

HHS Public Access

Birth Defects Res. Author manuscript; available in PMC 2024 November 14.

Published in final edited form as:

Author manuscript

Birth Defects Res. 2024 January ; 116(1): e2294. doi:10.1002/bdr2.2294.

Factors associated with infant sex and preterm birth status for selected birth defects from the National Birth Defects Prevention Study, 1997–2011

Eva M. Williford1, **Wei Yang**2, **Meredith M. Howley**1, **Chen Ma**2, **Ronnie T. Collins**2, **Kari A. Weber**3, **Dominique Heinke**4, **Julie M. Petersen**4, **A. J. Agopian**5, **Natalie P. Archer**6, **Andrew F. Olshan**7, **Lindsay A. Williams**8, **Marilyn L. Browne**1,9, **Gary M. Shaw**2, **National Birth Defects Prevention Study**

¹Birth Defects Registry, New York State Department of Health, Albany, New York, USA

²Division of Neonatal and Developmental Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, California, USA

³Arksansas Center for Birth Defects Research and Prevention and Department of Epidemiology, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

⁴Center for Birth Defects Research and Prevention, Massachusetts Department of Public Health, Boston, Massachusetts, USA

⁵Department of Epidemiology, Human Genetics and Environmental Sciences, UTHealth School of Public Health, Houston, Texas, USA

⁶Environmental Epidemiology and Disease Registries Section, Texas Department of State Health Services, Austin, Texas, USA

⁷Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

⁸Division of Epidemiology and Clinical Research, Department of Pediatrics, University of Minnesota, Minneapolis, Minnesota, USA

⁹Department of Epidemiology and Biostatistics, School of Public Health, University at Albany, Rensselaer, New York, USA

Abstract

Background: Birth defects and preterm birth co-occur, with some overlapping risk factors. Many birth defects and preterm births tend to have a male preponderance. We explored potential

Correspondence: Eva M. Williford, Birth Defects Registry, New York State Department of Health, Albany, New York, USA. eva.williford@health.ny.gov. Marilyn L. Browne and Gary M. Shaw are co-senior authors.

CONFLICT OF INTEREST STATEMENT

The authors report no conflicts of interest.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

risk factors impacting sex and preterm (<37 weeks of gestation) birth differences among infants with selected birth defects delivered from 1997 to 2011 using data from the National Birth Defects Prevention Study (NBDPS).

Methods: The NBDPS was a large multisite, population-based case–control study. Using random forests, we identified important predictors of male preterm, female preterm, and male term, each compared with female term births for each birth defect. Using logistic regression, we estimated odds ratios for associations between important predictors and sex-preterm birth status by birth defect.

Results: We examined 11,379 infants with nine specific birth defects. The top 10 most important predictors of sex-preterm birth status from the random forests varied greatly across the birth defects and sex-preterm comparisons within a given defect group, with several being novel factors. However, one consistency was that short interpregnancy interval was associated with sex-preterm birth status for many of the studied birth defects. Although obesity has been identified as a risk factor for preterm birth and birth defects in other research, it was not associated with sex-preterm birth status for any of the examined defects.

Conclusions: We confirmed expected associations for sex-preterm birth status differences and found new potential risk factors for further exploration among the studied birth defects.

Keywords

birth defects; preterm birth; random forests; sex

1 | INTRODUCTION

In the United States, birth defects occur in approximately 3% of live births (Prevention, 2008). Preterm birth affects 10% of all births but 21% of infants with birth defects (Purisch & Gyamfi-Bannerman, 2017; Rasmussen et al., 2001; Reefhuis et al., 2015). Understanding birth defects and preterm births can be complex, as risk factors for many birth defects overlap with risk factors for preterm birth, including short interpregnancy interval (time from end of one pregnancy to the start of the next), pre-pregnancy obesity, no folic acidcontaining supplement intake, pre-gestational diabetes, and maternal smoking (Dolan et al., 2009; Shaw, 2015).

Additionally, a preponderance of male sex has been observed for both preterm birth and some birth defects (Michalski et al., 2015; Shaw et al., 2003, 2021). This male excess is particularly striking among deliveries before 32 weeks of gestation (Shaw et al., 2021). Historically, a female preponderance has been observed among infants with neural tube defects; however, in recent years this difference has narrowed (Poletta et al., 2018; Shaw et al., 2020). To our knowledge, a comprehensive inquiry has not been made of potential maternal perinatal risk factors for the joint outcome of infant sex (male vs. female) and preterm birth status for a given birth defect. Identifying factors associated with sex and preterm birth differences for a given birth defect phenotype may stimulate new hypotheses regarding the etiologies of birth defects as well as preterm birth. More broadly, such consideration may enhance the understanding of sexual dimorphism and human development. In this study, we conducted exploratory analyses for potential risk

factors impacting jointly defined sex and preterm birth groupings among selected birth defects by investigating data from the National Birth Defects Prevention Study (NBDPS).

