Pediatric RESEARCH

www.nature.com/pr

REVIEW ARTICLE

The promise and pitfalls of precision medicine to resolve black–white racial disparities in preterm birth

Heather H. Burris[1], Clyde J. Wright[2], Haresh Kirpalani[1], James W. Collins Jr[3], Scott A. Lorch[1], Michal A. Elovitz[4] and Sunah S. Hwang[2]

Differences in preterm birth rates between black and white women are the largest contributor to racial disparities in infant mortality. In today's age of precision medicine, analysis of the genome, epigenome, metabolome, and microbiome has generated interest in determining whether these biomarkers can help explain racial disparities. We propose that there are pitfalls as well as opportunities when using precision medicine analyses to interrogate disparities in health. To conclude that racial disparities in complex conditions are genetic in origin ignores robust evidence that social and environmental factors that track with race are major contributors to disparities. Biomarkers measured in omic assays that may be more environmentally responsive than genomics, such as the epigenome or metabolome, may be on the causal pathway of race and preterm birth, but omic observational studies suffer from the same limitations as traditional cohort studies. Confounding can lead to false conclusions about the causal relationship between omics and preterm birth. Methodological strategies (including stratification and causal mediation analyses) may help to ensure that associations between biomarkers and exposures, as well as between biomarkers and outcomes, are valid signals. These epidemiologic strategies present opportunities to assess whether precision medicine biomarkers can uncover biology underlying perinatal health disparities.

Pediatric Research (2020) 87:221–226; https://doi.org/10.1038/s41390-019-0528-z

INTRODUCTION

In the United States, major racial disparities in health occur across the lifespan resulting in higher mortality rates among non-Hispanic black versus non-Hispanic white Americans.[1,2] Perinatal disparities are among the most striking; black infants are three times more likely than white infants to die in the first year of life.[3] The largest contributor to infant mortality differences between black and white infants is the wide racial disparity in preterm birth (PTB) rate, which is 50% higher among black women.[4–6] As molecular assays generate large-scale biomarker data to profile genomes, epigenomes, metabolomes, immunomes, and microbiomes for individualized precision medicine approaches, interest has grown in whether these "omic" assays can shed light on differences in the pathophysiology of health conditions,[7] including PTB, by race. However, before determining that variation in these molecular profiles is on the causal pathway to racial disparities in health, potential pitfalls and strategies to avoid them should be considered.

The basic pitfall is to assume that black race is a genetic construct and thus not modifiable.[8] The assumption that race genetically determines disease susceptibility arises from extrapolations beyond two facts: first that continental ancestral genetic patterns are somewhat associated with self-identified race; and second, that Mendelian diseases such as sickle cell disease track (although incompletely) along racial lines. Neither of these facts warrant the assumption that race is an unmodifiable biologic construct. That false assumption ignores the large body of literature demonstrating that race (as opposed to ancestry) is largely a social construct with biologic impacts that affect disease risk. The study of genetics to determine susceptibility or resilience to a causative exposure is important. However, applying genetics to the social construct of race is problematic. Using self-identified race will misclassify the genetics of many people who do not map exactly to ancestral patterns that are most common in their self-identified racial group. Disaggregating self-identified race, which may change susceptibility to disease due to societal context, from ancestral patterns of genetic sequence is critical to truly understanding why disparities exist. This is especially so in complex, multifactorial diseases, such as PTB, that reflect variation in lived experiences between racial groups.[9] This is demonstrated by international studies where the consequences of race vary by society. For example, McKinnon et al. showed that black women living in Canada have lower PTB rates than black women living in the United States (8.9% vs 12.7%, respectively).[10] Better birth outcomes among foreign-born black women in the US compared to US-born black women[11,12] highlight that the genetics that determine skin color do not determine adverse birth outcomes. Hence, concurrent differences in continental ancestral genetic sequences and differences in PTB risk by race do not mean that genetic sequence causes differences in PTB risk.

