JAMIA Open, 2024, **7(1)**, ooae024
https://doi.org/10.1093/jamiaopen/ooae024
Research and Applications

Research and Applications

Diving into CDC pregnancy data in the United States: longitudinal study and interactive application

Andrea O. Clark-Sevilla, MS*,[1], Yun C. Lin, MS[1], Arnav Saxena, MS[2], Qi Yan, PhD[3], Ronald Wapner, MD[3], Anita Raja, PhD[4], Itsik Pe'er, PhD[1], Ansaf Salleb-Aouissi, PhD[1]

[1]Department of Computer Science, Columbia University, New York, NY 10027, United States, [2]Department of Data Science, Columbia University, New York, NY 10027, United States, [3]Department of Obstetrics and Gynecology, Columbia University, New York, NY 10027, United States, [4]Department of Computer Science, CUNY Hunter College, New York, NY 10065, United States

*Corresponding author: Andrea Clark-Sevilla, MS, Department of Computer Science, Columbia University, 1214 Amsterdam Ave, 6LW1 Schapiro CEPSR, New York, NY 10027, United States (aoc2111@columbia.edu)

Abstract

Objective: Preterm birth (PTB) is a major determinant of neonatal mortality, morbidity, and childhood disability. In this article, we present a longitudinal analysis of the risk factors associated with PTB and how they have varied over the years: starting from 1968 when the CDC first started, reporting the natality data, up until 2021. Along with this article, we are also releasing an RShiny web application that will allow for easy consumption of this voluminous dataset by the research community. Further, we hope this tool can aid clinicians in the understanding and prevention of PTB.

Materials and Methods: This study used the CDC Natality data from 1968 to 2021 to analyze trends in PTB outcomes across the lens of various features, including race, maternal age, education, and interval length between pregnancies. Our interactive RShiny web application, CDC NatView, allows users to explore interactions between maternal risk factors and maternal morbidity conditions and the aforementioned features.

Results: Our study demonstrates how CDC data can be leveraged to conduct a longitudinal analysis of natality trends in the United States. Our key findings reveal an upward trend in late PTBs, which is concerning. Moreover, a significant disparity exists between African American and White populations in terms of PTB. These disparities persist in other areas, such as education, body-mass index, and access to prenatal care later in pregnancy.

Discussion: Another notable finding is the increase in maternal age over time. Additionally, we confirm that short interpregnancy intervals (IPIs) are a risk factor for PTBs. To facilitate the exploration of pregnancy risk factors, infections, and maternal morbidity, we developed an open-source RShiny tool called CDC NatView. This software offers a user-friendly interface to interact with and visualize the CDC natality data, which constitutes an invaluable resource.

Conclusion: In conclusion, our study has shed light on the rise of late PTBs and the persistent disparities in PTB rates between African American and White populations in the US. The increase in maternal age and the confirmation of a short IPI as a risk factor for PTB are noteworthy findings. Our open-source tool, CDC NatView, can be a valuable resource for further exploration of the CDC natality data to enhance our understanding of pregnancy risk factors and the interaction of PTB outcomes and maternal morbidities.

Lay Summary

This work utilized CDC Natality data spanning from 1968 to 2021 to uncover longitudinal trends and to further explore risk factors of preterm birth (PTB). An interactive RShiny web application, CDC NatView, is also being made available to the greater research community to facilitate easy access of this extensive dataset. The study revealed concerning trends, including an increase in late PTBs and persistent disparities between African American and White populations regarding PTB rates, education, body-mass index, and access to prenatal care. Additionally, it highlighted the rising maternal age and confirmed that short intervals between pregnancies increase the risk of PTB. Overall, this research enhances our understanding of PTB and provides a valuable tool for further exploration and prevention efforts in maternal and child health.

Key words: CDC; data analysis; maternal morbidity; R Shiny; preterm birth; longitudinal study.

Background

The prediction of preterm birth (PTB) has been a challenging problem, mainly due to the inherent complexity of its multifaceted etiology. The World Health Organization defines a PTB as one occurring before 37 weeks of gestation and can be further broken down into subcategories: extreme preterm, occurring before 28 weeks, severe preterm, occurring between 28 and 31 weeks, moderate preterm occurring between 32 and 33 weeks, and late preterm, occurring between 34 and 36 weeks. PTB is the leading cause of mortality and long-term disabilities among neonates, ranging from visual/hearing impairment, cerebral palsy, and mental retardation to increased likelihood of cardiovascular disease, hypertension, and diabetes later in life.[1]

The exact underlying factors contributing to PTB are unknown but are believed to stem from various mechanisms, including inflammation or infection, uteroplacental ischemia or hemorrhage, uterine distension, maternal stress, and the

Received: March 31, 2023; **Revised:** September 20, 2023; **Editorial Decision:** February 2, 2024; **Accepted:** March 5, 2024

immune pathway.[2] In addition, several maternal characteristics have been identified as being influential to PTB, including nutrition, pregnancy history, adverse behaviors such as drinking and smoking, cervical length, and other genetic markers.[3]

Despite having insight into possible risk factors, PTB is still a growing threat to infant mortality. In the United States, the rate of PTB is stated to be around 12%-13%, but most importantly, has been on the rise. From 1981 to 2005, this rate increased from 9.5% to 12.7%.[3] Furthermore, over 26 billion dollars are spent annually on the delivery and care of infants who are born prematurely in the United States.[4] A major challenge is being able to identify both women who are at the highest risk and lowest risk for very early PTB, to narrow the target demographic and avoid unnecessary and costly interventions.[5,6]

Motivated by the need to study large cohorts to understand the complexity of the PTB mechanism, we selected the CDC Natality data to bridge this gap.[7] Despite this data being publicly available and not having as fine-grained feature availability as other natality datasets, such as the Nulliparous Pregnancy Outcomes Study: Monitoring Mothers-to-Be (nuMoM2b),[8,9] this data have been recorded across nearly sixty years for millions of women each year, thereby making it incredibly rich as a longitudinal cohort study for stakeholders. As the CDC keeps updating this data yearly to incorporate new birth records, this additional data will continue to inspire the development of our CDC NatView tool, encouraging a continued exploration of this rich data for the foreseeable future. Our use of the tool has so far shed interesting insights, such as a clearer understanding of the racial component of PTB outcome through its association with education, body-mass index (BMI), and prenatal care, the role maternal age plays in increasing the risk of PTB incidence, and how increasing the interval between pregnancies can be a preventive measure against PTB. Furthermore, our application provides an accessible means through which to interact with this voluminous dataset, which might otherwise deter potential users from manipulating and reaping insights from this rich source of natality data.

