Preterm Birth and Mortality and Morbidity
A Population-Based Quasi-experimental Study

Brian M. D'Onofrio, PhD; Quetzal A. Class, BS; Martin E. Rickert, PhD; Henrik Larsson, PhD; Niklas Långström, MD, PhD; Paul Lichtenstein, PhD

IMPORTANCE Preterm birth is associated with increased mortality and morbidity. However, previous studies have been unable to rigorously examine whether confounding factors cause these associations rather than the harmful effects of being born preterm.

OBJECTIVE To estimate the extent to which the associations between early gestational age and offspring mortality and morbidity are the result of confounding factors by using a quasi-experimental design, the sibling-comparison approach, and by controlling for statistical covariates that varied within families.

DESIGN, SETTING, AND PARTICIPANTS A population-based cohort study, combining Swedish registries to identify all individuals born in Sweden from 1973 to 2008 (3 300 708 offspring of 1 736 735 mothers) and link them with multiple outcomes.

MAIN OUTCOMES AND MEASURES Offspring mortality (during infancy and throughout young adulthood) and psychiatric (psychotic or bipolar disorder, autism, attention-deficit/hyperactivity disorder, suicide attempts, substance use, and criminality), academic (failing grades and educational attainment), and social (partnering, parenthood, low income, and social welfare benefits) outcomes through 2009.

RESULTS In the population, there was a dose-response relationship between early gestation and the outcome measures. For example, extreme preterm birth (23-27 weeks of gestation) was associated with infant mortality (odds ratio, 288.1; 95% CI, 271.7-305.5), autism (hazard ratio [HR], 3.2; 95% CI, 2.6-4.0), low educational attainment (HR, 1.7; 1.5-2.0), and social welfare benefits (HR, 1.3; 1.2-1.5) compared with offspring born at term. The associations between early gestation and mortality and psychiatric morbidity generally were robust when comparing differentially exposed siblings and controlling for statistical covariates, whereas the associations with academic and some social problems were greatly or completely attenuated in the fixed-effects models.

CONCLUSIONS AND RELEVANCE The mechanisms responsible for the associations between preterm birth and mortality and morbidity are outcome-specific. Associations between preterm birth and mortality and psychiatric morbidity are largely independent of shared familial confounds and measured covariates, consistent with a causal inference. However, some associations, particularly predicting suicide attempt, educational attainment, and social welfare benefits, are the result of confounding factors. The findings emphasize the importance of both reducing preterm birth and providing wraparound services to all siblings in families with an offspring born preterm.
Preterm birth is associated with increased risk of mortality during infancy and through young adulthood. Shortened gestational age (GA) also predicts offspring morbidity across the lifespan, including psychiatric disorders, academic problems, and social difficulties. Precise estimates of the sequelae of shortened GA are critical for helping physicians and patients balance the benefits and risks of various interventions during pregnancy, and properly understanding the etiologic mechanisms is crucial for designing effective prevention efforts. Most researchers have made strong causal inferences regarding the consequences of early GA. Research suggests that physical and immunologic immaturity account for increased mortality and that brain abnormalities mediate the associations with cognitive and psychiatric problems. However, GA is associated with numerous environmental risks, such as poverty, that are themselves predictive of subsequent difficulties. Family- and twin-based studies also indicate that genetic factors, primarily inherited from the mother, influence GA. Environmental confounding and shared genetic liability, therefore, could account for part or all of the increased mortality and morbidity associated with GA.

Human research has relied primarily on controlling for statistical covariates to account for confounding factors, which provides only qualified support for causal inferences because of the inability to account for unmeasured confounding factors. Randomized studies in humans are impossible and animal studies of parturition have limited generalizability. Researchers, therefore, must use other methods to rule out plausible confounding by genetic and environmental factors. Prestigious scientific working groups in medicine and researchers across a number of other disciplines, including psychiatry, psychology, epidemiology, sociology, and economics, have stressed that quasi-experimental research, studies that use design features to account for confounding factors, play an essential role for drawing strong causal inferences. However, we know of only one study of GA that used such an approach, a sibling comparison study that found an independent association with offspring attention-deficit/hyperactivity disorder (ADHD) medication in a single year.