However, precision medicine approaches extend beyond genetic sequence analysis. Assays that measure multiple biomarkers can now be used to interrogate several key steps in the

[1]Division of Neonatology, The Children's Hospital of Philadelphia at the Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [2]Section of Neonatology, Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, CO, USA; [3]Division of Neonatology, Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL, USA and [4]Maternal and Child Health Research Center, Department of Obstetrics and Gynecology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
Correspondence: Heather H. Burris (burrish@email.chop.edu, hburris@pennmedicine.upenn.edu)

Received: 15 April 2019 Revised: 24 July 2019 Accepted: 25 July 2019
Published online: 5 August 2019

gene–environment interaction: transcription (transcriptome),[13,14] translation (proteome),[15] regulation of gene expression (epigenome),[16] host–microbial interactions (immunome and microbiome),[17] exposures and their metabolites (metabolome),[18–20] and their combined synergies (multiomics).[21] Results of these assays could highlight processes by which exposures lead to increased risk of PTB (mechanistic discovery of causal pathways) or may generate biomarkers of risk of PTB (useful for prediction). These two potential applications (mechanism identification and prediction) could be useful for developing therapeutic targets and, diagnostic tests, as well as identifying populations at particularly high risk in whom interventions might be most important. If omic data are to be used to explain the biology underlying disparities in PTB, they must be involved in the causal pathway leading to PTB, especially among black women. A conceptual model (Fig. 1a) shows how omic assays may shed light on how exposures or ancestral sequences could lead to differential biomarker results by race, with the caveat that the relationship between omic biomarkers and outcomes may be confounded by race (Fig. 1b). We describe below how omic assays common in "precision medicine" or "precision public health" may, or may not, be useful to understand and ultimately ameliorate racial disparities in perinatal health.

REDUCING RACIAL DISPARITIES IN PTB AT A POPULATION LEVEL

Resolving racial disparities in health will require two phases of discovery. The first is to determine why there are population differences in risk by race in order to identify causal factors for groups of women at risk (primary prevention). The second is to identify populations at risk and intervene on these populations in the biomedical setting to prevent PTB (secondary prevention).

We first consider how population differences in harmful exposures could lead to racial disparities in health outcomes, such as PTB. There are two potential ways: (1) differences in prevalence of external and universally potent causal factors or (2) differences in susceptibility to causal factors.

Exposure differences by race: the case of universal potency but differing prevalence

One risk factor with similar impacts on black and white women, but which is more prevalent among black women is a short inter-pregnancy interval. The association of intervals <6 months, compared to 18–23 months, has been shown to elevate PTB risk (odds ratio (OR) = 1.22, 99% confidence interval (CI) = 1.11–1.35) even in a within-family study,[22] which avoids confounding by variables that remain constant between pregnancy.[23] In the United States, a study of >4.8 million births examined inter-pregnancy interval examined the odds of low birth weight (<2.5 kg) within racial/ethnic strata in the United States from 1989 to 1991. Among births with a preceding inter-pregnancy interval of <6 months compared to intervals >12 months, higher odds were observed for PTB in non-Hispanic black (OR: 1.64, 95% CI 1.60–1.68) and white (OR: 1.67, 95% CI 1.63–1.70) women,[24] demonstrating similar rates in each racial/ethnic group. The etiology of any disparity resulting from short inter-pregnancy intervals, therefore, results from the higher prevalence of short inter-pregnancy interval among black women (10.5%) compared with white women (5.7%). In a separate analysis, investigators found that the potential impact of increasing inter-pregnancy intervals to 18–23 months would reduce the excess PTB by up to 8% among black and 4% among white women.[25]

Population differences by race: the case of differential susceptibility due to social factors

