# The Heritability of Gestational Age in a Two-million Member Cohort: Implications for Spontaneous Preterm Birth

**Wilfred Wu**[1], **David J. Witherspoon**[1], **Alison Fraser**[2], **Erin A. S. Clark**[3], **Alan Rogers**[4], **Gregory J. Stoddard**[5], **Tracy A. Manuck**[3], **Karin Chen**[6], **M. Sean Esplin**[3], **Ken R. Smith**[2], **Michael W. Varner**[3], and **Lynn B. Jorde**[1]

[1]Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[2]Population Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[3]Department of Obstetrics and Gynecology, University of Utah School of Medicine, and, Maternal-Fetal Medicine, Intermountain Healthcare, Salt Lake City, UT, USA

[4]Department of Anthropology, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[5]Study Design and Biostatistics Center, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[6]Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

## Abstract

Preterm birth (PTB), defined as birth prior to a gestational age (GA) of 37 completed weeks, affects more than 10% of births worldwide. PTB is the leading cause of neonatal mortality and is associated with a broad spectrum of lifelong morbidity in survivors. The etiology of spontaneous PTB (SPTB) is complex and has an important genetic component. Previous studies have compared monozygotic and dizygotic twin mothers and their families to estimate the heritability of SPTB, but these approaches cannot separate the relative contributions of the maternal and the fetal genomes to GA or SPTB. Using the Utah Population Database, we assessed the heritability of GA in more than 2 million post-1945 Utah births, the largest familial GA dataset ever assembled. We estimated a narrow-sense heritability of 13.3% for GA and a broad-sense heritability of 24.5%. A maternal effect (which includes the effect of the maternal genome) accounts for 15.2% of the variance of GA, and the remaining 60.3% is contributed by individual environmental effects. Given the relatively low heritability of GA and SPTB in the general population, multiplex SPTB pedigrees are likely to provide more power for gene detection than will samples of unrelated individuals. Furthermore, nongenetic factors provide important targets for therapeutic intervention.

Correspondence: Lynn B. Jorde, lbj@genetics.utah.edu, Phone: 801-581-4566, FAX: 801-585-9148.

**Conflict of interest**
The authors declare that there are no conflicts of interest.

## Keywords

Quantitative genetics; Heritability; Complex trait/complex disease; Preterm birth; Preterm labor

## Introduction

Preterm birth (PTB), which is defined as birth prior to 37 weeks of gestational age (GA), is a major public health problem that affects more than 10% of all births worldwide [1, 2] and 11.7% of all births in United States [3]. Globally, an estimated 15 million babies are born pre-term each year [1, 2]. Although PTB affects all socioeconomic strata, rates vary significantly within and across different income groups [1, 2], implying that there are socioeconomic as well as genetic influences. Most importantly, PTB is the leading cause of neonatal mortality in otherwise unaffected newborns [1, 3–5], and premature infants who survive the neonatal period are more likely to be left with a broad spectrum of lifelong morbidity, including neurodevelopmental delay, cerebral palsy, blindness, deafness, and chronic lung disease [6–8]. Some PTBs are iatrogenic (i.e., they can be attributed to obstetric or medical interventions). Non-iatrogenic PTBs are known as spontaneous PTBs (SPTB). The etiology of SPTB is complex and multifactorial, and genetic factors are thought to be important causal components [8–12].

Twin studies have been used previously to estimate the heritability SPTB [13–15], using mothers of SPTBs as index cases. In this approach, a "maternal" SPTB phenotype is designated when the mother has a child born before 37 weeks' GA. Thus, SPTB is a diagnosis that applies a threshold to the underlying quantitative trait, GA [16]. In some of these studies, the mother's phenotype was defined as SPTB based on the GA of the first-born offspring [13–15]; in other studies, a mother was assigned the SPTB phenotype if any of her offspring had an early GA [13, 14]. Typically, concordance rates for SPTB in monozygotic versus dizygotic twin mothers have been used to estimate heritability, and these estimates ranged from 19% to 34% [13–15].