The aim of the present study was to explore the associations between GA and numerous indices of mortality and morbidity in what we believe to be the largest population-based cohort study of GA to date. We also sought to rigorously rule out confounding factors by comparing differentially exposed siblings to account for all genetic and environmental factors that make siblings similar and controlling for measured covariates that vary within families. Finally, we conducted several sensitivity analyses using various approaches to examine whether assumptions and limitations in the sibling comparison design accounted for the results.

Methods

Study Design
After approval by the institutional review boards at Karolinska Institutet and Indiana University to analyze the de-identified data (for which informed consent was unnecessary), the data for this national cohort were obtained by linking information available in the following population-based registries: (1) the Medical Birth Registry includes data on more than 99% of pregnancies in Sweden since 1973, (2) the Multi-Generation Register contains information about biological relationships for all individuals living in Sweden since 1933, (3) the Migration Register supplies information on dates for migration in or out of Sweden, (4) the Cause of Death Register contains information on dates and causes of all deaths since 1958, (5) the Patient Registry provides diagnoses for all inpatient hospital admissions since 1973 and outpatient care since 2001, (6) the National Crime Register includes detailed information about all criminal convictions since 1973, (7) the National School Register includes grades in all subjects for all students at the end of grade 9 since 1983, (8) the Education Register contains information on highest level of completed formal education through 2008, and (9) the Longitudinal Integration Database for Health Insurance and Social Studies (LISA) contains yearly assessments of income, marital status, social welfare status, and educational level for all individuals aged 15 years or older since 1990. More details on these and additional registries are available from the authors on request.

The present study consisted of singleton offspring born in Sweden between January 1, 1973, and December 31, 2008. Birth-related data for 3,619,712 offspring were obtained from the Swedish Medical Birth Registry. We sequentially removed multiple births (86,273), children with missing data on GA (8290), those with a recorded GA less than 23 weeks (153) or more than 42 weeks and 6 days (41,440), missing maternal identification numbers (4070), invalid or missing sex (2), invalid parity (23), and those who emigrated from Sweden (178,753) during this period. The resulting cohort of 3,300,708 offspring represents 91.2% of all recorded births to a total of 1,736,735 biological mothers. Most offspring had siblings in the data set (2,665,666 [80.8%]); they were in families of mothers (1,101,693) representing 91.2% of all recorded births to a total of 1,736,735 biological mothers. Most offspring had siblings in the data set (2,665,666 [80.8%]); they were in families of mothers (1,101,693) with more than one offspring. Sibling comparisons were made among this subset of the population.

Measures

Gestational Age
The analyses used 2 different representations for GA. For the ordinal representation, children were divided into 5 subgroups: (1) 23 weeks to 27 weeks 6 days, (2) 28 weeks to 30 weeks 6 days, (3) 31 weeks to 33 weeks 6 days, (4) 34 weeks to 36 weeks 6 days, and (5) 37 weeks to 42 weeks 6 days. These groupings are consistent with those of previous studies. For continuous assessment, we converted GA to a linear scale that was referenced at 40 weeks and ranged from -17.0 weeks (raw gestational age, 23 weeks) to +2.9 weeks (42 weeks 6 days).

Offspring Outcomes
Two mortality outcomes were created from the Cause of Death Registry. Infant mortality indexed children who were born alive but died before their first birthday. A separate right-censored variable was used to index mortality after 1 year (up to age 36 years).

Six indices of psychiatric morbidity were modeled. Psychotic or bipolar disorder (up to age 37 years) was measured
as age at the first inpatient hospitalization for schizophrenia, bipolar disorder, or other nonorganic psychotic disorders according to International Classification of Diseases (ICD) Eighth, Ninth, and Tenth Revisions (ICD-8, -9, and -10, respectively) criteria, which are valid indices of these disorders. Autism and ADHD were identified using inpatient and outpatient diagnoses according to ICD-9 and ICD-10 for individuals born between 1980 and 2001 (up to age 19 years). The diagnoses of autism and ADHD have been validated. Age at first suicide attempt (up to age 37 years) was identified using the ICD-8, -9, and -10 codes for any primary or secondary diagnosis for individuals aged 12 years or older in the Patient Registry. Substance use problem (up to age 37 years) was defined as first inpatient hospitalization involving a primary or secondary diagnosis of alcohol or any other nonnicotine substance use disorder for individuals aged 12 years or older. Criminality was indexed by the age at the first occurrence of any criminal conviction (from 15 years, the age of legal responsibility in Sweden, to 37 years). More details about the measurement of psychiatric morbidity are available upon request.