2 | METHODS

The NBDPS was a large multisite, population-based case–control study that included data from pregnancies with estimated delivery dates (EDD) from October 1997 through December 2011 ascertained through birth defects surveillance programs from selected geographic regions in 10 states (Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah) (Reefhuis et al., 2015). Institutional review board approval was obtained for each study site, and participants provided informed consent. Clinical geneticists reviewed medical record information on each case to determine eligibility and to classify cases as having isolated (one major birth defect), multiple (two or more major defects in more than one organ system), or complex birth defects (Reefhuis et al., 2015). In addition, birth defects attributed to a known chromosomal abnormality or single-gene condition were excluded. Infants with congenital heart defects (CHDs) were classified based on cardiac phenotype, complexity, and presence of non-cardiac birth defects (Botto et al., 2007).

Our analysis included infants with at least one of the following birth defects: spina bifida, D-transposition of the great arteries, tetralogy of Fallot, cleft palate without cleft lip, cleft lip with or without cleft palate, longitudinal/intercalary limb deficiency, transverse limb deficiency, craniosynostosis, or gastroschisis. We chose these nine birth defects ($n = 12,276$) for their well-defined phenotypic classification and larger sample sizes. Cases included live births, stillbirths (spontaneous loss at 20 weeks of gestation or later), and induced terminations. We included cases classified as having isolated, multiple, or complex defects in our analyses. We excluded cases with ambiguous or missing sex, or unknown gestational age at delivery. Our analysis excludes data from controls. For each specific birth defect, we categorized the infants into four distinct groups by sex (male or female) and gestational age at delivery dichotomized as term or preterm $(37 \text{ or } 37 \text{ weeks of gestation})$. We obtained gestational age from the infant's birth or medical record. The outcome variable for this study was a composite of infant sex and gestational age at delivery with four levels: male preterm, female preterm, male term, and female term. We compared each sex-preterm birth combination to a common sex-preterm birth referent category (female term births).

As pre-gestational diabetes (i.e., type I or II) has been reported to be associated with these selected birth defects, infants whose mothers had these conditions were excluded from the analyses ($n = 254, 2.1\%$) (Correa et al., 2008). We further excluded multiple births from the analyses ($n = 560, 4.7\%$ cases) since the etiologies of preterm birth may differ between singleton and multiple births (Tingleff et al., 2023).

Trained interviewers conducted computer-assisted telephone interviews in English or Spanish with participating women between 6 weeks and 24 months after their EDD. Women answered questions about their demographics, pregnancy history, health conditions, and other exposures before or during pregnancy. We considered four potential risk factors selected a priori for differences in sex and gestational age at delivery: (1) folic acid-

containing supplement intake (yes or no) in early pregnancy, (2) maternal smoking (yes or no) in early pregnancy, (3) pre-pregnancy obesity (body mass index (BMI) 30 kg/m^2 or $\langle 30 \text{ kg/m}^2 \rangle$, and (4) interpregnancy interval (no previous pregnancies, 12 months, or >12 months) (Dolan et al., 2009; Shaw et al., 2021). We defined interpregnancy interval as the difference in months between the date of conception of the index pregnancy and the end date of the last previous pregnancy before the index pregnancy. Early pregnancy is the critical period in embryonic development associated with most structural defects and was defined as the month before conception through the second month of pregnancy. The month before conception was included, as it is difficult to determine the exact date of conception.

For each birth defect, we used a multinomial, multivariable logistic regression model to analyze the association between the four-level outcome and the four potential risk factors selected a priori described above. We calculated adjusted odds ratios (aORs) and associated 95% confidence intervals (CIs) for each model using female term as the referent outcome and adjusting for three other a priori potential risk factors. We used likelihood ratio tests to investigate all possible two-way statistical interactions between the four potential risk factors.

To further understand predictors of joint sex and preterm birth status among infants with birth defects, we conducted an exploratory (hypothesis-generating) analysis with random forests. We used this data-mining procedure to identify potential risk factors by birth defect. Random forests are a supervised machine-learning method that models a number of decision trees to classify observations as, for example male preterm versus female term, based on a set of predictors (Strobl et al., 2009). We ran three separate models (male preterm vs. female term, female preterm vs. female term, and male term vs. female term) for each defect. For our analysis, we modeled 5000 conditional inference trees (sufficiently large number of trees to achieve stable results) with 15 variables (square-root of the number of variables) randomly sampled to determine each split in a given tree with a minimum sum of weights in a node of five (results were stable across different random seeds) (Strobl et al., 2009). Since the variables differed in scale of measurement and to remove bias towards correlated variables, we utilized conditional inference trees, as they produce unbiased trees and use an adequate resampling method (Hothorn et al., 2006; Strobl et al., 2009). The variables were ranked based on the metric mean decrease accuracy (MDA) that was calculated for each variable as a measure of variable importance (Strobl et al., 2009).

We included a total of 241 variables in the random forests including dietary, demographic, and behavioral characteristics (Appendix A). For interview questions that asked about specific timing before and during pregnancy, separate variables for each month during early pregnancy were included (the month before conception, the first month of pregnancy, and the second month of pregnancy). Overall, missingness was low, ranging from 0% to 10% across the 241 variables, with only five variables having missing values for more than 5% of the women. We excluded women with missing responses for more than 10% of the variables considered ($n = 752, 6.6\%$). For variables with 10% missing values, missingness was imputed with the most frequent response for categorical variables and the median for continuous variables (Schafer, 1999; Weber et al., 2018).