Methods

The CDC offers natality datasets that report various features of every live birth in the United States from the Natality Records 1995 to 2021 provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.[7] The data have been aggregated by various demographics, health, and behavioral characteristics. A lay user wanting to query the data for longitudinal trends and interactions between risk factors, demographics, and so forth, would face challenges working with the data in its raw format. The CDC has updated its reporting methodology throughout the years, which has led to inconsistencies in how key features are encoded, presenting a major obstacle for researchers wanting to utilize this data. Currently, the CDC provides a tool that allows public access to the natality data, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER).[7] The data are presented in pre-specified year brackets (1995-2002, 2003-2006, 2007-2021, 2016-2021), and allows the user to perform a series of aggregations on a standardized format of the data. This tool can bypass the issue regarding the inconsistencies in data reporting throughout the years by providing a processed set of features that have been curated to remove missing entries and mitigate the reporting of features that have changed or have been included throughout the years, importantly race, maternal risk factors, BMI, and prenatal care.[7] Using aggregated views provided by the CDC WONDER tool, we offer a unique perspective on this data in the form of an RShiny[10] web application, which we call CDC NatView.

Our tool presents users with a curated subset of features identified as diagnostically relevant by clinicians, including demographic features such as race and maternal age, as well as details about the pregnancy itself, such as the interval since the last pregnancy [interpregnancy interval (IPI)], delivery method, and prenatal care. As many features of the earliest issues of the CDC Natality data were not reported or reported in a manner inconsistent with subsequent years, years before 1995 were excluded from our tool. Our application leverages the CDC WONDER preprocessing mechanism. The interface provides several key features, including maternal demographics such as age, residence, race, Hispanic origin, and education, details about the birth, such as birth weight, plurality, delivery method, and gestational age based on the Last Menstrual Period or Obstetrician's Estimate, and maternal risk factors, such as chronic hypertension, diabetes, eclampsia, and smoking. Users interested in the inclusion criteria and reporting mechanism for the included features are encouraged to visit the data summary page provided by the CDC (https://wonder.cdc.gov/wonder/help/natality.html).

CDC NatView allows users to dive deeper into known pregnancy risk factors, infections, and maternal morbidity conditions longitudinally specified in the aggregation bracket. Furthermore, users can visually explore the interactions of said maternal conditions with demographic information about the mother, including race, pregnancy outcome, gestational age, and the start of prenatal care (eg, incidence of pre-pregnancy diabetes by race). The trend of PTB incidence can also be spatially inspected in the map in the first panel, with raw counts and percentages to account for the possibility that pregnancy reporting might be more widespread and consistent in later issues of the Natality data. CDC NatView also provides a panel for conducting odds ratio analysis to measure the strength of association between said demographic features and the provided maternal conditions. The hope is that such a tool will provide researchers interested in understanding the multifaceted challenge of PTB with a quick and easy way to explore interactions between previously reported risk factors and maternal demographics.

The visualizations have been generated using the highcharter package[11] and the odds ratio analysis using the epitools package.[12] The source code for CDC NatView is freely available on GitHub (https://github.com/PRAISE-Lab-Repository/CDCNatView) and the web application can be accessed through the ShinyApps server (https://mmo7d7-adam-lin.shinyapps.io/CDCNatView/).

Figure 1 shows a screenshot of the interface showing the risk factor of pre-pregnancy diabetes by BMI (lower, left panel) and its effect on the obstetric estimate of gestational age across the years (upper, right panel). A user wanting to explore a specific risk factor or PTB outcome should first select a year bracket of interest corresponding to one of the four databases listed on the top panel. The user can then select the desired condition from the Patient Condition dropdown menu, and select Update Dashboard at the top left

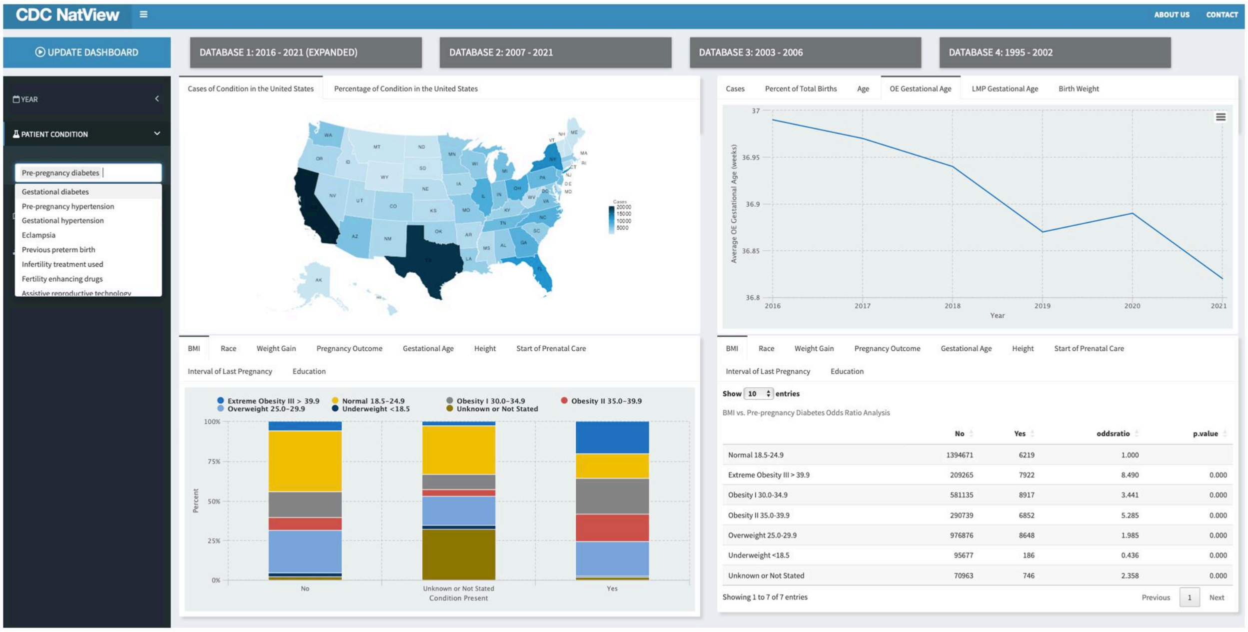

Figure 1. Screenshot of CDC NatView navigator.

The CDC NatView application contains a state incidence visualizer (top-left), a longitudinal plot of a handful of features (top-right), a bar plot to allow understanding of feature interaction with PTB outcome (bottom-right), and an odds ratio analysis to measure the strength of association of a given feature to PTB outcome.

