 2000), but in line with Rutter et al. (2003)(Rutter, et al., 2003), results from the current study show that birth weight has no or little effect on developmental scores throughout the first 2 years. It is possible that the little variance in birth weight in our study prevented us from finding any association with developmental skills.

5 Limitations

We used the ASQ-3 in this study, because they are easy parent-completed questionnaires and require very little time to complete. However, the questionnaires are limited by a ceiling effect, where children cannot score higher than 'normal development' on each subscale. The use of other (observational or interactive) assessment tools is necessary to confirm the current findings.

Although sample size varied at each age, there were no significant differences in average and range for birth weight, gestational age at birth and maternal age between the follow-up times. Nevertheless, a larger, more complete sample and a singleton control group would be desirable to confirm the results in this study in a controlled, longitudinal analysis and to investigate whether there is indeed a twin-singleton difference in predictors of developmental skills.

Further follow-up is also needed to determine within-twin pair difference beyond infancy and the effect of environmental factors in childhood, such as twin-twin interaction(Rutter & Redshaw, 1991), type of childcare, socioeconomic status and parent-infant interaction(E. Fergusson, et al., 2008; Rutter & Redshaw, 1991). Communicative skills in particular could be affected by parent-child and twin-twin interaction that is specific to raising young same-aged children(Rutter, et al., 2003).

Ethnicity was not included in our analyses, as the addition of ethnicity did not significantly improve the regression model. However, previous studies have found that ethnicity influences parents' child rearing choices, which in turn have an effect on child development(Bradley & Corwyn, 2002; E. Fergusson, et al., 2008; McLoyd, 1998). It seems that ethnicity might indirectly influence child development through the abovementioned environmental factors. As with twin-singleton and within twin pair differences, these environmental influences could emerge later in life. We recommend that future studies to include ethnicity and environmental factors as potential predictors of developmental skills.

6 Conclusion

Although the data in the current study did not allow for longitudinal analyses, healthy twins did score below the normal range on current singleton norms for cognitive, communicative, social and emotional, and psychomotor development at every follow-up in the first year of life. Considering the previous literature on the development of twins compared with singletons, our finding that twins are delayed in early infancy seems to be reasonable. Whether these results are comparable to preterm singletons, or whether there are twin-specific issues involved, should be further investigated in a study that uses a matched singleton control group. Nonetheless, twins should be considered at a higher risk for developmental delay and those who formally assess the development of twins at this early age should be aware of this when reporting their findings to parents. Birth weight and Apgar scores do not seem to be as influential as maternal age at this stage, but may become more important later in life. Similarly, further investigation is needed into environmental influences on child development as they emerge at older ages.

Results should be confirmed by using more comprehensive assessments of child development as well as possible influential factors. While the association between unfavourable perinatal outcomes and developmental delays is well documented, as described in the introduction, only little information is available on the early development of twins, who are at higher risk for unfavourable perinatal outcomes. The current study is therefore a valuable addition to the literature and provides direction for future research.

Acknowledgements

The authors would like to thank the School of Psychology and Speech Pathology at Curtin University for hosting CN, while conducting this research. We are also grateful to all of the BiRTHS families who participated in the study, and hospital staff who have helped in recruitment and data collection. Ethical approval for this study was obtained from the Birmingham North, East and Solihull Research Ethics Committee in 2006 [06/Q2706/87].

Financial disclosure

This research was funded by the Birmingham Children's Hospital Research Foundation [BCHRF102f].

Conflict of interest

The authors declare no conflicts of interest.

List of references

Alexander, G. M., & Wilcox, T. (2012). Sex Differences in Early Infancy. *Child Development Perspectives, 6,* 400-406.

Anderson, V., Northam, E., Hendy, J., & Wrennal, J. (2005). Cognitive development. In *Developmental Neuropsychology*. New York, NY: Psychology Press Ltd.

Asbury, K., Almeida, D., Hibel, J., Harlaar, N., & Plomin, R. (2008). Clones in the classroom: a daily diary study of the nonshared environmental relationship between monozygotic twin differences in school experience and achievement. *Twin Research and Human Genetics, 11*, 586-595.