Twin studies have several well-known biases and limitations, including potential overestimation of heritability [17–20]. Although both maternal and fetal genomes have been shown to influence the risk of SPTB [12, 21–29], twin-based designs using mothers as index cases cannot separate their effects. In addition, the use of dichotomous traits, such as PTB, usually offers less power to detect underlying causal loci than does the use of the full distribution of a quantitative trait such as GA [30]. To provide guidance for future gene discovery studies of PTB and GA, we focus here on the heritability of GA, using more than 2 million singleton births derived from the Utah Population Database (UPDB). By designating offspring rather than mothers as index cases and assessing the correlation of GA between offspring and their full and half-siblings, we were able to distinguish the potential contributions of inherited genes, maternal and paternal effects, and environmental effects to GA.

## Materials and Methods

The central component of the UPDB (http://www.huntsmancancer.org/groups/ppr) is an extensive set of Utah family histories, with many genealogies extending eight to ten generations. The 7.3 million individuals represented in the UPDB are linked to medical and demographic information, including hospital records, the Utah Cancer Registry, and birth and death certificates. SPTB was analyzed using each individual's GA, which was calculated based on the date of the mother's last menstrual period (LMP). We use this measure because it is consistent across the full span of records we analyzed, from 1945 through 2010. Clinical estimates (based on ultrasound or other clinical judgment) were only used when the date of last menses was missing or incomplete or if it generated an out-of-range gestational age. In 86% of cases, the clinical estimate of GA was identical to the LMP-based GA, implying that the clinical estimate records are usually based on LMP. There is an excess of LMP dates that fall on certain days of the month (1, 5, 10, 15, 20, 25, 28, with a slight dearth on days 29 and 31), presumably since those dates are used when the actual date is not exactly recalled. The excess on these days compared to the average on other days accounts for approximately 13% of records. With this extensive database, we were able to estimate the heritability, maternal effect, and environmental contributions to GA by designating offspring as index cases and applying standard analytic methods [17, 18]. Statistical analyses were carried out using STATA Statistical Software: Release 12 [31]. This study was approved by University of Utah's Institutional Review Board (IRB number: 00015034).

## Results

### Data Overview

We assessed GA in 2,158,880 post-1945 singleton births in the state of Utah. We first excluded conditions in which the births were unlikely to have occurred spontaneously, including major congenital anomalies in the fetus and high-risk obstetric conditions in mothers (Online Resource 1). The records of all fetal diagnoses were visually inspected to ensure appropriate inclusion or exclusion. In total, 374,708 iatrogenic births were identified and removed, leaving 1,784,172 non-iatrogenic births. We did not exclude stillbirths with no known anomalies (0.2% of all births.) The distribution of GA for these non-iatrogenic births is unimodal, with an average of 39.11 weeks and standard deviation of 2.11 weeks (Online Resource 2).

When SPTB is defined as a GA less than 37 weeks, the overall rate of SPTB in our sample is 7.6%. The full-sibling recurrence risk of SPTB, given one prior SPTB in the sibship, is 38.3%. Using a cutoff of 34 weeks GA to define SPTB, the overall SPTB rate is 2.0%, with a sibling risk of 33.4%.. The SPTB rate was 6.2% during the first decade following 1945, and it increased to 11.4% after 2005. In the thirty-year period following 1980, when ultrasound came to be used commonly in estimating GA, the SPTB rate increased steadily, from 6.2% in 1980 to 10.9% in 2010. This is consistent with national trends [8].

## Parent-offspring comparisons

Heritability can be estimated by comparing the offspring's GA with those of the parents [17, 18]. In our data set, more than one million parent-offspring pairs are available for comparison (Table 1). The heritability estimated from regression analysis of GA for father-offspring pairs is 6.77%, with a 95% confidence interval (CI) of 6.06% – 7.49% (Table 1). The heritability estimated from mother-offspring regression is 14.21% (95% CI 13.52% – 14.90%) (Table 1). Further sex-limited analysis did not show significant differences among daughters and sons (Online Resource 3).

Quantile-quantile analysis (Fig. 1) revealed a generation effect in which both paternal-offspring and maternal-offspring QQ plots showed a pattern with skewing below the diagonal. More specifically, before the GA reaches 40 weeks (term pregnancy), the GA in the parental generation is consistently larger than GA in the offspring for each percentile. There are two plausible and non-exclusive causes for this bias: more early induced labor and Caesarean section births in the recent offspring cohorts; and low survival and reproduction of preterm-born persons in older parental cohorts. There is some evidence that preterm birth is associated not just with neonatal mortality, but with higher childhood and adult mortality as well as decreased reproduction [32, 33]. Regardless of the cause, this difference between two generations resulted in a bias towards individuals with larger GA as parents, which can introduce bias in the heritability estimate when using a parent-offspring approach [17, 18, 34].