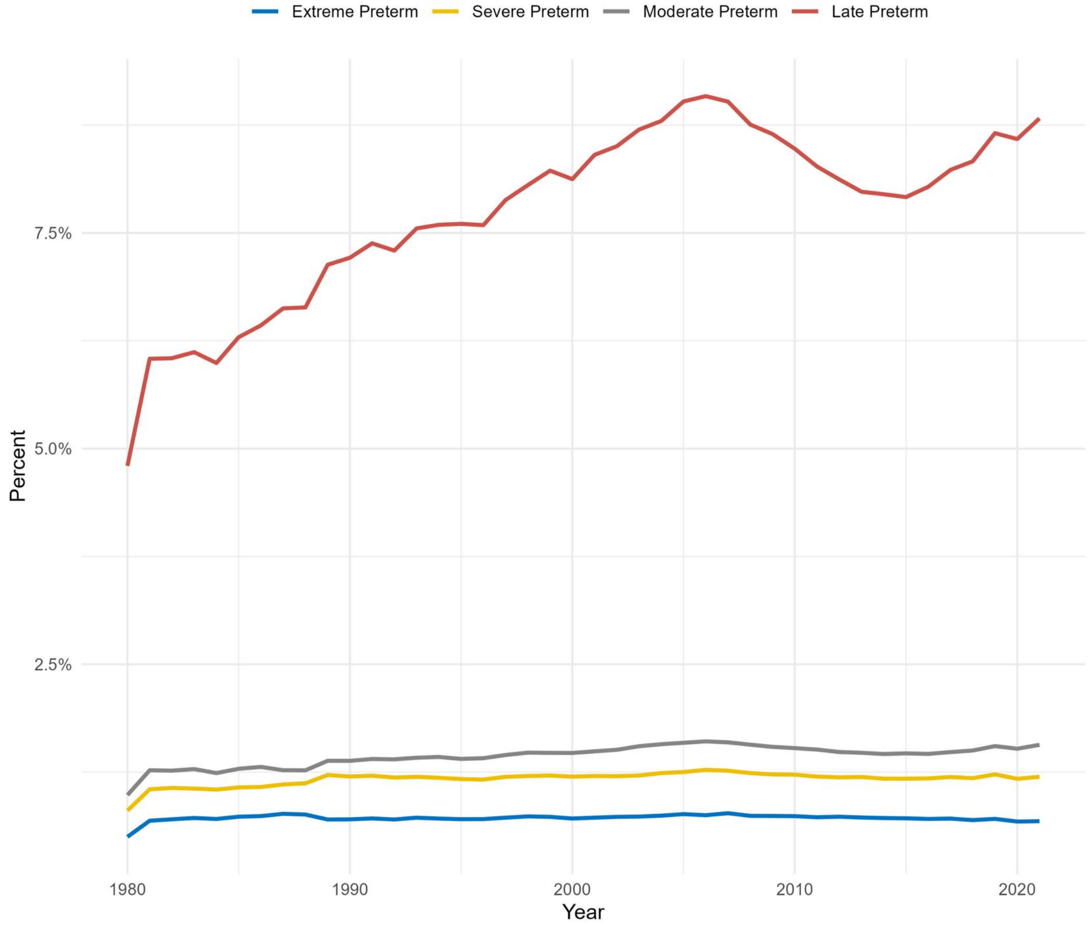

Figure 2. PTB incidence (1980-2021).

The rate of PTB has been rising since the 1970s, with the late preterm category exhibiting the sharpest increase.

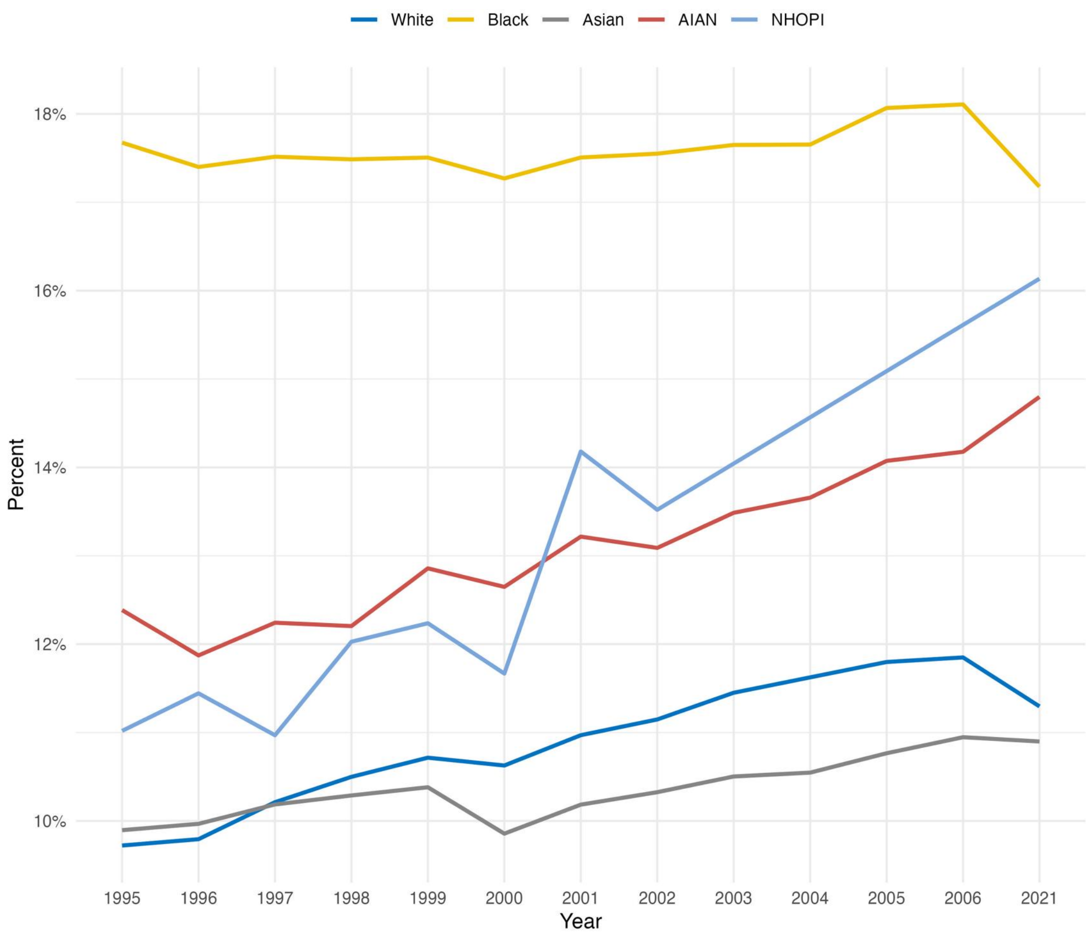

Figure 3. PTB incidence by race (1995-2021).
The figure shows the count normalized rate of PTB incidence by race, with the African American population experiencing the highest incidence by a large margin.

corner of the navigator. The application will update the four panels accordingly. The first panel in the top, left corner showcases the incidence of the selected patient condition at the national level, for instance, gestational diabetes, which can be displayed as either raw counts or percentages by selecting the corresponding tab at the top of this panel. The second panel in the top, right corner visualizes longitudinal trends for the selected database, including raw counts of the selected patient condition, percentages of total births, gestational age, and birth weight. The third panel in the bottom left corner displays demographic information, as well as information relating to the pregnancy, such as the start of prenatal care and the interval between the last pregnancy. The fourth panel in the bottom, right corner displays odd ratios and corresponding *P* values to measure the strength of association between the selected patient condition and a desired feature of interest, including BMI, race, and start of prenatal care. A desired feature can be toggled by selecting it from the tabs directly above the generated charts for panels 2-4. The development will continue with suggestions and improvements coming from its users.