Three indices of academic problems were included. Failing grades indexed poor school performance in grade 9 (when the offspring were approximately aged 15 years), commensurate with a mean failing grade across 16 academic subjects. Highest level of educational attainment was available in the Education Register. Education of less than 10 years was an index of low educational attainment. The higher education group completed 3 or more years of postsecondary education; only individuals born between 1973 and 1983 whose age made it possible to achieve that level were included in the analysis of higher education.

Three indices of social adversity were incorporated, which included assessments of individuals up to age 38 years. First, parenthood was indexed as age when they first became biological parents. Second, whether an individual was ever partnered was based on age at first civil or marital partnership using information recorded in the LISA. Third, the social welfare benefits variable was based on age at first receiving government social welfare subsidies during the previous year in the LISA.

Statistical Analysis
We used Cox survival analyses for right-censored outcomes and logistic regression analyses for dichotomous outcomes. We fitted a series of models for each outcome. All models controlled for offspring sex and birth order, and the logistic models also controlled for offspring year of birth. First, we used the ordinal assessment of GA to provide estimates of increased risks consistent with previous research. Second, we compared a linear and quadratic model using the continuous representation of GA as a baseline model; model selection was based on the Akaike information criterion fit statistic. We refer to this analysis as the baseline model, which estimated the associations between GA and each outcome in the population. Third, we included both offspring-specific (sex, birth order, and year of birth, as well as maternal and paternal age at childbearing) and parental covariates (maternal and paternal highest level of education and history of criminal conviction) to account statistically for the measures; we refer to that analysis as the adjusted model. Fourth, we fit a fixed-effects model at the maternal level that accounted for all factors that siblings share, including all genetic and environmental factors that make siblings similar, while controlling for offspring-specific covariates (we refer to the analysis as the fixed-effects model). The final model, therefore, compared siblings born at different gestational ages and statistically controlled for measured covariates that varied among siblings. We also ran several sensitivity analyses to test assumptions in sibling comparison studies, and we examined whether historical changes throughout the study period altered our conclusions concerning infant mortality.

Results
The sample is presented in the Supplement (eTable) by different GA categories and the number of cases for each outcome. The eTable in the Supplement illustrates how the covariates and outcome variables differed across the ordinal subgroups of GA.

Mortality
The initial analyses, which used the ordinal assessment of GA to provide estimates consistent with previous research, are presented in Figure 1 (all ordinal parameter estimates are available on request). There was a strong association between GA and risk of infant mortality in the population. For example, offspring born at 23 to 27 weeks of gestation had much higher odds of infant mortality (odds ratio [OR], 288.1; 95% CI, 271.7-305.5) compared with offspring born at term. Offspring born at 28 to 30 (OR, 72.8; 68.6-77.3), 31 to 33 (OR, 24.6; 23.3-26.1), and 34 to 36 weeks (OR, 6.9; 6.6-7.2) of gestation also had higher odds of mortality.