## Sibling analysis

To avoid inter-generational effects, we focused our analysis on sibling correlations. The sibling analysis incorporates the effects of dominance variance and common environment, such as a maternal effect [17, 18; see Table 2]. We compared intra-class correlations among full siblings, paternal half siblings, and maternal half siblings. In our dataset, there are more than 1.27 million individuals with a full sibling, grouped under 433,150 father-mother pairs (Table 3). The intra-class correlation (ICC) for full-sib GA is 0.2467 (bootstrapped 95% CI 0.2436 – 0.2498) (Table 3).

We further dissected the genetic and environmental components by analyzing half siblings. There are 151,047 individuals in the UPDB who have a paternal half-sib, with 90,509 mothers and 43,563 fathers (Table 3). The ICC for paternal half-sib GA is 0.0333 (95% CI = 0.0274 – 0.0392) (Table 3). There are 144,527 individuals in the UPDB who have a maternal half sib, with 99,635 fathers and 47,854 mothers (Table 3). The ICC for maternal half-sib GA is 0.1855 (95% CI 0.1776 – 0.1936) (Table 3).

We next partitioned out the different components that contribute to variation in GA (Table 2). Narrow-sense heritability is defined as the proportion of phenotypic variance ($V_P$) variance due to additive genetic variance ($V_A$), which can be calculated as four times the paternal half-sib correlation [17, 18]. This estimate was 13.33% (95% CI 10.87% – 15.79%) (Table 2). The difference between the maternal and paternal half-sib correlation is the maternal effect, which is 15.23% (95% CI 14.26% – 16.19%) (Table 2). The maternal genome, the mitochondrial genome, differential maternal or paternal imprinting, or other

non-genetic maternal factors that are constant over the time can contribute to the maternal effect. Since the maternal womb is the only environment that a newborn has been exposed to when GA is measured, the maternal effect also represents the proportion of common environmental variance ($V_C$; Table 2). Segregation analysis using multiplex families with 2 or more SPTB probands also supports the importance of the mother's genome or maternally-inherited genes in influencing GA [28]. The proportion of dominance genetic variance ($V_D$, if we assume no epistasis, which cannot be separated from dominance variance in this design) can then be calculated from the full-sib correlation as 11.12% (95% CI 6.69% – 15.55%; Table 2). The remaining unexplained 60.33% of the variance (95% CI 57.47% – 63.19%) is the proportion of individual environmental variance ($V_E$; Table 2).

## Discussion

For many quantitative traits, such as human height, the great majority of genetic variance is accounted for by the additive component, $V_A$ [35–37]. In contrast, our analysis of GA shows that $V_A$ is relatively small (13.1%) and is similar to the dominance variance, $V_D$, of 11.1%. This may reflect the action of natural selection, which reduces additive variance but can increase dominance variance [34, 38–40]. There is evolutionary evidence for natural selection on SPTB [28]. Our inter-generational data (Fig. 1) also suggest selection against SPTB, although the increasing prevalence of induced labor and Caesarean section births may contribute to that pattern.

Given that the narrow-sense heritability of GA is 13.33% and the $V_D$ is 11.12%, the broad-sense heritability is 24.45%. Previous correlation studies of GA in mother-offspring pair and in sibs also derived relatively low heritability estimates [41, 42]. These estimates imply that studies of unrelated SPTB cases and controls may have limited power to detect causal loci, even if the sample sizes are very large [34, 43, 44]. Attempts over the past several decades to isolate specific genes associated with SPTB have not been reproducible, consistent with the idea that SPTB is a final common outcome of multiple etiologies. It is also likely that a dose-dependent effect exists, consistent with the concept that SPTB is the end result of interactions between the components of multiple biologic systems[45]. The use of high-risk pedigrees with multiple familial cases may increase the likelihood of identifying causal genes and should be an important target of future genetic analyses of SPTB.