Results

Here we present a sampling of key features that have yielded interesting insights and have been addressed through the use of CDC NatView.

PTB incidence

Figure 2 indicates that while PTB has overall been on the rise, the major PTB category driving this increase is the Late Preterm category, which has seen no decline since the 1970s. A Chi-Square Goodness of Fit test was performed to examine the role of maternal race on PTB outcome. The relation between these variables was significant, $\chi^2(5,\ N=136\,479\,230) = 859\,301$, $P<.001$. Further, Figure 3 highlights a difference between the African American and White populations, and all other races for that matter, but since around 2010, this gap has begun to close. Beginning from around 2010, it appears that the Asian, Native Hawaiian, and Other Pacific Islander (NHOPI) population has experienced a rather drastic increase in PTB cases, which warrants further investigation as to the exact cause.

Education

Figure 4 showcases the relationship between maternal race and the highest educational attainment, with snapshots from years ranging from 1970 to 1990 and 2010 to 2020. A Chi-Square Goodness of Fit test showed a statistically significant difference between race and highest educational attainment, $\chi^2(30,\ N=46\,196\,010) = 1\,989\,596$, $P<.001$. The disparity between the scholastic achievement between African American and White mothers is quite evident when it comes to higher education. From years ranging from 2010 to 2020, 11.79% of African American mothers reported having completed a college degree and 6.24% completing a graduate or

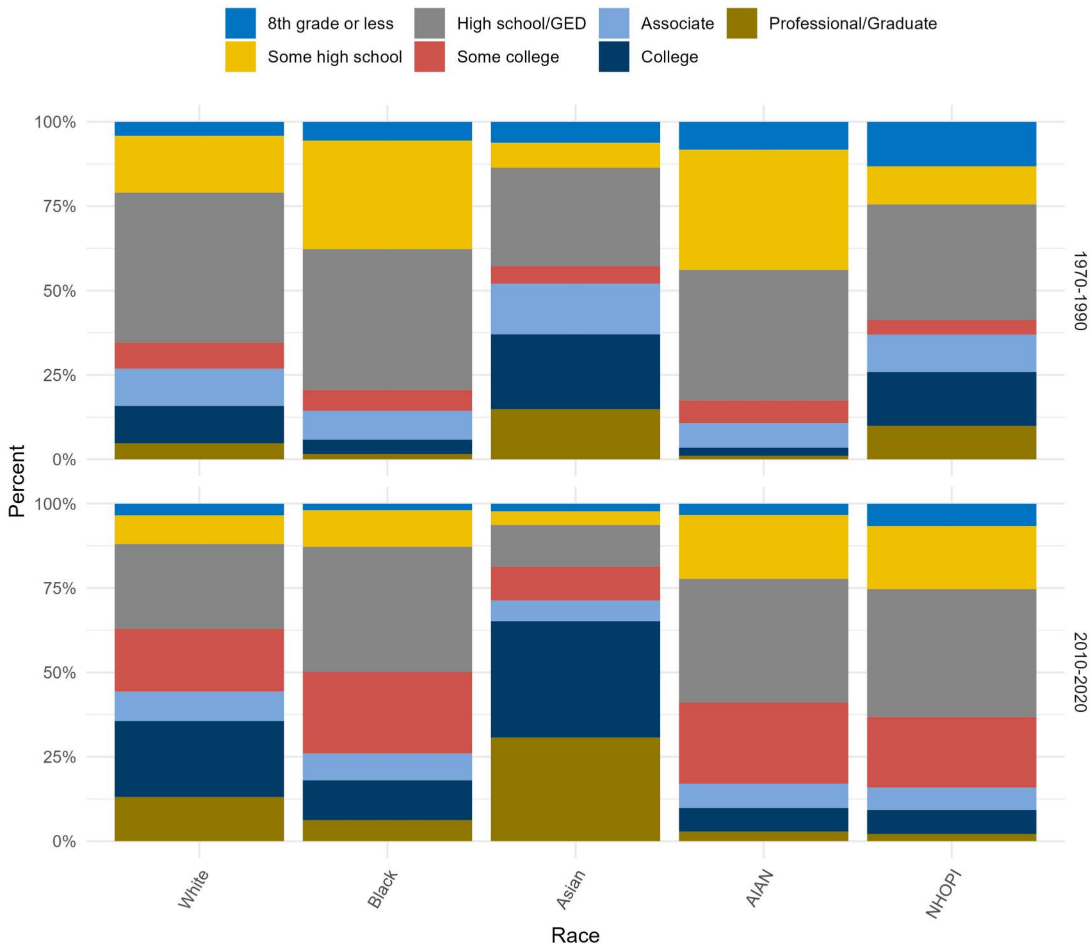

Figure 4. Highest educational attainment by race (1970-1990, 2010-2020).

The bar plot shows the count normalized proportions of the highest educational attainment when grouped by race. There is a clear trend toward women receiving more education, as witnessed by the increased proportion of College and Professional/Graduate studies from 2010-2020, as compared to 1970-1990. The Asian and White populations exhibit the highest proportions of advanced education attainment (College, Professional/Graduate), and the AIAN and NHOPI have the lowest such proportions.

professional degree, versus 22.56% and 13.06% of White mothers, respectively. In contrast, Asian mothers exhibit the highest educational attainment (34.51% completed college, 30.68% completed graduate or professional degree). When looking at educational attainment versus PTB outcome, a Chi-Square Goodness of Fit test yielded a significant statistical difference, $\chi^2(6, N = 98\,248\,260) = 229\,910$, $P < .001$.