Bayley, N. (2006). *Bayley scales of infant and toddler development* (3rd ed.). San Antonia, TX: Harcourt Assessment, Inc.
Berglund, E., Eriksson, M., & Westerlund, M. (2005). Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandivian Journal of Psychology, 46*, 485-491.
Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annu Rev Psychol, 53*, 371-399.
Brouwer, S. I., van Beijsterveldt, T. C., Bartels, M., Hudziak, J. J., & Boomsma, D. I. (2006). Influences on achieving motor milestones: a twin-singleton study. *Twin Res Hum Genet, 9*, 424-430.
Cooke, R. (2010). Does neonatal and infant neurodevelpmental morbidity of multiples and singletons differ? *Seminars in Fetal & Neonatal Medicine, 15*, 362 - 366.
Datar, A., & Jacknowitz, A. (2009). Birth weight effects on children's mental, motor, and physical development: evidence from twins data. *Maternal and Child Health Journal, 13*, 780-794.
de Kieviet, J. F., Piek, J. P., Aarnoudse-Moens, C. S., & Oosterlaan, J. (2009). Motor development in very preterm and very low-birth-weight children from birth to adolescence: a meta-analysis. *Journal of the American Medical Association, 302*, 2235-2242.
de Moura, D. R., Costa, J. C., Santos, I. S., Barros, A. J., Matijasevich, A., Halpern, R., Dumith, S., Karam, S., & Barros, F. C. (2010). Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatric and Perinatal Epidemiology, 24*, 211-221.
Drillien, C. M. (1969). School disposal and performance for children of different birthweight born 1953-1960. *Archives of Disease in Childhood, 44*, 562-570.
Emerson, E., & Einfeld, S. (2010). Emotional and behavioural difficulties in young children with and without developmental delay: a bi-national perspective. *J Child Psychol Psychiatry, 51*, 583-593.
Erdemoglu, E., Mungan, T., Tapisiz, O. L., Ustunyurt, E., & Caglar, E. (2003). Effect of inter-twin delivery time on Apgar scores of the second twin. *Australian and New Zealand Journal of Obstetrics and Gynaecology, 43*, 203-206.
Fergusson, D. M., & Lynskey, M. T. (1993). Maternal age and cognitive and behavioural outcomes in middle childhood. *Paediatr Perinat Epidemiol, 7*, 77-91.
Fergusson, E., Maughan, B., & Golding, J. (2008). Which children receive grandparental care and what effect does it have? *Journal of Child Psychology & Psychiatry, 49*, 161-169.
. Fertility treatment in 2011: Trends & Figures. In. (2013): Human Fertilisation & Embryology Authority.
Gasson, N., & Piek, J. P. (2003). A longitudinal investigation of the development of fullterm, preterm and 'at-risk' infants from birth to four years using a parent report measure. In M. Katsikitis (Ed.), *38th Annual APS Conference* (pp. 78 - 82). Perth.
Glascoe, F. P. (2010). *Parents' Evaluation of Developmental Status (PEDS)*: Nolensville, TN: PEDSTest.com, LLC.
Gopnik, A., & Meltzhoff, A. (1987). The Development of Categorization in the Second Year and It Relation to Other Cognitive and Linguistic Developments. *Child Dev, 58*, 1523-1531.