The largest component (60.33%) contributing to GA is $V_E$, the individual environment. Factors that may be different from individual to individual, and birth to birth, such as infection and inflammation, parental socio-economic status, stress, nutrition, and prenatal exposure to tobacco, alcohol, or other environmental pollutants, may play a role in determining SPTB risk [9, 12, 46–49].

The nature of our data imposes some limitations on our analyses. We used the date of last menstrual period to estimate gestational age because it is the only measure that is consistently reported across the long time span of our records (see Materials and Methods.) Although there can be no doubt that the advent of obstetric ultrasound has substantially improved gestational age assessment, it was not universally utilized until the 1980s. The error variance introduced by LMP-based GA estimates will appear as non-genetic variance,

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

causing our estimates of the heritability of GA to be conservative. In addition, since we do not have data on environmental variables that might affect preterm birth (e.g. tobacco smoking, nutrition, socioeconomic status, and others [50, 51]), nor genome-wide genotypes for these individuals, we cannot estimate the interaction effects of genes with the environment or genes with other genes (epistasis; although interactions of alleles within a locus are estimated as dominance genetic variance) [17]. To the extent that a mother's smoker status, nutritional level, and similar environmental variables remain constant across offspring, these effects should be captured in the maternal effects, and should not affect the heritability we estimate.

Future genetic studies of SPTB thus need to control for environmental risk factors. At the same time, population-based efforts to modify environmental risk factors may have the biggest impact on SPTB, and may inform prevention and intervention strategies. Our findings suggest that family-based studies, controlling for non-genetic risk factors, may optimize our ability to better characterize the etiology of SPTB

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## References

1. Reich ES. Pre-term births on the rise. Nature. 2012; 485:20. [PubMed: 22552070]

2. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet. 2012; 379:2162–72. [PubMed: 22682464]

3. Hamilton BE, Hoyert DL, Martin JA, Strobino DM, Guyer B. Annual summary of vital statistics: 2010–2011. Pediatrics. 2013; 131:548–58. [PubMed: 23400611]

4. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2003 period linked birth/infant death data set. National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System. 2006; 54:1–29.

5. Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012; 379:2151–61. [PubMed: 22579125]

6. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. The New England journal of medicine. 2000; 343:378–84. [PubMed: 10933736]

7. Damus K. Prevention of preterm birth: a renewed national priority. Current opinion in obstetrics & gynecology. 2008; 20:590–6. [PubMed: 18989136]

8. Muglia LJ, Katz M. The enigma of spontaneous preterm birth. The New England journal of medicine. 2010; 362:529–35. [PubMed: 20147718]

9. Collins JW Jr, David RJ, Handler A, Wall S, Andes S. Very low birthweight in African American infants: the role of maternal exposure to interpersonal racial discrimination. American journal of public health. 2004; 94:2132–8. [PubMed: 15569965]

10. Kistka ZA, Palomar L, Lee KA, Boslaugh SE, Wangler MF, Cole FS, et al. Racial disparity in the frequency of recurrence of preterm birth. American journal of obstetrics and gynecology. 2007; 196:131, e1–6. [PubMed: 17306652]

11. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008; 371:75–84. [PubMed: 18177778]

12. Wu W, Clark EA, Stoddard GJ, Watkins WS, Esplin MS, Manuck TA, et al. Effect of interleukin-6 polymorphism on risk of preterm birth within population strata: a meta-analysis. BMC genetics. 2013; 14:30. [PubMed: 23617681]

13. Clausson B, Lichtenstein P, Cnattingius S. Genetic influence on birthweight and gestational length determined by studies in offspring of twins. BJOG: an international journal of obstetrics and gynaecology. 2000; 107:375–81. [PubMed: 10740335]

14. Treloar SA, Macones GA, Mitchell LE, Martin NG. Genetic influences on premature parturition in an Australian twin sample. Twin research: the official journal of the International Society for Twin Studies. 2000; 3:80–2. [PubMed: 10918619]

15. Kistka ZA, DeFranco EA, Ligthart L, Willemsen G, Plunkett J, Muglia LJ, et al. Heritability of parturition timing: an extended twin design analysis. American journal of obstetrics and gynecology. 2008; 199:43, e1–5. [PubMed: 18295169]

16. Kieler H, Axelsson O, Nilsson S, Waldenstrom U. The length of human pregnancy as calculated by ultrasonographic measurement of the fetal biparietal diameter. Ultrasound in obstetrics & gynecology: the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 1995; 6:353–7.