BMI

A Chi-Square Goodness of Fit test showed that the proportion of PTB differs by maternal race $\chi^2(4, N = 136\,479\,230) = 857\,721$, $P < .001$. Figure 5 exhibits a significant difference between African American and White mothers for the obese and extremely obese weight categories, with approximately 18.92% and 22.18% of African American mothers falling into these classifications, respectively, versus 15.59% and 14.68% of White mothers. The American Indian and Alaskan Native (AIAN) Native Hawaiian and Other Pacific Islander (NHOPI) report the highest proportions of mothers in the obese to extremely obese categories (21.21% of the NHOPI population falls into the extremely obese category), while on the other spectrum, Asian mothers consistently fall into the lower-weight brackets, with only 8.71% and 5.03% of the Asian population falling into the obese category and extremely obese categories, respectively. Looking at the relationship of maternal BMI on PTB outcome, a Welch Two Sample *t*-test yielded a statistically significant difference between the prepregnancy BMI of women who delivered full-term ($M = 26.3$, $SD = 5.62$) versus that of women who delivered preterm ($M = 26.7$, $SD = 5.90$), $t(5\,008\,777) = -138$, $P < .001$.

Maternal age

Figure 6 exhibits the overall trend of maternal age increasing over time, beginning from 1970 up until 2020. The average maternal age from 1970 to 1990 was 25.3, whereas, from 2010 to 2020, this rose to 28.5. Women who delivered preterm babies were found to be younger on average ($M = 26.4$, $SD = 6.3$) than did women who delivered full-term babies ($M = 26.9$, $SD = 5.9$), $t(29\,467\,851) = 308$, $P < .001$. In addition, cases of extreme PTB were found to be associated with lower-than-average maternal age (24.9) versus full-term cases (27.26).

Start of prenatal care & interpregnancy interval

A Chi-Square Goodness of Fit test showed that race plays a statistically significant role in whether prenatal care is started early (during the first trimester) or late (after the first

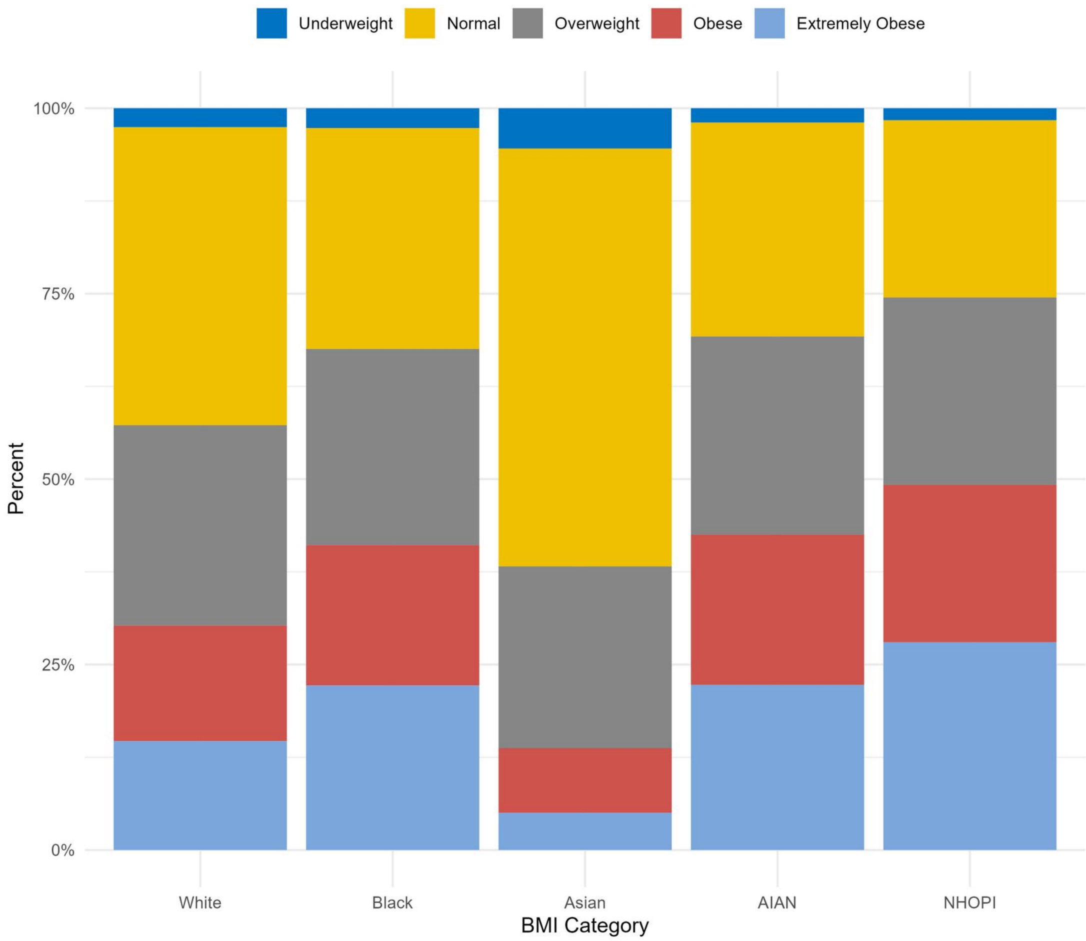

Figure 5. BMI by race (1980-2021).

The bar plot shows the count normalized proportions of BMI categories when grouped by race. The African American, AIAN, and NHOPI populations exhibit the highest proportions of elevated BMI (overweight, obese, extremely obese), with the Asian population showing the lowest proportion.

trimester or not at all) $\chi^2(6, N = 1\,654\ 388) = 2$, $P < .001$. Further, the relationship between whether prenatal care was started early or late was significant to PTB outcome, $\chi^2(1, N = 155\,683\ 639) = 78\,378$, $P < .001$. A Welch Two Sample *t*-test on the data from 1968 up until 2021 revealed that mothers who delivered full-term were associated with a higher IPI ($M = 43.9$, $SD = 33.7$) than mothers who had a preterm delivery ($M = 41.0$, $SD = 38.0$), $t(8\,289\ 593) = 193$, $P < .001$. Indeed, we found a consistent pattern between extreme PTB cases and mothers who had an IPI of less than 25 months, as Figure 7 depicts.

Discussion

The CDC data provides a rich basis from which to frame and develop research questions. A few key features were analyzed for this discussion, and to showcase potential use cases for the CDC NatView application.

PTB incidence

Previous studies have corroborated our findings that PTB is on the rise. Shapiro-Mendoza and Lackritz found that of all PTBs, 72% of these births are from late PTBs.[13] While infant survival is higher for those born closer to a full-term birth, late preterm babies still face higher rates of infant morbidity and mortality, as well as higher risks of childhood disabilities.[13] Further, another challenge PTB poses is the apparent disparity between racial groups.