Goyen, T. A., & Lui, K. (2002). Longitudinal motor development of "apparently normal" high-risk infants at 18 months, 3 and 5 years. *Early Human Development, 70*, 103-115.
Gutierrez, R. G. (2008). Tricks of the trade: Getting the most out of xtmixed. In *North American Stata Users Group Meeting*.
Hart, R., & Norman, R. J. (2013). The longer-term health outcomes for children born as a result of IVF treatment. Part II-Mental health and development outcomes. *Hum Reprod Update, 19*, 244-250.
Harvey, E. (1999). Short-term and long-term effects of early parental employment on children of the National Longitudinal Survey of Youth. *Developmental Psychology, 35*, 445-459.
Hediger, M. L., Overpeck, M. D., Ruan, W. J., & Troendle, J. F. (2002). Birthweight and gestational age effects on motor and social development. *Paediatric and Perinatal Epidemiology, 16*, 33-46.
Hvidtjorn, D., Schieve, L., Schendel, D., Jacobsson, B., Svaerke, C., & Thorsen, P. (2009). Cerebral palsy, autism spectrum disorders, and developmental delay in children born after assisted conception: a systematic review and meta-analysis. *Arch Pediatr Adolesc Med, 163*, 72-83.
Iverson, J. M. (2010). Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language, 37*, 229-261.
Janssen, A. J., Akkermans, R. P., Steiner, K., de Haes, O. A., Oostendorp, R. A., Kollee, L. A., & Nijhuis-van der Sanden, M. W. (2011). Unstable longitudinal motor performance in preterm infants from 6 to 24 months on the Bayley Scales of Infant Development--Second edition. *Research in Developmental Disabilities, 32*, 1902-1909.
Jeng, S. F., Yau, K. I., Liao, H. F., Chen, L. C., & Chen, P. S. (2000). Prognostic factors for walking attainment in very low-birthweight preterm infants. *Early Human Development, 59*, 159-173.
Ketterlinus, R. D., Henderson, S., & Lamb, M. E. (1991). The effects of maternal age-at-birth on children's cognitive development. *Journal of Research on Adolescence, 1*, 173-188.
Klamer, A., Lando, A., Pinborg, A., & Greisen, G. (2005). Ages and Stages Questionnaire used to measure cognitive deficit in children born extremely preterm. *Acta Paediatrica*, *94*, 1327-1329.
Krone, R. E., Ewer, A. K., Barrett, T. G., Moy, R. J., Bakour, S., Maher, E. R., Thangaratinam, S., Cox, P., Martin, B., Khan, K. S., & Zeegers, M. P. (2006). The Birmingham Registry for Twin and Heritiability Studies (BiRTHS). *Twin Research and Human Genetics, 9*, 907-912.
Laucht, M., Esser, G., & Schmidt, M. H. (1997). Developmental outcome of infants born with biological and psychosocial risks. *J Child Psychol Psychiatry, 38*, 843-853.
Leung, G. P., Chan, C. C., Chung, R. C., & Pang, M. Y. (2010). Determinants of activity and participation in preschoolers with developmental delay. *Research in Developmental Disabilities, 32*, 289-296.
Limbos, M. M., & Joyce, D. P. (2011). Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *Journal of Developmental & Behavioral Pediatrics, 32*, 499-511.