To quantify the associations between the top 10 most important predictors from the random forests and sex-preterm birth status, conventional aORs and 95% CIs were estimated from logistic regression models with Firth penalization (Firth, 1993).

Random forest analyses were performed using the Party Package in R software (V4.1.3). All other analyses were performed in SAS version 9.4 (SAS Institute, Cary, NC).

3 | RESULTS

We analyzed nine birth defects among 11,379 infants in total; the number of infants with each birth defect ranged from 465 to 2930 (Table 1). Infants with D-transposition of the great arteries had the lowest proportion of preterm birth (9.2%), whereas infants with gastroschisis had the highest (62.0%). Among control infants, 4.1% ($n = 468$) were male preterm, 3.8% ($n = 428$) were female preterm, 46.9% ($n = 5311$) were male term, and 45.1% $(n = 5108)$ were female term births.

3.1 | A priori selected potential risk factor results

The a priori selected potential risk factors were tabulated by sex and preterm birth status for each birth defect (Table 2). Among infants with spina bifida with mothers not taking a folic acid-containing supplement during early pregnancy, there was a higher proportion of males compared with females for both preterm (17.2% vs. 10.0%) and term (40.5% vs. 32.3%) births. Among infants with spina bifida, there was a higher proportion of male preterm births among mothers who had an interpregnancy intervals of 12 months (17.1%) compared with mothers in the other interpregnancy interval categories (12.4% for no previous pregnancies and 10.8% for >12 months). A similar pattern was observed among gastroschisis (36.7% male preterm births among mothers with interpregnancy intervals of

12 months, compared with 30% among those with no previous pregnancies and 30% among those with interpregnancy intervals >12 months). Among infants with longitudinal/ intercalary limb deficiency, regardless of sex, there was a higher proportion of preterm births among obese mothers (14.8% for males and 16.1% for females) compared with non-obese mothers (11.6% for males and 9.7% for females).

Multinomial, multivariable logistic regression results are presented in Table 3. For infants with spina bifida, compared with female term births, mothers of male preterm births were more likely to have not used folic acid-containing supplements (aOR [95% CI]: 1.95 [1.25– 3.04]) and less likely to have pre-pregnancy obesity (aOR [95% CI]: 0.59 [0.36–0.97]). Additionally, mothers of male term births were less likely to have interpregnancy intervals of 12 months compared with mothers of female term births (aOR [95% CI]: 0.67 [0.46– 0.97]). For infants with D-transposition of the great arteries, compared with female term births, mothers of female preterm births were more likely to be smokers (aOR [95% CI]: 4.22 $[1.50-11.92]$) or have an interpregnancy interval of $[12 \text{ months (aOR } 95\% \text{ CI}]$: 5.14 [1.54–17.21]). For infants with cleft palate without cleft lip, compared with female term births, mothers of male preterm and male term births were more likely to have an interpregnancy interval 12 months (aOR [95% CI]: 2.01 [1.16–3.48] and 1.66 [1.22–2.27], respectively). For infants with craniosynostosis, compared with female term births, mothers of male preterm births were more likely to have an interpregnancy interval of 12 months

(aOR [95% CI]: 1.76 [1.05–2.94]) and mothers of male term births were more likely to take folic acid-containing supplements (aOR [95% CI]: 0.70 [0.51–0.95]).

3.2 | Random forests results

Our analyses employing random forests sought to identify variables associated with the outcome in each of the three models (male preterm vs. female term, female preterm vs. female term, and male term vs. female term) by birth defect. Table 4 provides the top 10 most important predictors from the random forests, with female term as the referent outcome for each model by birth defect. Tables S1–S9 present the aORs and corresponding 95% CIs for each logistic regression model with the top 10 most important predictors from the random forests for the studied defects.

Household smoke exposure was among the top 10 most important predictors from the random forests for two of the models (in the male preterm vs. female term model and the male term vs. female term model) for infants with spina bifida (Table 4). Household smoke exposure in early pregnancy was identified as a risk factor for male preterm births (aOR [95% CI]: 1.34 [0.80–2.21]) but had a reduced aOR for male term births compared with female term births (aOR [95% CI]: 0.69 [0.45–1.06]) (Table S1). Consistent with the multinomial, multivariable logistic regression model with the a priori potential risk factors (Table 3), mothers of male preterm births were more likely to not take folic acid-containing supplements during the second month of pregnancy (aOR [95% CI]: 1.81 [1.14–2.85]) compared with mothers of female term births (Table S1).

There were no commonalities in the top 10 most important predictors from the random forests across the three models for infants with D-transposition of the great arteries or tetralogy of Fallot (Table 4). Mothers of female preterm births were more likely to smoke during the first 2 months of pregnancy compared with mothers of female term births (aOR [95% CI]: 3.17 [1.04–9.31]) (Table S2), consistent with the multinomial, multivariable logistic regression model for that association among infants with D-transposition of the great arteries (Table 3). Among infants with tetralogy of Fallot, mothers of male preterm births were more likely to have no previous pregnancies versus $a > 12$ month interpregnancy interval (aOR [95% CI]: 1.73 [1.04–2.87]) compared with mothers of female term births (Table S3).