A large disparity between the African American and White populations has been reported concerning PTB outcome, with the former experiencing nearly 1.5-1.6 times higher than the latter.[14] In 2008, late PTB rates were highest for infants of non-Hispanic, African American mothers (11.3%), followed by American Indian or Alaskan Natives (AIAN) (9.7%), Hispanics (8.8%), non-Hispanic whites (8.2%), and Asian or Pacific Islanders (7.9%).[13] Our findings similarly confirmed this racial gap, with the African American population having a significantly greater proportion of PTB outcomes than that of the Asian, AIAN, and NHOPI populations.

Education

Low socioeconomic and educational status has been previously linked to PTB.[15–17] Our findings aimed at exploring the connection between race and PTB incidence bolstered the credibility of association studies made between low educational status and PTB incidence, as East Asian women typically present the lowest PTB rates across races.[3]

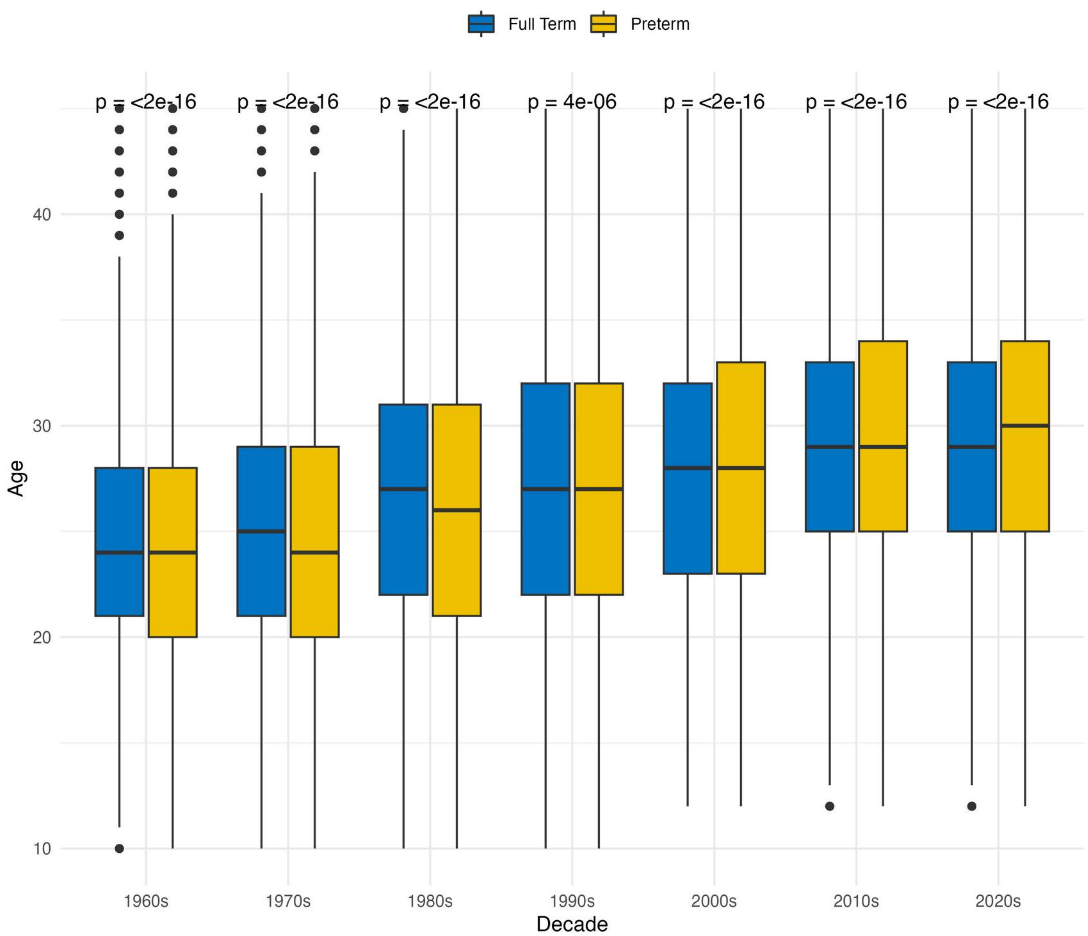

Figure 6. Maternal age by year (1970-2021).

Maternal age is on the rise, as can be seen in the boxplot. The average maternal age in recent years (2020s) is approximately 28.5, as compared to approximately 25.3 from the 1970s to 1990s.

BMI

Studies related to the association of pre-pregnancy BMI and PTB have been inconsistent. Some studies report that women with low prepregnancy BMI are at an increased risk for PTB,[18,19] while women with higher BMI (overweight or obese) have been reported to have the same or reduced risk as women within the normal BMI range,[20–22] while other studies claim there is an increased risk for PTB.[23,24]

A CDC National Health Statistics Report on obesity rates from 2017 to 2020 reported that obesity was most prevalent among non-Hispanic African American adults (49.9%), versus non-Hispanic White (41.4%) and non-Hispanic Asian (16.1%), with no difference observed between men and women overall. However, a higher obesity rate was reported for non-Hispanic African American women, as compared to their male counterparts.[25] Our exploration with CDC NatView yielded a significant difference between African American and White mothers for the obese and extremely obese categories, which bolsters findings that suggest obesity rates are higher for the African American population. However, further exploration needs to be done on the association of race and BMI to validate whether or not this racial trend in BMI underlies PTB outcome.

Maternal age

Previous studies have discovered a "U-shaped" association between maternal age and PTB incidence, in which mothers that are at either age extreme are at an increased risk for PTB.[26–28] Furthermore, maternal age has been increasing over the years, which might be placing older mothers (≥35 years) at a higher risk for PTB incidence.[29,30] Our exploration with CDC NatView found an increasing trend for maternal age and an increased incidence of PTB for women on extremes of the age spectrum, which provides further evidence to bolster the aforementioned studies.

Start of prenatal care & interpregnancy interval

Mothers lacking in prenatal care are at a higher risk of PTB, hence early start to prenatal care can be seen as a preventive measure for preterm delivery. In fact, the absence of prenatal care has been associated with a 2-fold increase in the risk of delivering a preterm baby.[31] Our analysis of the data found a statistically significant difference between racial groups and when prenatal care began. This finding is rather unique in that it has a clear actionable point, whether in providing more accessible prenatal care or through the healthcare community stressing the importance of consistent and early prenatal care for a healthy delivery.