Lung, F. W., Shu, B. C., Chiang, T. L., & Lin, S. J. (2009). Twin-singleton influence on infant development: a national birth cohort study. *Child: Care, Health and Development, 35*, 409-418.
McLean, M., McCormick, K., & Baird, S. (1991). Concurrent validity of the Griffiths' Mental Develpment Scales with a population of children under 24 months. *Journal of Early Intervention, 15*, 338 - 344.
McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *Am Psychol, 53*, 185-204.
Mousiolis, A., Baroutis, G., Sindos, M., Costalos, C., & Antsaklis, A. (2013). Maternal age as a predictive factor of pre-term birth. An epidemiological study from 1999 to 2008 in Greece. *J Obstet Gynaecol, 33*, 28-31.
Murray, G. K., Veijola, J., Moilanen, K., Miettunen, J., Glahn, D. C., Cannon, T. D., Jones, P. B., & Isohanni, M. (2006). Infant motor development is associated with adult cognitive categorisation in a longitudinal birth cohort study. *Journal of Child Psychology & Psychiatry, 47*, 25-29.
Odd, D. E., Rasmussen, F., Gunnell, D., Lewis, G., & Whitelaw, A. (2008). A cohort study of low Apgar scores and cognitive outcomes. *Archives of Disease in Childhood - Fetal Neonatal Edition, 93*, F115-120.
Piek, J. P., Dawson, L., Smith, L. M., & Gasson, N. (2008). The role of early fine and gross motor development on later motor and cognitive ability. *Human Movement Science, 27*, 668 - 681.
Pike, A., Iervolino, A. C., Eley, T. C., Price, T., & Plomin, R. (2006). Environmental risk and young children's cognitive and behavioral development. *International Journal of Behavioral Development, 30*, 55 - 66.
Ronalds, G. A., De Stavola, B. L., & Leon, D. A. (2005). The cognitive cost of being a twin: evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *British Medical Journal, 331*, 1306.
Rossi, A., Mullin, P., & Chmait, R. (2011). Neonatal outcomes of twins according to birth order, presentation and mode of delivery: a systematic review and meta-analysis*. *BJOG, 118*, 523-532.
Roze, E., Meijer, L., Van Braeckel, K. N., Ruiter, S. A., Bruggink, J. L., & Bos, A. F. (2010). Developmental trajectories from birth to school age in healthy term-born children. *Pediatrics, 126*, e1134-1142.
Rutter, M., & Redshaw, J. (1991). Annotation: growing up as a twin: twin-singleton differences in psychological development. *Journal of Child Psychology & Psychiatry, 32*, 885-895.
Rutter, M., Thorpe, K., Greenwood, R., Northstone, K., & Golding, J. (2003). Twins as a natural experiment to study the causes of mild language delay: I: Design; twin-singleton differences in language, and obstetric risks. *Journal of Child Psychology & Psychiatry, 44*, 326-341.
Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*, 532-545.
Schonhaut, L., Armijo, I., Schonstedt, M., Alvarez, J., & Cordero, M. (2013). Validity of the ages and stages questionnaires in term and preterm infants. *Pediatrics, 131*, e1468-1474.
Sherlock, R. L., Synnes, A. R., & Koehoorn, M. (2008). Working mothers and early childhood outcomes: lessons from the Canadian National Longitudinal Study on Children and Youth. *Early Human Development, 84*, 237-242.
Siegel, L. S., Saigal, S., Rosenbaum, P., Morton, R. A., Young, A., Berenbaum, S., & Stoskopf, B. (1982). Predictors of development in preterm and full-term

infants: a model for detecting the at risk child. *Journal of Pediatric Psychology, 7,* 135-148.
Smyth, R. M., Spark, P., Armstrong, N., & Duley, L. (2009). Magpie Trial in the UK: methods and additional data for women and children at 2 years following pregnancy complicated by pre-eclampsia. *BMC Pregnancy and Childbirth, 9,* 15.
Squires, J. T., Bricker, D., & Potter, L. (2009). *Excerpts from ASQ-3 User's Guide*: Brookes Publishing Co. Inc.
StataCorp. (2009). Stata Statistical Software: Release 11. In T. College Station (Ed.): StataCorp LP.
Sullivan, M. C., & Margaret, M. M. (2003). Perinatal morbidity, mild motor delay, and later school outcomes. *Developmental Medicine & Child Neurology, 45,* 104-112.
Sutcliffe, A. G., & Derom, C. (2006). Follow-up of twins: health, behaviour, speech, language outcomes and implications for parents. *Early Human Development, 82*, 379-386.
The Magpie Trial Follow Up Study Management, G., & The Magpie Trial Follow Up Study Collaborative, G. (2004). The Magpie Trial follow up study: outcome after discharge from hospital for women and children recruited to a trial comparing magnesium sulphate with placebo for pre-eclampsia [ISRCTN86938761]. *BMC Pregnancy and Childbirth, 4,* 5.
Thorpe, K. (2006). Twin children's language development. *Early Human Development, 82,* 387-395.
Thun-Hohenstein, L., Largo, R. H., Molinari, L., Kundu, S., & Duc, G. (1991). Early fine motor and adaptive development in high-risk appropriate for gestational age preterm and healthy term children. *European Journal of Pediatrics, 150,* 562-569.
van Haastert, I. C., de Vries, L. S., Helders, P. J., & Jongmans, M. J. (2006). Early gross motor development of preterm infants according to the Alberta Infant Motor Scale. *Journal of Pediatrics, 149,* 617-622.
Waldfogel, J., Han, W. J., & Brooks-Gunn, J. (2002). The effects of early maternal employment on child cognitive development. *Demography, 39,* 369-392.
Wechsler, D. (2003). *Wechsler intelligence scale for children - fourth edition: technical and interpretative manual*. San Antonio, TX: The Psychological Corporation.
Yu, L. M., Hey, E., Doyle, L. W., Farrell, B., Spark, P., Altman, D. G., Duley, L., & Magpie Trial Follow-Up Study Collaborative, G. (2007). Evaluation of the Ages and Stages Questionnaires in identifying children with neurosensory disability in the Magpie Trial follow-up study. *Acta Paediatrica, 96*, 1803-1808.