There were no commonalities in the top 10 most important predictors from the random forests across the three models for infants with cleft palate without cleft lip (Table 4). Mothers of male term births were more likely to have an interpregnancy interval of 12 months (aOR [95% CI]: 1.55 [1.13–2.12]) compared with mothers of female term births (Table S4), consistent with the multinomial, multivariable logistic regression model with the a priori potential risk factors (Table 3). Two commonalities, though not statistically significant, were observed in the top 10 most important predictors from the random forests among infants with cleft lip with or without cleft palate (Table 4): eating cantaloupe (in the male preterm vs. female term model and the female preterm vs. female term model) and paternal race/ethnicity (in the male preterm vs. female term model and the male term vs. female term model).

Among infants with longitudinal/intercalary limb deficiency, vitamin E as alpha-tocopherol intake was identified among the top 10 most important predictors from the random forests in more than one model (in the male preterm vs. female term model and the female preterm vs. female term model) (Table 4). Mothers of male preterm births were more likely to have had fertility treatment compared with mothers of female term births (aOR [95% CI]: 3.97 [1.15–15.53]) among infants with longitudinal/intercalary limb deficiency (Table S6). For infants with transverse limb deficiency, total carbohydrate, caffeine from soda, vitamin C intake, maternal race/ethnicity, paternal race/ethnicity, and vitamin E intake were identified in the top 10 most important predictors from the random forests in more than one model (Table 3). Mothers of male preterm births were more likely to consume caffeine (mg per day) from soda (aOR [95% CI] for a 10-unit change: 1.05 [1.02–1.10]) compared with mothers of female term births (Table S7). The association of caffeine from soda with female preterm versus female term was not statistically significant.

For infants with craniosynostosis, frequency of baths at home and study site were identified in more than one model (in the male preterm vs. female term model and the female preterm vs. female term model) among the top 10 most important predictors from the random forests (Table 4). Among infants with craniosynostosis, compared with female term births, mothers of male preterm birth were less likely to not use folic acid-containing supplements during the first month of pregnancy (aOR [95% CI]: 0.62 [0.38–0.98]) (Table S8). Folic acid-containing supplement use in early pregnancy was found to be statistically significant for male term, but not for male preterm births compared with female term births in the a priori selected potential risk factors analysis (Table 3). Among infants with gastroschisis, caffeine consumption from coffee (mg per day) was identified as a top 10 most important predictor from the random forests by all three models.

4 | DISCUSSION

The objective of this study was to perform exploratory and hypothesis generating analyses of sex and preterm birth differences to identify areas for future research. We investigated known and agnostically identified factors that might contribute to the differences in sex and gestational age at delivery for specific birth defects. Short interpregnancy interval (12) months) was associated with some sex-preterm birth status differences among many of the birth defects studied. While obesity has been associated with select birth defects and preterm birth in prior research (Challis et al., 2013; Liu et al., 2019; Stothard et al., 2009), it was not associated with any sex-preterm birth status differences.

The top 10 most important predictors identified agnostically from the random forests varied greatly across sex-preterm comparisons within each birth defect and across birth defects within a given sex-preterm birth comparison. Overall, the results from the logistic regression models suggested non-null associations with the top 10 most important predictors identified from the random forests. Findings were most consistent between the random forests and the multinomial, multivariable logistic regression models with a priori potential risk factors for sex-preterm birth status among infants with spina bifida, D-transposition of the great arteries, or tetralogy of Fallot. However, it is important to note that these modeling strategies

have important differences (e.g., logistic regression is parametric whereas random forests are nonparametric), so we would not necessarily expect them to always agree.

In this exploratory analysis, we utilized random forests, a non-hypothesis-driven datamining algorithm. This approach allowed us to investigate a large number of potential risk factors simultaneously and rank the variables for each model based on the MDA. Random forests confirmed some of the four a priori identified associations and allowed us to observe new potential risk factors for further exploration of sex-preterm birth status differences in future birth defects studies.

Caffeine consumption was identified as an important predictor for sex-preterm birth status multiple models among infants with transverse limb deficiency (caffeine from soda) or gastroschisis (caffeine from coffee). Previous NBDPS analyses have observed some small, elevated effect estimates between pre-pregnancy total caffeine consumption and birth defects, including transverse limb deficiency (Williford et al., 2023). We observed that the association of caffeine consumption from soda differed between male preterm and female preterm compared with female term births among infants with transverse limb deficiency. Among infants with gastroschisis the associations between caffeine consumption from coffee and sex-preterm birth status were not statistically significant. Inconsistent results have been observed in the literature for the association of caffeine consumption during pregnancy and risk of preterm delivery (Maslova et al., 2010).

Study site was identified among the top 10 most important predictors in two out of the three models among infants with craniosynostosis (male preterm vs. female term and female preterm vs. female term). Arkansas was more likely to have male preterm and female preterm births compared with many of the other study sites adjusted for all other predictors in the models. Among infants with craniosynostosis, 18.7% were preterm births in Arkansas (43% with gestational age at delivery of 36 weeks among preterm births). This is a larger percentage of craniosynostosis cases that were preterm births than other sites, which ranged from 4.1% (New York) to 14.0% (Georgia). The percentage of infants with craniosynostosis classified as isolated was consistent across study sites, ranging from 88% (California) to 93% (Arkansas). Case ascertainment for some pregnancy outcomes and prenatal diagnosis procedures differed over time for some sites (Reefhuis et al., 2015). This hypothesis generating analysis has identified study site as an area of future work in sex and preterm differences among infants with craniosynostosis.