Figure 1 also summarizes the results from the continuous analyses of GA in the baseline model, where we present the results from either the linear or quadratic model of GA, depending on which model fit the best based on fit indices (available on request). A quadratic model fit significantly better than the linear model when predicting infant mortality. As can be seen with the parameter estimates in the Table, the estimated association from the baseline model was based on a quadratic model ($b_{linear} = -0.363; P < .001; b_{quadratic} = 0.004; P < .001$). The solid line closely follows the point estimates from the ordinal analyses when plotted in Figure 1, providing a similar interpretation to the results from the ordinal analysis. The association remained robust when controlling for covariates in
2.3-4.4), autism (HRGA: 23-27 weeks, 2.3; 2.0-2.8 for ADHD diagnosis; commensurate with those of a previous study,10 are available on request). When predicting psychotic or bipolar disorder, the magnitude of the association with earlier GA was slightly attenuated in the adjusted model, and the association was further attenuated in the fixed-effects model, suggesting that confounding factors account for some, but not all, of the increased risk with earlier GA. The adjusted and fixed-effects models for autism and ADHD found that the associations with GA were principally independent of the measured covariates and familial factors shared by siblings. A different pattern occurred when predicting suicide attempts (Figure 2). Extremely preterm GA was associated with increased risk of suicide attempts in the population (HRGA: 23-27 weeks, 1.7; 95% CI, 1.2-2.4). The association was slightly attenuated but still robust in the adjusted model. In contrast, the association between GA and suicide attempts was completely attenuated when comparing differentially exposed siblings, suggesting that shared familial confounding factors account for the statistical association in the population.

Early GA was associated with decreased risk for problematic substance use (HRGA: 23-27 weeks, 0.5; 95% CI, 0.4-0.7) and criminality (HRGA: 23-27 weeks, 0.7; 0.6-0.8) in the population (Figure 2). These decreased associations remained robust to the statistical controls and the comparison of siblings, suggesting that GA had a specific relationship with lower odds of substance use problems and criminality.

Academic Problems
The figures for academic problems are presented in the Supplement (eFigure 1). Early GA was also associated with multiple indicators of academic problems in the population,
including greater risk of failing grades (HR\textsubscript{GA: 23-27 weeks} \(2.0; 95\%\ CI, 1.7-2.3\)), odds of completing less than 10 years of education (HR\textsubscript{GA: 23-27 weeks} \(0.5; 0.4-0.6\)), and lower likelihood of parenthood (HR\textsubscript{GA: 23-27 weeks} \(0.7; 95\%\ CI, 0.6-0.9\)) and ever being married/in a registered partnership (HR\textsubscript{GA: 23-27 weeks} \(0.5; 0.4-0.6\)). When controlling for statistical covariates and shared familial confounds in the fixed-effects models did not reduce these associations. These findings are in contrast to those with receiving social welfare benefits. Earlier GA predicted social welfare benefits in the population (HR\textsubscript{GA: 23-27 weeks} \(1.3; 95\%\ CI, 1.2-1.5\)). The magnitude of the association was reduced in the adjusted model, and the association was largely attenuated when controlling for statistical covariates and shared familial confounds in the fixed-effects model (comparable results when predicting income are available on request).

**Social Adversity**

Early GA was strongly associated with social adversity, such as decreased likelihood of parenthood (HR\textsubscript{GA: 23-27 weeks} \(0.7; 95\%\ CI, 0.6-0.9\)) and ever being married/in a registered partnership (HR\textsubscript{GA: 23-27 weeks} \(0.2; 0.2-0.3\)) (Figure 3). Controlling for measured covariates in the adjusted models and shared familial confounding factors in the fixed-effects models did not reduce these associations. These findings are in contrast to those with receiving social welfare benefits. Earlier GA predicted social welfare benefits in the population (HR\textsubscript{GA: 23-27 weeks} \(1.3; 95\%\ CI, 1.2-1.5\)). The magnitude of the association was reduced in the adjusted model, and the association was largely attenuated when controlling for statistical covariates and shared familial confounds in the fixed-effects model (comparable results when predicting income are available on request).