17. Falconer, DS.; Mackay, TFC. Introduction to quantitative genetics. Essex, England: Longman; 1996.

18. Lynch, M.; Walsh, B. Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer; 1998.

19. Rahman I, Bennet AM, Pedersen NL, de Faire U, Svensson P, Magnusson PK. Genetic dominance influences blood biomarker levels in a sample of 12,000 Swedish elderly twins. Twin research and human genetics: the official journal of the International Society for Twin Studies. 2009; 12:286–94. [PubMed: 19456221]

20. Zaitlen N, Kraft P, Patterson N, Pasaniuc B, Bhatia G, Pollack S, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS genetics. 2013; 9:e1003520. [PubMed: 23737753]

21. Porter TF, Fraser AM, Hunter CY, Ward RH, Varner MW. The risk of preterm birth across generations. Obstetrics and gynecology. 1997; 90:63–7. [PubMed: 9207815]

22. Winkvist A, Mogren I, Hogberg U. Familial patterns in birth characteristics: impact on individual and population risks. International journal of epidemiology. 1998; 27:248–54. [PubMed: 9602406]

23. Li DK. Changing paternity and the risk of preterm delivery in the subsequent pregnancy. Epidemiology. 1999; 10:148–52. [PubMed: 10069250]

24. Vatten LJ, Skjaerven R. Effects on pregnancy outcome of changing partner between first two births: prospective population study. Bmj. 2003; 327:1138. [PubMed: 14615338]

25. Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. American journal of epidemiology. 2007; 165:734–41. [PubMed: 17311798]

26. Plunkett J, Muglia LJ. Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. Annals of medicine. 2008; 40:167–95. [PubMed: 18382883]

27. Wilcox AJ, Skjaerven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. American journal of epidemiology. 2008; 167:474–9. [PubMed: 18048376]

28. Plunkett J, Feitosa MF, Trusgnich M, Wangler MF, Palomar L, Kistka ZA, et al. Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. Human heredity. 2009; 68:209–19. [PubMed: 19521103]

29. Haataja R, Karjalainen MK, Luukkonen A, Teramo K, Puttonen H, Ojaniemi M, et al. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. PLoS genetics. 2011; 7:e1001293. [PubMed: 21304894]

30. Williams JT, Blangero J. Power of variance component linkage analysis-II. Discrete traits. Annals of human genetics. 2004; 68:620–32. [PubMed: 15598220]

31. Stata Corp. STATA Statistical Software: Release 12. College Station, TX: 2011.

32. Swamy GK, Ostbye T, Skjaerven R. Association of preterm birth with long-term survival, reproduction, and next-generation preterm birth. Jama. 2008; 299:1429–36. [PubMed: 18364485]

33. Crump C, Sundquist K, Winkleby MA, Sundquist J. Early-term birth (37–38 weeks) and mortality in young adulthood. Epidemiology. 2013; 24:270–6. [PubMed: 23337240]

34. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era--concepts and misconceptions. Nature reviews Genetics. 2008; 9:255–66.

35. Visscher PM, Macgregor S, Benyamin B, Zhu G, Gordon S, Medland S, et al. Genome partitioning of genetic variation for height from 11,214 sibling pairs. American journal of human genetics. 2007; 81:1104–10. [PubMed: 17924350]

36. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. PLoS genetics. 2008; 4:e1000008. [PubMed: 18454194]

37. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. Nature reviews Genetics. 2013; 14:139–49.