Our findings are consistent with some of the prior research of infant sex and identified risk factors, although existing work has not examined the combined outcome of sex-preterm birth status among infants with birth defects. Others have observed that maternal cigarette smoking may negatively impact growth in male fetuses more than female fetuses (Shaw et al., 2003). In our analysis, we found maternal cigarette smoking in early pregnancy to be associated with female preterm births compared with female term births among infants with D-transposition of the great arteries. We did not find maternal cigarette smoking to be associated with male preterm or term births for any of the studied birth defects. A previous analysis explored folic acid use and infant sex among neural tube defects and reported more females in the two studies with pregnancies before mandatory folate fortification in the

United States (Shaw et al., 2020). However, for infants with spina bifida, we observed that mothers of male preterm births were more likely to not be taking folic-acid supplementation during early pregnancy compared with mothers of female term births. Another study found older paternal age to be associated with female births among infants with cleft lip with or without cleft palate, and gravidity to be associated with female births among infants with spina bifida (Rittler et al., 2004). We did not identify either of those variables among the top 10 most important predictors from the random forests for the models with either cleft lip or spina bifida. Discrepancies between our findings and those of previous research may be explained by a wide range of factors, including differences in data sources, methods, outcome definitions, and exposure assessments.

This study has many strengths including the large multi-site population-based design, the clinical classification of birth defects, and a detailed standardized questionnaire. The use of random forests allowed for the exploration of a large number of variables, capitalizing on the breadth of potential risk factors captured in these data, without concerns regarding correlations between the variables. Limitations of this study include that exposure data were collected after delivery, which could impact recall, and the potential for selection bias due to participation refusals and non-response. Our observations may be biased if a particular birth defect and sex combination was more likely to result in a pregnancy loss that was not ascertained (spontaneous pregnancy loss before 20 weeks gestation). Such bias could be further amplified if studied factors also increased the likelihood of the particular birth defect and sex combination to result in a pregnancy loss. In addition, our observations could be biased for some of the birth defects (e.g., craniosynostosis, D-transposition of the great arteries, or tetralogy of Fallot) that may not have been diagnosed in pregnancy terminations or losses which would bias the gestational age to term births (Heinke et al., 2020; Liberman et al., 2023; McPherson et al., 2017). For infants with gastroschisis, preterm delivery may be initiated by the provider due to the presence of the defect (Friedman et al., 2016); however, in our data we are unable to distinguish if the preterm delivery was spontaneous or provider-initiated. The literature suggests that spontaneous preterm birth occurs frequently and the optimal timing of delivery is not conclusive for infants with gastroschisis (Baer et al., 2019; Friedman et al., 2016; Goldstein et al., 2022). Thus, results may be biased and should interpreted with caution for defects such as gastroschisis, where there may be a preference for an early delivery at some facilities. We defined preterm birth using the standard definition of less than 37 weeks gestation at delivery. In our analysis of 11,379 infants with selected birth defects, 4.6% ($n = 525$) were delivered at 35 weeks gestation and 6.7% ($n = 759$) were delivered at 36 weeks gestation. There are two potential limitations of our definition of preterm birth, which could be explored in future work. We may have observed different results using other definitions of preterm birth (e.g., less than 32 weeks gestation) or by focusing on spontaneous preterm births (the reason for preterm birth was not collected in NBDPS). In addition, some of the outcome categories within each birth defect were small, resulting in imprecise estimates. The impact of predictor classification errors on the performance of random forests is unclear without formal bias analysis (Jiang et al., 2021). Lastly, there are alternative approaches for calculating variable importance for random forests (e.g., Gini impurity importance) (Strobl et al., 2007). We did not evaluate whether modifying the variable importance measure or tuning parameters might affect our

results from random forests. A limitation of evaluating so many risk factors at once is the concern of multiple testing (365 estimates across the three models for the nine birth defects); some observed associations may be due to chance. We did not perform multiple comparison adjustment methods and presented all estimates and confidence intervals calculated as recommended by Rothman (1990) and Greenland (2008). Results should therefore be interpreted cautiously.

5 | CONCLUSION

Our findings suggest that there are differences in infant sex and preterm birth status and their predictors among the studied birth defects. Our analysis confirmed some known risk factors for preterm birth and birth defects (short interpregnancy interval and no folic acid-containing supplement use). Further exploration of the newly identified factors may help advance understanding of sex differences among birth defects and preterm birth.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This project was supported through Centers for Disease Control and Prevention (CDC) cooperative agreements under PA #96043, PA #02081, FOA #DD09-001, FOA #DD13-003, and NOFO #DD18-001 to the Centers for Birth Defects Research and Prevention participating in the National Birth Defects Prevention Study (NBDPS) and/or the Birth Defects Study To Evaluate Pregnancy exposureS (BD-STEPS) and the New York Center for Birth Defects Research and Prevention U01 DD001227. We thank the California Department of Public Health Maternal Child and Adolescent Health Division for providing data. We thank the participating families and staff from the NBDPS sites. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the California Department of Public Health.