**Sensitivity Analyses**

The sibling comparison design includes many limitations and assumptions that could influence the interpretation of the results.\textsuperscript{23,36-38,50} To address concerns about the general-
izability of findings from offspring with siblings to offspring without siblings, we ran 3 sets of sensitivity analyses. First, we compared the population estimates in families with multiple children with the estimates in families with 1 child (available on request). The population estimates were not lower in offspring with siblings (except for when predicting criminality), which indicates that the lower fixed-effects estimates (when they occurred) were not associated with lower population estimates in the subset of data that included offspring with siblings. Second, we ran the fixed-effects analyses with the ordinal distribution of GA to test whether mispecification of the shape of the models (eg, linear or quadratic) could account for the findings (available on request). The sensitivity analysis relaxed the assumption about the shape of the analytical models in families with multiple children, because only these families can provide information for the sibling comparison models. The results of the ordinal analyses gave interpretations commensurate with those in the main analyses, which...
We examined the possibility of exposure of one sibling influencing the outcome of another (ie, carryover effects) by conducting 2 sets of analyses. First, we fit bidirectional, case-crossover models (available on request), which explored whether different patterns of early GA within families (ie, either the first- or second-born offspring had lower GA) moderated the sibling comparison results. The analyses compared the sibling comparison estimates in families in which the first child had an earlier GA with families in which the second child had an earlier GA. The results were consistent only with a carryover effect for one outcome variable: low educational attainment. The bidirectional case-crossover model fitting, however, suggested the opposite pattern for 2 outcomes—infant mortality and psychotic and bipolar disorders—although the effect sizes were large in both types of sibling pairs. The results imply that the analyses suggesting carryover effects of early GA for the first-born sibling on the low educational attainment of the second-born sibling may be a chance finding. To further test for the possibility of carryover effects, we relied on the cousin comparison models (Supplement [eAppendix and eFigure 2]), in which carryover effects are less of a concern. The cousin comparisons again gave a pattern of results commensurate with those in the main analyses. The sensitivity analyses, therefore, do not support the hypothesis that carryover effects account for the attenuation of the associations in the sibling comparison models.

Sibling comparisons do not test for moderating factors. As such, we ran supplemental analyses (Supplement [eAppendix]) to examine whether year of birth decreased the association between GA and infant mortality, which has been reported elsewhere. The analyses support the overall conclusions regarding infant mortality. Finally, sibling comparisons are sensitive to measurement error. To begin to address this concern, we removed observations with extreme values for birth weight relative to GA because there may be misclassifications. The results of the baseline and fixed-effects models based on the subset of the data were commensurate with those presented in the main analyses (Supplement [eAppendix]).

**Discussion**

This large population-based cohort study replicates previous reports: early GA is associated with increased risk of early mortality and psychiatric, academic, and social prob-
lems. Early GA also was associated with decreased likelihood of criminality and substance use problems, consistent with some but not all previous research. The present study used a sibling comparison design and controlled for measured covariates to examine the degree to which confounding factors account for the associations. Several of the statistical associations (eg, with mortality during infancy and through young adulthood, autism, ADHD, substance use problems, criminality, parenthood, and ever partnered) were largely independent of shared familial confounding factors and the statistical covariates, consistent with a causal inference. The findings support theories associated with the mediating role in independent of shared familial confounding factors and the criminality, parenthood, and ever partnered) were largely young adulthood, autism, ADHD, substance use problems, associations (eg, with mortality during infancy and through brain development, on subsequent mortality and morbidity. In contrast, the associations between GA and other outcomes were either greatly (eg, with suicide and receiving social welfare benefits) attenuated. The latter results, therefore, suggest that confounding factors, such as environmental factors correlated with early GA, and not early GA in itself, account for these statistical associations. The findings for grades, educational attainment, suicide, and social welfare benefits contradict the results of previous studies and meta-analyses of the associations between GA and these outcomes as well as the general conclusions in reviews of the field, although no previous studies of these outcomes used a quasi-experimental approach.

The present study provides critical insight into the consequences associated with early GA because of 6 key advances. First and foremost, the study combined design features to rule out all confounding factors shared by siblings with statistical controls to rule out plausible alternative hypotheses for the observed associations. To our knowledge, this is one of the first studies of GA to use a quasi-experimental design, which is essential for drawing stronger causal inferences. The statistical associations between GA and many outcomes (eg, receiving social welfare benefits) were attenuated only in the fixed-effects models, which highlights the limitations of relying solely on statistical covariates to control for confounding factors. Second, the study explicitly tested several assumptions about sibling comparison studies, including the generalizability from offspring with siblings with statistical controls to rule out plausible alternative hypotheses for the observed associations. To our knowledge, this is one of the first studies of GA to use a quasi-experimental design, which is essential for drawing stronger causal inferences. The sensitivity analyses suggest that these alternative explanations do not account for the general conclusions, which further strengthen the inferences we were able to draw. However, additional quasi-experimental research, relying on methods with different assumptions and limitations, and research in other populations is necessary to strengthen causal inferences.