Table 1. Means and z-scores for Ages and Stages Questionnaires subscales[a] at each follow-up between 3 and 24 months.

	Communication	Gross motor	Fine motor	Problem solving	Personal-social
3 months (N = 66)					
Mean (SD)	42.82 (13.61)	40.15 (11.67)	28.47 (14.28)	30.46 (14.97)	35.31 (14.97)
z-score (SD)	-1.18 (1.54)*	-1.79 (1.44)*	-2.11 (1.30)*	-2.48 (1.59)*	-1.77 (1.60)*
6 months (N = 96)					
Mean (SD)	46.00 (11.25)	29.74 (14.74)	38.75 (17.49)	42.97 (15.54)	36.35 (16.25)
z-score (SD)	-0.30 (1.17)*	-1.36 (1.26)*	-0.86 (1.47)*	-0.66 (1.37)*	-1.04 (1.42)*
9 months (N = 90)					
Mean (SD)	35.83 (14.65)	29.39 (16.74)	47.00 (15.19)	41.50 (16.83)	39.78 (13.34)
z-score (SD)	-0.22 (1.19)*	-1.20 (1.16)*	-0.51 (1.45)*	-0.77 (1.62)*	-0.23 (1.13)*
12 months (N = 100)					
Mean (SD)	44.15 (14.32)	36.65 (21.73)	50.50 (12.30)	45.66 (14.70)	43.15 (14.71)
z-score (SD)	0.07 (1.04)	-0.93 (1.53)*	-0.19 (1.39)	-0.31 (1.36)*	-0.22 (1.23)*
18 months (N = 67)					
Mean (SD)	40.30 (15.05)	54.03 (12.92)	53.06 (10.37)	44.40 (12.81)	46.04 (11.20)
z-score (SD)	-0.14 (1.03)	-0.16 (1.43)	0.07 (1.14)	-0.16 (1.26)	-0.18 (1.08)
24 months (N = 45)					
Mean (SD)	51.44 (14.45)	54.22 (12.48)	52.00 (9.32)	46.67 (13.78)	46.67 (11.92)
z-score (SD)	0.02 (1.11)	-0.06 (1.50)	0.04 (1.13)	-0.28 (1.40)	-0.46 (1.22)

Raw score for all ASQ-3 subscales range from 0 to 60, z-scores are based on singleton population means and standard deviations for each age and subscale. Significant differences with the norm scores are displayed with *p<0.001. N=sample size, SD=standard deviation.

Table 2. Expected increase in mean Ages and Stages z-scores for each additional kilo (kg) in birth weight, year in maternal age and point in Apgar score, displayed for each subscale at each follow-up between 3 and 24 months.