It has been found that a short IPI, which is defined as a lapse of fewer than 18 months between pregnancies, is a risk factor for PTB in women of extreme age groups.[32,33] Studies on a cohort of women from the Tennessee Birth Statistical files found that women with a short IPI were on average

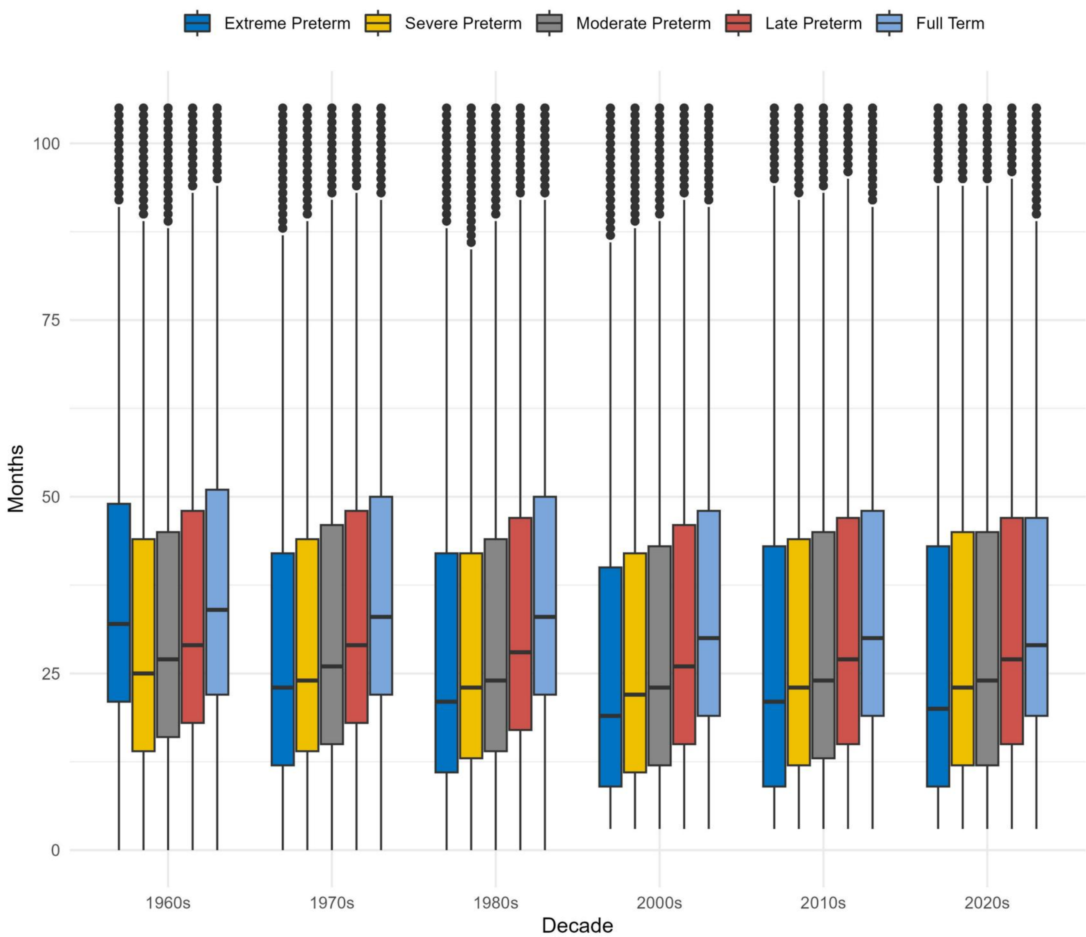

Figure 7. IPI (1960s-2020s).

A shorter IPI is associated with a greater risk of a PTB. With the exception of the 1960s, the shorter the IPI, the more likely the infant is associated with a preterm category of increasing severity.

younger, had lower educational attainment, lower income, higher BMIs, and were more likely to be unmarried, smokers, and have begun prenatal care later.[34] The CDC data confirmed this pattern, showing that mothers who delivered early were associated with a shorter IPI, specifically with cases of extreme PTB.

As can be appreciated from the relevant features and their findings, the CDC NatView tool shows a lot of potential for better understanding reported risk factors for PTB and providing a significant cohort with which to confirm the validity of such findings. However, the tool is not without its limitations. The CDC Natality data presented many encoding inconsistencies throughout the years, importantly for how race was encoded. Further, the 2003 revision implementing the U.S. standard Certificate of Live Birth rendered some features incompatible with previous revisions, resulting in significant changes to the features available thereafter.[7] For this reason, the data aggregations done by year brackets were a necessary component of CDC NatView, as is provided by CDC WONDER's data querying tool. An early version of our application used a standardized version of the data from 1960 to 2021 we curated, but many inconsistencies appeared longitudinally for it to be useful as a visual tool. Having the inability to query the entire span of the data undermines the ability to truly appreciate longitudinal trends, such as the incidence of PTB across years. However, despite this limitation, the data are vast enough at a nationwide level for such patterns to emerge, which is a major strong point of this tool.

Further, the tool provides a powerful, exploratory, and visual complement to the interface provided by CDC WONDER. CDC NatView provides ready-to-use features that the user can explore rendered, interactive camera-ready charts. The CDC WONDER tool requires the user to have to bypass a series of aggregation filters to acquire the necessary data, which would deter many users from interacting with it.

Another point of improvement remaining for our tool would be the incorporation of querying by more than one feature at a time. For instance, we found a significant association between BMI and PTB outcome, but the current version of the tool would not allow grouping by race as well, to understand if race and BMI interact together in a significant way to drive PTB outcome. This feature will be integrated into a subsequent version of our application, as the need for such multifaceted analysis is crucial for more sophisticated research questions that could potentially be addressed by the data.

Conclusion

In this work, we analyzed CDC natality data from 1968 to 2020, drawing insights from interactions of various risk factors previously identified as driving PTB, and developed a web application for supporting this exploration process using the CDC WONDER aggregation tables for portability to the RShiny framework and quick response time to user queries. The hope is that the larger research community will find the interface helpful in exploring the otherwise dense but incredibly rich dataset, leveraging the vast number of samples to uncover hidden associations between maternal morbidity factors and demographic features at a statistically significant level. CDC NatView has been developed entirely in R, an open source programming language, making it free to use, share, and modify according to its users' needs. Future work will involve scraping data from CDC WONDER directly through an API connected to CDC NatView to provide more patient conditions to filter by, thereby adding more potential interactions between demographics features, in addition to providing instant access to the latest data.

Author contributions

The CDC Natality dataset was compiled and interpreted by Andrea O. Clark-Sevilla and Arnav Saxena, the CDC NatView RShiny web application and its data acquisition process was developed by Yun C. Lin and Andrea O. Clark-Sevilla. Ansaf Salleb-Aouissi conceived the project idea and provided significant guidance in framing research questions. All authors contributed to the web application design and reviewed and approved the final manuscript.

Funding

This work was supported by NIH/NLM (www.nlm.nih.gov) grant number R01LM013327.

Conflict of interests

All authors declare no financial or non-financial competing interests.