38. Harris DL. Genotypic Covariances between Inbred Relatives. Genetics. 1964; 50:1319–48. [PubMed: 14239792]

39. Bulmer MG. The Effect of Selection on Genetic Variability. The American Naturalist. 1971; 105:201–11.

40. Shaw RG, Byers DL, Shaw FH. Genetic components of variation in Nemophila menziesii undergoing inbreeding: morphology and flowering time. Genetics. 1998; 150:1649–61. [PubMed: 9832540]

41. Karn MN, Lang-Brown H, MacKenzie, Penrose LS. Birth weight, gestation time and survival in sibs. Annals of eugenics. 1951; 15:306–22. [PubMed: 24540311]

42. Magnus P, Bakketeig LS, Skjaerven R. Correlations of birth weight and gestational age across generations. Annals of human biology. 1993; 20:231–8. [PubMed: 8489198]

43. Myking S, Boyd HA, Myhre R, Feenstra B, Jugessur A, Devold Pay AS, et al. X-chromosomal maternal and fetal SNPs and the risk of spontaneous preterm delivery in a Danish/Norwegian genome-wide association study. PloS one. 2013; 8:e61781. [PubMed: 23613933]

44. Wu W, Clark EA, Manuck TA, Esplin MS, Varner MW, Jorde LB. A Genome-Wide Association Study of spontaneous preterm birth in a European population. F1000Research. 2013

45. Hood L, Tian Q. Systems approaches to biology and disease enable translational systems medicine. Genomics, proteomics & bioinformatics. 2012; 10:181–5.

46. Watts DH, Krohn MA, Hillier SL, Eschenbach DA. The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. Obstetrics and gynecology. 1992; 79:351–7. [PubMed: 1738513]

47. Howard DL, Marshall SS, Kaufman JS, Savitz DA. Variations in low birth weight and preterm delivery among blacks in relation to ancestry and nativity: New York City, 1998–2002. Pediatrics. 2006; 118:e1399–405. [PubMed: 17079541]

48. Behrman, RE.; SBA. Preterm Birth: Causes, Consequences, and Prevention. Washington, DC: The National Academies Press; 2007.

49. Onderdonk AB, Hecht JL, McElrath TF, Delaney ML, Allred EN, Leviton A, et al. Colonization of second-trimester placenta parenchyma. American journal of obstetrics and gynecology. 2008; 199:52, e1–e10. [PubMed: 18313635]

50. Svensson AC, Sandin S, Cnattingius S, Reilly M, Pawitan Y, Hultman CM, et al. Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. American journal of epidemiology. 2009; 170:1365–72. [PubMed: 19854802]

51. Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal contributions to preterm delivery. American journal of epidemiology. 2009; 170:1358–64. [PubMed: 19854807]
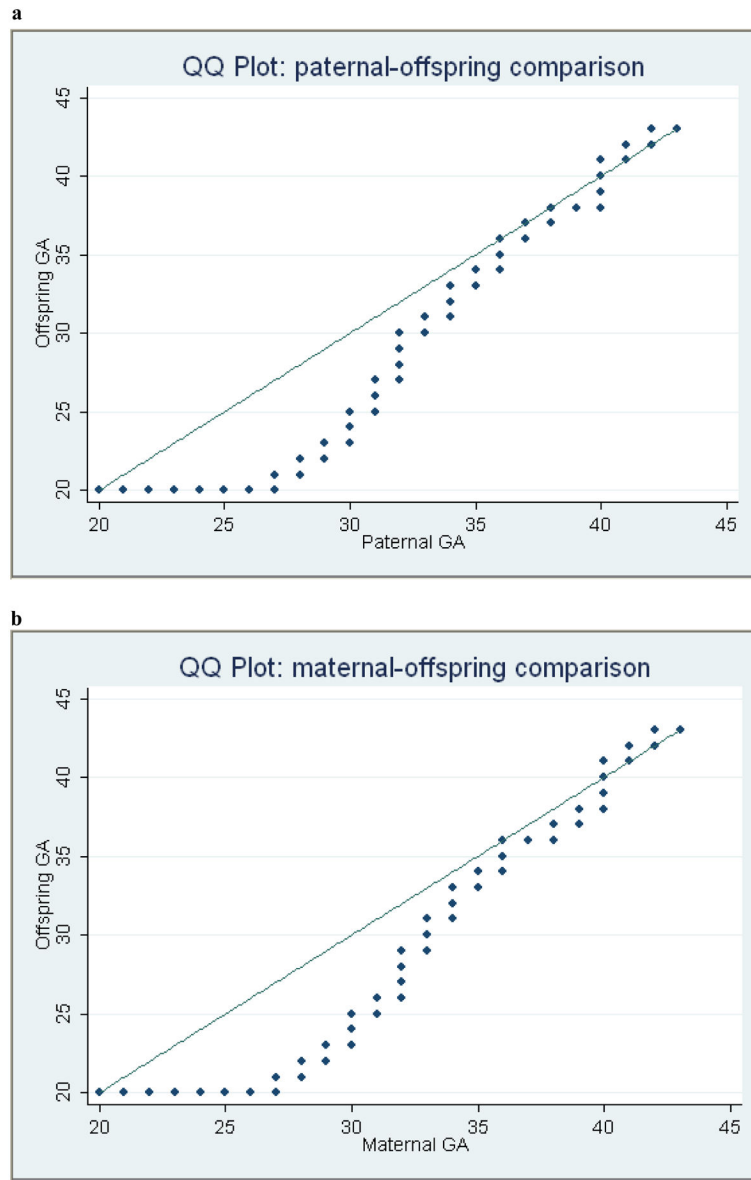
**Figure 1. Quantile-Quantile (QQ) plots for gestational age (GA) across two generations**
(a) Paternal GA versus offspring GA. (b) Maternal GA versus offspring GA. In both panels, GA in the parent is represented on the x- axis and GA of the offspring on the y-axis. Each dot represents the value of (x,y) for different cumulative percentiles. The diagonal line represents the expected y = x relationship. Both the paternal and maternal plots show skewing below the diagonal, resulting in a bias towards individuals with larger GA as parents.