Funding information

Centers for Disease Control and Prevention, Grant/Award Numbers: PA #96043, PA #02081, FOA #DD09-001, FOA #DD13-003, NOFO #DD18-001; New York Center for Birth Defects Research and Prevention, Grant/Award Number: U01 DD001227

DATA AVAILABILITY STATEMENT

The study questionnaires and process for accessing the data used in this study is described at [https://www.cdc.gov/ncbddd/birthdefects/nbdps-public-access-procedures.html.](https://www.cdc.gov/ncbddd/birthdefects/nbdps-public-access-procedures.html)

APPENDIX A: Variables from the National Birth Defects Prevention Study included in random forests.

Variable

Variable

Variable

Abbreviations: B1, month before conception; P1, first month of pregnancy; P2, second month of pregnancy; P3, third month of pregnancy.

REFERENCES

- Baer RJ, Chambers CD, Ryckman KK, Oltman SP, Rand L, & Jelliffe-Pawlowski LL (2019). High risk of spontaneous preterm birth among infants with gastroschisis. American Journal of Medical Genetics, Part A, 179(1), 37–42. 10.1002/ajmg.a.60675 [PubMed: 30549407]
- Botto LD, Lin AE, Riehle-Colarusso T, Malik S, & Correa A (2007). Seeking causes: Classifying and evaluating congenital heart defects in etiologic studies. Birth Defects Research. Part A, Clinical and Molecular Teratology, 79(10), 714–727. 10.1002/bdra.20403 [PubMed: 17729292]
- Challis J, Newnham J, Petraglia F, Yeganegi M, & Bocking A (2013). Fetal sex and preterm birth. Placenta, 34(2), 95–99. 10.1016/j.placenta.2012.11.007 [PubMed: 23261268]
- Correa A, Gilboa SM, Besser LM, Botto LD, Moore CA, Hobbs CA, Cleves MA, Riehle-Colarusso TJ, Waller DK, & Reece EA (2008). Diabetes mellitus and birth defects. American Journal of Obstetrics and Gynecology, 199(3), 237.e231–237.e239. 10.1016/j.ajog.2008.06.028
- Dolan SM, Callaghan WM, & Rasmussen SA (2009). Birth defects and preterm birth: Overlapping outcomes with a shared strategy for research and prevention. Birth Defects Research. Part A, Clinical and Molecular Teratology, 85(11), 874–878. 10.1002/bdra.20634 [PubMed: 19824057]
- Firth D (1993). Bias reduction of maximum likelihood estimates. Biometrika, 80(1), 27–38. 10.2307/2336755
- Friedman AM, Ananth CV, Siddiq Z, D'Alton ME, & Wright JD (2016). Gastroschisis: Epidemiology and mode of delivery, 2005–2013. American Journal of Obstetrics and Gynecology, 215(3), 348.e341–348.e349. 10.1016/j.ajog.2016.03.039
- Goldstein MJ, Bailer JM, & Gonzalez-Brown VM (2022). Preterm vs term delivery in antenatally diagnosed gastroschisis: A systematic review and meta-analysis. American Journal of Obstetrics & Gynecology MFM, 4(4), 100651. 10.1016/j.ajogmf.2022.100651 [PubMed: 35462060]
- Greenland S (2008). Multiple comparisons and association selection in general epidemiology. International Journal of Epidemiology, 37(3), 430–434. 10.1093/ije/dyn064 [PubMed: 18453632]
- Heinke D, Nestoridi E, Hernandez-Diaz S, Williams PL, Rich-Edwards JW, Lin AE, Van Bennekom CM, Mitchell AA, Nembhard WN, Fretts RC, Roberts DJ, Duke CW, Carmichael SL, Yazdy MM, & National Birth Defects Prevention Study. (2020). Risk of stillbirth for fetuses with specific birth defects. Obstetrics and Gynecology, 135(1), 133–140. 10.1097/aog.0000000000003614 [PubMed: 31809437]
- Hothorn T, Hornik K, & Zeileis A (2006). Unbiased recursive partitioning: A conditional inference framework. Journal of Computational and Graphical Statistics, 15(3), 651–674. 10.1198/106186006X133933
- Jiang T, Gradus JL, Lash TL, & Fox MP (2021). Addressing measurement error in random forests using quantitative bias analysis. American Journal of Epidemiology, 190(9), 1830–1840. 10.1093/aje/kwab010 [PubMed: 33517416]
- Liberman RF, Heinke D, Lin AE, Nestoridi E, Jalali M, Markenson GR, Sekhavat S, & Yazdy MM (2023). Trends in delayed diagnosis of critical congenital heart defects in an era of enhanced screening, 2004–2018. The Journal of Pediatrics, 257, 113366. 10.1016/j.jpeds.2023.02.012 [PubMed: 36858148]
- Liu B, Xu G, Sun Y, Du Y, Gao R, Snetselaar LG, Santillan MK, & Bao W (2019). Association between maternal pre-pregnancy obesity and preterm birth according to maternal age and race or ethnicity: A population-based study. The Lancet Diabetes and Endocrinology, 7(9), 707–714. 10.1016/s2213-8587(19)30193-7 [PubMed: 31395506]
- Maslova E, Bhattacharya S, Lin S-W, & Michels KB (2010). Caffeine consumption during pregnancy and risk of preterm birth: A meta-analysis. The American Journal of Clinical Nutrition, 92(5), 1120–1132. 10.3945/ajcn.2010.29789 [PubMed: 20844077]
- McPherson E, Nestoridi E, Heinke D, Roberts DJ, Fretts R, Yazdy MM, & Lin AE (2017). Alternatives to autopsy for fetal and early neonatal (perinatal) deaths: Insights from the Wisconsin stillbirth service program. Birth Defects Research, 109(18), 1430–1441. 10.1002/bdr2.1112 [PubMed: 28898573]
- Michalski AM, Richardson SD, Browne ML, Carmichael SL, Canfield MA, VanZutphen AR, Anderka MT, Marshall EG, & Druschel CM (2015). Sex ratios among infants with birth defects, National