Data availability

The dataset used in the analysis and made accessible through the CDC NatView RShiny web application is available through the National Bureau of Economic Research (NBER) at https://data.nber.org/data/natality.html

References

1. Zerhouni EA. Prematurity Research at the NIH. 2008. Accessed August 10, 2022. https://www.nichd.nih.gov/sites/default/files/publications/pubs/documents/nih_prematurity_research_2008.pdf
2. Romero RJ, Espinoza J, Kusanovic JP, et al. The preterm parturition syndrome. *BJOG*. 2006;113(Suppl 3):17-42.
3. Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet*. 2008;371(9606):75-84.
4. Institute of Medicine (U.S.) Committee on Understanding Premature Birth and Assuring Healthy Outcomes. In: Behrman RE, Butler AS, eds. *Preterm Birth: Causes, Consequences, and Prevention*. National Academies Press; 2007. http://books.google.com/books?id=9c_7kxBsKzIC
5. Vovsha I, Rajan A, Salleb-Aouissi A, et al. *Predicting Preterm Birth is Not Elusive: Machine Learning Paves the Way to Individual Wellness*. AAAI Spring Symposia; 2014.
6. Vovsha I, Salleb-Aouissi A, Raja A, et al. Using Kernel Methods and Model Selection for Prediction of Preterm Birth. In *Proceedings of the 1st Machine Learning for Healthcare Conference (PMLR)*. *Proceedings of Machine Learning Research*, Vol. 56. 2016:55-72. https://proceedings.mlr.press/v56/Vovsha16.html
7. CDC Wonder. Accessed August 10, 2022. https://wonder.cdc.gov/natality.html
8. Silver R. *Nulliparous Pregnancy Outcomes Study: Monitoring Mothers-to-Be (Version 1) [Dataset]*. NICHD Data and Specimen Hub; 2019. https://doi.org/10.57982/gjxm-yz73
9. Goretsky A, Dmitrienko A, Tang I, et al. Data Preparation of the nuMoM2b Dataset. *medRxiv*. 2021.
10. Shiny. Accessed August 10, 2022. https://shiny.rstudio.com/
11. Kunst J. *Highcharter: A Wrapper for the "Highcharts" Library*. 2022. Accessed August 10, 2022. https://CRAN.R-project.org/package=highcharter
12. Aragon TJ. *Epitools: Epidemiology Tools*. 2020. Accessed August 10, 2022. https://CRAN.R-project.org/package=epitools
13. Shapiro-Mendoza CK, Lackritz EM. Epidemiology of late and moderate preterm birth. *Semin Fetal Neonatal Med*. 2012;17(3):120-125.
14. Braveman P, Dominguez TP, Burke W, et al. Explaining the Black-White disparity in preterm birth: a consensus statement from a Multi-Disciplinary scientific work group convened by the march of dimes. *Front Reprod Health*. 2021;3:684207.
15. Smith L, Draper ES, Manktelow BN, et al. Socioeconomic inequalities in very preterm birth rates. *Arch Dis Childhood—Fetal Neonatal Ed*. 2006;92(1):F11-F14.
16. Brett KM, Strogatz DS, Savitz DA. Employment, job strain, and preterm delivery among women in North Carolina. *Am J Public Health*. 1997;87(2):199-204.
17. Thompson JM, Irgens LM, Rasmussen S, et al. Secular trends in socio-economic status and the implications for preterm birth. *Paediatr Perinat Epidemiol*. 2006;20(3):182-187.
18. Sebire NJ, Jolly M, Harris J, et al. Is maternal underweight really a risk factor for adverse pregnancy outcome? A population-based study in London. *BJOG*. 2001;108(1):61-66.
19. Siega-Riz AM, Adair LS, Hobel CJ. Maternal underweight status and inadequate rate of weight gain during the third trimester of pregnancy increases the risk of preterm delivery. *J Nutr*. 1996;126(1):146-153.
20. Sebire NJ, Jolly M, Harris J, et al. Maternal obesity and pregnancy outcome: a study of 287 213 pregnancies in london. *Int J Obes*. 2001;25(8):1175-1182.
21. Yaacob S, Saad FA, Sharara H, et al. The effect of obesity in pregnancy on perinatal outcome in Qatar. *Qatar Med J*. 2002;2002(2):16.
22. Doherty DA, Magann EF, Francis J, et al. Pre-pregnancy body mass index and pregnancy outcomes. *Int J Gynecol Obstet*. 2006;95(3):242-247.
23. Cedergren MI. Maternal morbid obesity and the risk of adverse pregnancy outcome. *Obstet Gynecol*. 2004;103(2):219-224.
24. Baeten JM, Bukusi E, Lambe M. Pregnancy complications and outcomes among overweight and obese nulliparous women. *Am J Public Health*. 2001;91(3):436-440.
25. National Health Statistics Reports. Accessed August 10, 2022. https://stacks.cdc.gov/view/cdc/106273
26. Fuchs F, Monet B, Ducruet T, et al. Effect of maternal age on the risk of preterm birth: a large cohort study. *PLoS One*. 2018;13(1):e0191002.
27. Ferré C, Callaghan W, Olson C, et al. Effects of maternal age and Age-Specific preterm birth rates on overall preterm birth rates United States 2007 and 2014. *MMWR Morb Mortal Wkly Rep*. 2016;65(43):1181-1184. https://doi.org/10.15585/mmwr.mm6543a1
28. Esposito G, Mauri PA, Cipriani S, et al. The role of maternal age on the risk of preterm birth among singletons and multiples: a retrospective cohort study in lombardy Northern Italy. *BMC Pregnancy Childbirth*. 2022;22(1):234. https://doi.org/10.1186/s12884-022-04552-y

29. Aoyama K, Pinto R, Ray JG, et al. Association of maternal age with severe maternal morbidity and mortality in Canada. *JAMA Network Open*. 2019;2(8):e199875.
30. Koo Y-J, Ryu HM, Yang JH, et al. Pregnancy outcomes according to increasing maternal age. *Taiwan J Obstet Gynecol*. 2012;51 (1):60-65.
31. Vintzileos AM, Ananth CV, Smulian JC, et al. The impact of prenatal care in the United States on preterm births in the presence and absence of antenatal high-risk conditions. *Am J Obstet Gynecol*. 2002;187(5):1254-1257.
32. Nerlander L, Callaghan WM, Smith RA, et al. Short interpregnancy interval associated with preterm birth in US adolescents. *Matern Child Health J*. 2015;19(4):850-858.
33. Ihongbe TO, Wallenborn JT, Rozario SS, et al. Short interpregnancy interval and adverse birth outcomes in women of advanced age: a population-based study. *Ann Epidemiol*. 2018;28(9):605-611.
34. Appareddy S, Pryor JR, Bailey BA. Inter-pregnancy interval and adverse outcomes: evidence for an additional risk in health disparate populations. *J Matern-Fetal Neonatal Med*. 2017;30 (21):2640-2644.