**Table 1**

Sample size, regression coefficient, and heritability estimate by parent-offspring comparison

| Relationship | Number of pairs | beta | $h^2$ | 95% CI for $h^2$ |
|---|---|---|---|---|
| Father-Offspring | 549,777 | 0.0339 | 6.77% | 6.06% – 7.49% |
| Mother-Offspring | 610,253 | 0.0710 | 14.21% | 13.52% – 14.90% |

beta: regression coefficient. $h^2$: heritability. CI: confidence interval

**Table 2**

Equations and partition of components contributing to SPTB.

**A Intraclass covariances for full, maternal half, and paternal half sibs**

$$ICC_{phs} = \frac{1}{4} \cdot \frac{V_A}{V_P}$$

$$ICC_{mhs} = \frac{1}{4} \cdot \frac{V_A}{V_P} + \frac{V_C}{V_P}$$

$$ICC_{fs} = \frac{1}{2} \cdot \frac{V_A}{V_P} + \frac{1}{4} \cdot \frac{V_D}{V_P} + \frac{V_C}{V_P}$$

**B Phenotypic variance components**

$$V_P = V_A + V_D + V_C + V_E$$

$$1 = \frac{V_A}{V_P} + \frac{V_D}{V_P} + \frac{V_C}{V_P} + \frac{V_E}{V_P}$$

$$= \underset{(0.1087,\ 0.1579)}{0.1333} + \underset{(0.0669,\ 0.1555)}{0.1112} + \underset{(0.1426,\ 0.1619)}{0.1523} + \underset{(0.5747,\ 0.6319)}{0.6033}$$

$ICC_{ph}$: covariance of paternal half sibs. $ICC_{mhs}$: covariance of maternal half sibs. $ICC_{fs}$: covariance of full sibs. $V_P$: phenotypic variance. $V_A$: additive genetics variance. $V_D$: dominance genetic variance. $V_C$: common environmental variance. $V_E$: individual environmental variance. In the case of SPTB using newborns as index cases, the maternal environment (or maternal effect) is the common environment. 95% confidence interval by bootstrapping are shown in parentheses on the bottom row.

**Table 3**

Sample size and intra-class correlations of gestational age for three classes of sibling relationships

| Class | Number of fathers | Number of mothers | Number of offspring | ICC | 95% CI |
|---|---|---|---|---|---|
| Full-sibs | 433,150 | 433,150 | 1,279,321 | 0.2467 | 0.2436 – 0.2498 |
| Paternal half-sibs | 43,563 | 90,509 | 151,047 | 0.0333 | 0.0274 – 0.0392 |
| Maternal half-sibs | 99,635 | 47,854 | 144,527 | 0.1855 | 0.1776 – 0.1936 |

ICC: Intra-class correlation. CI: Confidence interval.