Birth Defects Prevention Study, 1997–2009. American Journal of Medical Genetics. Part A, 167a(5), 1071–1081. 10.1002/ajmg.a.36865 [PubMed: 25711982]

- Poletta FA, Rittler M, Saleme C, Campaña H, Gili JA, Pawluk MS, Gimenez LG, Cosentino VR, Castilla EE, & López-Camelo JS (2018). Neural tube defects: Sex ratio changes after fortification with folic acid. PLoS One, 13(3), e0193127. 10.1371/journal.pone.0193127 [PubMed: 29538416]
- Prevention C f. D. C. a. (2008). Update on overall prevalence of major birth defects—Atlanta, Georgia, 1978–2005. MMWR Weekly Report, 57(1), 1–5. [http://www.cdc.gov/mmwr/preview/mmwrhtml/](http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5701a2.htm) [mm5701a2.htm](http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5701a2.htm)
- Purisch SE, & Gyamfi-Bannerman C (2017). Epidemiology of preterm birth. Seminars in Perinatology, 41(7), 387–391. 10.1053/j.semperi.2017.07.009 [PubMed: 28865982]
- Rasmussen SA, Moore CA, Paulozzi LJ, & Rhodenhiser EP (2001). Risk for birth defects among premature infants: A population-based study. The Journal of Pediatrics, 138(5), 668–673. 10.1067/ mpd.2001.112249 [PubMed: 11343041]
- Reefhuis J, Gilboa SM, Anderka M, Browne ML, Feldkamp ML, Hobbs CA, Jenkins MM, Langlois PH, Newsome KB, Olshan AF, Romitti PA, Shapira SK, Shaw GM, Tinker SC, Honein MA, & Study t. N. B. D. P. (2015). The national birth defects prevention study: A review of the methods. Birth Defects Research Part A: Clinical and Molecular Teratology, 103(8), 656–669. 10.1002/bdra.23384 [PubMed: 26033852]
- Rittler M, López-Camelo J, & Castilla EE (2004). Sex ratio and associated risk factors for 50 congenital anomaly types: Clues for causal heterogeneity. Birth Defects Research. Part A, Clinical and Molecular Teratology, 70(1), 13–19. 10.1002/bdra.10131 [PubMed: 14745890]
- Rothman KJ (1990). No adjustments are needed for multiple comparisons. Epidemiology, 1(1), 43–46. [PubMed: 2081237]
- Schafer JL (1999). Multiple imputation: a primer. Statistical Methods in Medical Research, 8(1), 3–15. 10.1177/096228029900800102 [PubMed: 10347857]
- Shaw GM (2015). Epidemiology of structural birth defects and preterm birth. In Stevenson DK, Cohen RS, & Sunshine P (Eds.), Neonatology: Clinical practice and procedures. McGraw-Hill Education. accesspediatrics.mhmedical.com/content.aspx?aid=1109791069
- Shaw GM, Carmichael SL, Kaidarova Z, & Harris JA (2003). Differential risks to males and females for congenital malformations among 2.5 million California births, 1989–1997. Birth Defects Research. Part A, Clinical and Molecular Teratology, 67(12), 953–958. 10.1002/bdra.10129 [PubMed: 14745913]
- Shaw GM, Mayo JA, Eisenberg ML, Catalano R, & Stevenson DK (2021). Male-to-female ratios, race/ ethnicity, and spontaneous preterm birth among 11 million California infants. American Journal of Perinatology, 38(7), 683–689. 10.1055/s-0039-3400449 [PubMed: 31756757]
- Shaw GM, Yang W, & Finnell RH (2020). Male-to-female ratios among NTDs and women's periconceptional intake of folic acid. Birth Defects Research, 112(16), 1187–1193. 10.1002/ bdr2.1708 [PubMed: 32415919]
- Stothard KJ, Tennant PW, Bell R, & Rankin J (2009). Maternal overweight and obesity and the risk of congenital anomalies: A systematic review and meta-analysis. JAMA, 301(6), 636–650. 10.1001/ jama.2009.113 [PubMed: 19211471]
- Strobl C, Boulesteix AL, Zeileis A, & Hothorn T (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8, 25. 10.1186/1471-2105-8-25 [PubMed: 17254353]
- Strobl C, Malley J, & Tutz G (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychological Methods, 14(4), 323–348. 10.1037/a0016973 [PubMed: 19968396]
- Tingleff T, Räisänen S, Vikanes Å, Sandvik L, Sugulle M, Murzakanova G, & Laine K (2023). Different pathways for preterm birth between singleton and twin pregnancies: A population-based registry study of 481 176 nulliparous women. BJOG, 130(4), 387–395. 10.1111/1471-0528.17344 [PubMed: 36372962]
- Weber KA, Yang W, Carmichael SL, Lupo PJ, Dukhovny S, Yazdy MM, Lin AE, Van Bennekom CM, Mitchell AA, Shaw GM, & National Birth Defects Prevention Study. (2018). An application of

data mining to identify potential risk factors for anophthalmia and microphthalmia. Paediatric and Perinatal Epidemiology, 32(6), 545–555. 10.1111/ppe.12509 [PubMed: 30300919]