	3 months (N = 66)	6 months (N = 96)	9 months (N = 90)	12 months (N = 100)	18 months (N = 67)	24 months (N = 45)
	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)
Communication						
Birth weight (kg)	0.70 (-0.20 - 1.61)	0.18 (-0.42 - 0.77)	0.40 (-0.17 - 0.98)	-0.22 (-0.76 - 0.33)	0.24 (-0.44 - 0.91)	0.31 (-0.65 - 1.27)
Maternal age (years)	-0.02 (-0.10 - 0.06)	-0.01 (-0.05 - 0.04)	-0.08 (-0.13 - -0.03)**	-0.06 (-0.1 - -0.03)***	-0.05 (-0.09 - 0.00)*	-0.06 (-0.12 - -0.01)**
Apgar at 5 minutes	-0.48 (-0.99 - 0.03)	-0.15 (-0.42 - 0.11)	0.04 (-0.13 - 0.40)	0.21 (-0.02 - 0.43)	0.17 (-0.13 - 0.46)	0.26 (-0.21 - 0.73)
Gross motor						
Birth weight (kg)	0.25 (-0.57 - 1.08)	0.35 (-0.31 - 1.00)	0.35 (-0.24 - 0.93)	0.00 (-0.81 - 0.82)	0.18 (-0.61 - 0.97)	0.32 (-0.89 - 1.53)
Maternal age (years)	-0.02 (-0.10 - 0.05)	0.03 (-0.02 - 0.08)	-0.05 (-0.10 - 0.01)	-0.06 (-0.11 - -0.01)**	-0.05 (-0.10 - 0.00)	-0.05 (-0.12 - 0.03)
Apgar at 5 minutes	0.19 (-0.27 - 0.66)	-0.08 (-0.37 - 0.22)	0.08 (-0.28 - 0.45)	0.14 (-0.20 - 0.48)	0.81 (0.51 - 1.11)***	0.90 (0.34 - 1.46)**
Fine motor						
Birth weight (kg)	-0.03 (-0.78 - 0.72)	0.35 (-0.33 - 1.04)	0.15 (-0.51 - 0.80)	0.14 (-0.57 - 0.85)	0.23 (-0.51 - 0.97)	-0.16 (-0.94 - 0.63)
Maternal age (years)	-0.05 (-0.11 - 0.02)	-0.03 (-0.09 - 0.02)	-0.06 (-0.12 - 0.00)*	-0.05 (-0.09 - 0.00)*	-0.05 (-0.09 - 0.00)	-0.05 (-0.10 - 0.00)*
Apgar at 5 minutes	0.15 (-0.27 - 0.58)	-0.01 (-0.32 - 0.29)	0.15 (-0.26 - 0.55)	0.20 (-0.09 - 0.48)	0.09 (-0.24 - 0.42)	0.45 (0.07 - 0.83)*
Problem solving						
Birth weight (kg)	-0.33 (-1.21 - 0.54)	0.27 (-0.33 - 0.87)	0.32 (-0.39 - 1.04)	-0.01 (-0.77 - 0.74)	-0.13 (-0.95 - 0.69)	0.26 (-0.75 - 1.26)
Maternal age (years)	-0.03 (-0.12 - 0.05)	-0.03 (-0.08 - 0.02)	-0.09 (-0.15 - -0.02)**	-0.03 (-0.08 - 0.01)	-0.03 (-0.09 - 0.02)	-0.07 (-0.13 - -0.01)**
Apgar at 5 minutes	0.37 (-0.13 - 0.86)	0.03 (-0.24 - 0.30)	-0.10 (-0.54 - 0.33)	-0.07 (-0.39 - 0.24)	0.08 (-0.29 - 0.44)	0.50 (0.02 - 0.98)*
Personal-social						
Birth weight (kg)	-0.32 (-1.22 - 0.58)	0.67 (0.02 - 1.32)*	0.23 (-0.29 - 0.75)	-0.18 (-0.82 - 0.46)	-0.19 (-0.97 - 0.60)	0.52 (-0.45 - 1.49)
Maternal age (years)	-0.09 (-0.18 - -0.01)*	-0.02 (-0.07 - 0.03)	-0.07 (-0.11 - -0.02)**	-0.04 (-0.08 - 0.00)*	-0.02 (-0.06 - 0.03)	-0.05 (-0.11 - 0.00)

Apgar at 5 minutes	0.42 (-0.08 - 0.93)	-0.01 (-0.31 - 0.30)	0.08 (-0.24 - 0.41)	0.18 (-0.09 - 0.45)	-0.05 (-0.39 - 0.30)	0.10 (-0.40 - 0.60)

All results are adjusted for gestational age at birth and gender. *p<0.05, **p<0.01, ***p<0.001. Coef. (95% CI)=regression coefficient (95% confidence interval).