Third, this is the largest epidemiologic study to date of GA, providing a comprehensive view from an entire country. The sample size and measurement allowed us to more precisely estimate the risks for rare outcomes that were difficult to predict in previous research (eg, autism). Fourth, the analyses explored associations with the continuum of GA. Therefore, the study sheds light on extremely preterm and very preterm births in addition to moderate and late preterm births.

Fifth, the inclusion of multiple valid indices of morbidity in the present study allowed us the opportunity to find converging evidence—commensurate results were found when using different indices of key constructs. For example, we obtained the same results when predicting ADHD diagnosis and when predicting prescriptions for treating ADHD; we likewise found comparable results when predicting school grades and IQ. In addition, we obtained the same pattern when predicting both low and high income as well as social welfare benefits. As such, the results do not appear to be dependent on single observations or indices of important constructs.

Sixth, predicting multiple domains of functioning with valid indices of morbidity allowed us to explore the specificity of the predictions and underlying etiologic mechanisms associated with early GA. As such, the present study provides novel insight because the mechanisms responsible for the associations with early GA are outcome specific. In particular, researchers need to explore risk factors shared by siblings that account for the statistical association between early GA and suicide attempts, educational outcomes, and the need for social welfare benefits.

The present study also has several limitations. The findings need to be replicated to examine whether the results from a country with universal health care coverage and the quality of prenatal care in Sweden generalize to other countries. Quasi-experimental studies are not randomized studies and, therefore, cannot rule out all confounding factors. The sibling comparison design does not account for offspring-specific genetic factors that could influence GA. Twin and family quantitative genetic studies, including in this cohort, have indicated that fetal-specific genetic factors do not account for much variability in GA, although recent research suggests that such genetic factors may play a larger role than previous estimates. We controlled for offspring-specific covariates, but as is true of all human studies of GA, the present study cannot rule out the possibility that medical problems could cause preterm birth and the offspring outcomes. Nevertheless, the results suggest that risks specifically associated with early GA influence subsequent mortality and morbidity. The present study also may have misestimated the magnitude of some associations because the measurement of GA can misclassify some offspring. Sibling and cousin comparisons are sensitive to random measurement error and bias from confounders shared by siblings that are unrelated to the outcomes. In addition, fixed-effects models have lower statistical power than do population-based estimates, but our use of a continuous index of GA helped us to more precisely estimate the associations.

The present study, one of the first quasi-experimental studies of GA, stresses the importance of prevention efforts aimed at reducing preterm birth as well as wraparound services that target familial risks that occur with preterm birth. The findings should inform etiologic theory, risk assessment, and follow-up practices to prevent adverse outcomes associated with preterm birth.
Preterm Birth and Mortality and Morbidity

ARTICLE INFORMATION
Submitted for Publication: November 6, 2012; final revision received February 2, 2013; accepted February 18, 2013.
Published Online: September 25, 2013.

Author Contributions: Dr Rickert had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: D’Onofrio, Class, Larsson, Lichtenstein.
Acquisition of data: Lichtenstein.
Analysis and interpretation of data: All authors.
Drafting of the manuscript: D’Onofrio, Class, Rickert.
Critical revision of the manuscript for important intellectual content: D’Onofrio, Class, Larsson, Långström, Lichtenstein.
Statistical analysis: D’Onofrio, Class, Rickert, Lichtenstein.
Obtained funding: D’Onofrio, Class, Larsson, Lichtenstein.
Administrative, technical, and material support: Larsson, Långström, Lichtenstein.
Study supervision: D’Onofrio, Larsson.

Conflict of Interest Disclosures: None reported.
Funding/Support: The study was supported by grant HD061877 from the National Institute of Child Health and Human Development, grant MH059401 from the National Institute of Mental Health, the Swedish Research Council (Medicine), and the Swedish Prison and Probation Services.

Role of the Sponsor: The study sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES
34. Freese J. Genetics and the social science explanation of individual outcomes. AJSP. 2006;144(suppl):S33-53.