Williford EM, Howley MM, Fisher SC, Conway KM, Romitti PA, Reeder MR, Olshan AF, Reefhuis J, Browne ML, & National Birth Defects Prevention Study. (2023). Maternal dietary caffeine consumption and risk of birth defects in the National Birth Defects Prevention Study, 1997–2011. Birth Defects Research, 115, 921–932. 10.1002/bdr2.2171 [PubMed: 36942611]

Distribution of preterm birth for selected defects. Distribution of preterm birth for selected defects.

A priori potential risk factors by sex and preterm birth status. a

 $83 (6.2)$ 9 (6.9) 9 (6.9) 83 (63.9) 30 30 31 (23.1)

 $8(6.2)$

 $9(6.9)$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Birth Defects Res. Author manuscript; available in PMC 2024 November 14.

Williford et al. Page 23

 \overline{a}

 Author ManuscriptAuthor Manuscript

Birth Defects Res. Author manuscript; available in PMC 2024 November 14.

Williford et al. Page 24

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Row percentages.

 b Numbers vary because of missing values. Numbers vary because of missing values.

 $\mathbf{\hat{c}_M}$ of the before conception through the second month of pregnancy. Month before conception through the second month of pregnancy.

Birth Defects Res. Author manuscript; available in PMC 2024 November 14.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 3

Adjusted odds ratios and 95% confidence intervals from multinomial regression models for associations between folic acid-containing supplement use,^{4,b} Adjusted odds ratios and 95% confidence intervals from multinomial regression models for associations between folic acid-containing supplement use, maternal smoking,^{4,c} pre-pregnancy obesity,^d interpregnancy interval,^e and sex and preterm birth status. e and sex and preterm birth status. d interpregnancy interval, ^{4, c} pre-pregnancy obesity, maternal smoking,

Author Manuscript

Author Manuscript

Note: The referent group is "female term". The estimates in bold have statistically significant adjusted odds ratios. 3 Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; IPI, interpregnancy interval. nâr Ki ង
រ

Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; IPI, interpregnancy interval.

 $^2\!M$ onth before conception through the second month of pregnancy. Month before conception through the second month of pregnancy.

 b djusted for maternal smoking, pre-pregnancy obesity, and interpregnancy interval. Adjusted for maternal smoking, pre-pregnancy obesity, and interpregnancy interval.

 Δ dijusted for folic acid-containing supplement use, pre-pregnancy obesity, and interpregnancy interval. Adjusted for folic acid-containing supplement use, pre-pregnancy obesity, and interpregnancy interval.

 $d_\mathbf{A}$ djusted for folic acid-containing supplement use, maternal smoking, and interpregnancy interval. Adjusted for folic acid-containing supplement use, maternal smoking, and interpregnancy interval.

Adjusted for folic acid-containing supplement use, maternal smoking, and pre-pregnancy obesity. Adjusted for folic acid-containing supplement use, maternal smoking, and pre-pregnancy obesity.

TABLE 4

Top 10 most important predictors from the random forests for each model by birth defect.

Cleft palate without cleft lip

during P1

Male preterm Female preterm Male term

Total choline (mg/day) Maternal education *Father employed*

Timing of pregnancy discovery Folic acid-containing supplement use

Longitudinal/intercalary limb deficiency

Cantaloupe (1/4 melon) *Lutein and zeaxanthin (μg/day)* Number of people supported with household income Parity Paternal age at delivery Paternal race/ethnicity Maternal age at delivery **Maternal race/ethnicity**

Cantaloupe (1/4 melon) Showers at home Broccoli (1/2 cup) Total carbohydrate (g/day) Chicken/turkey (4–6 oz.) Other cheese (slice or 1 oz.) Cigarette smoking during P2 Copper (mg/day) Magnesium (mg/day) Gravidity

Tomatoes/tomato juice

Peanut butter (1 Tbs)

Birth Defects Res. Author manuscript; available in PMC 2024 November 14.

Male preterm Female preterm Male term Total choline (mg/day) *Maternal race/ethnicity Oranges (1)*

Note: Predictors in bold have a statistically significant odds ratio less than one, and predictors in bold and italics have a statistically significant odds ratio greater than one (see Tables S1–S9 for all model estimates from the logistic regressions). The referent group is "female term".

Abbreviations: B1, month before conception; DFE, dietary folate equivalents; NSAIDs, nonsteroidal anti-inflammatory drugs; P1, first month of pregnancy; P2, second month of pregnancy; P3, third month of pregnancy; RAE, retinoic acid equivalents.

 α Household smoke exposure is defined by the presence of anyone in the household smoking cigarettes.