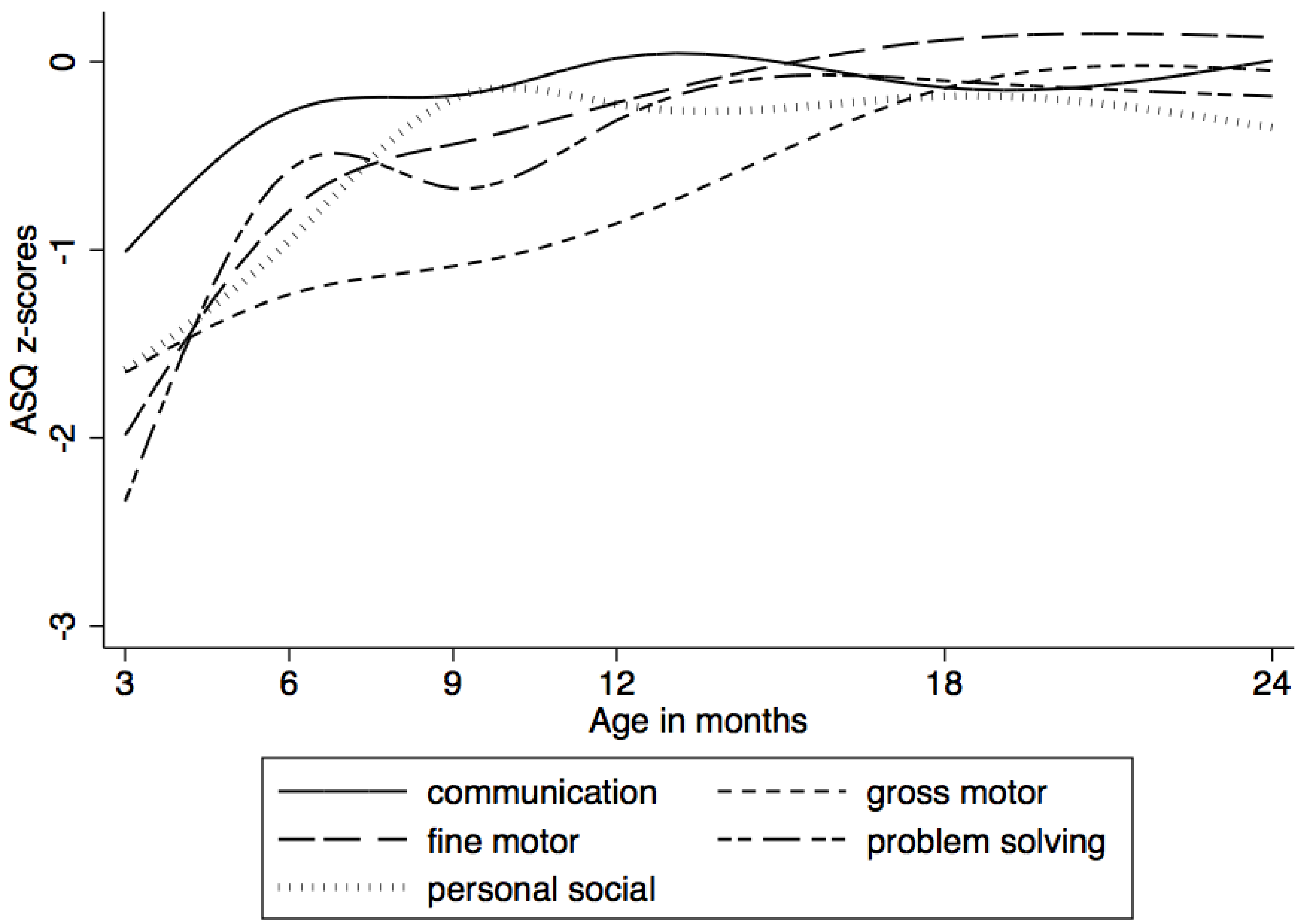

Figure 1. Developmental trajectories of healthy twins from 3 to 24 months, as assessed by the Ages and Stages Questionnaires.
(Footnote for Figure 1: Based on results from a multilevel linear regression adjusted for gestational age at birth, gender and maternal age.)