An example where women may have differential susceptibility to causal factors according to self-identified race may be the association of air pollution and PTB. While all people are susceptible to air pollution, it may be that other exposures that track with self-identified race, specifically socioeconomic disadvantage, may increase susceptibility to the adverse effects of air pollution.[26–28] Heightened susceptibility is not necessarily a genetic predisposition but could occur due to stress-induced immunologic alterations that alter the body's ability to handle an additional physical stressor—analogous to increased susceptibility to viral infections during periods of stress.[29] It is known that black

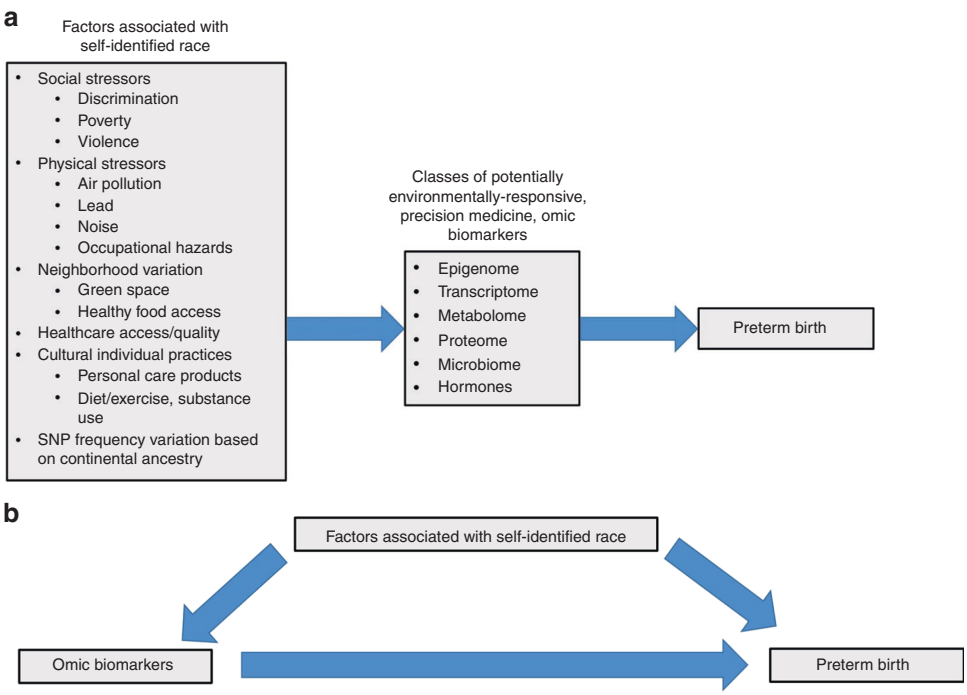

Fig. 1 Conceptual framework of how race could lead to differential omic biomarkers and be associated with racial disparities in preterm birth (**a**) and how the relationship between omic biomarkers could be confounded by race-associated factors (**b**)

families are more likely than white families to have lower socioeconomic position as well as experience discrimination and stress.[30–32] Hence, black women may have increased susceptibility to the health effects of air pollution, including PTB. A recent study of 53,843 births in California demonstrated that PTB risk was associated with environmental pollution especially in areas of low socioeconomic status.[28] Given that air pollution exposures are higher among urban black communities[33–36] than white communities (different doses) and that black women are more likely to be of low socioeconomic position that might heighten susceptibility to air pollution, reduction in air pollution overall could lead to decreased racial disparities in PTB. In fact, there has been some recent evidence that closures of coal power plants in California reduced PTB rates in surrounding areas, particularly among minorities.[37] However, no study has assessed the causal pathway between air pollution reduction and racial disparities in PTB.

Taken together, identifying causal factors either with higher prevalence or heightened susceptibility among black families because of the association between race/ethnicity and additional social factors may help to reduce racial disparities in PTB. However, as physicians and patients, we cannot wait for policy efforts to reduce harmful exposures to go into effect. We need to find ways to help individual women maximize the chance at a healthy, full-term birth. It is here that the promise of "precision medicine" is so exciting.

PRECISION MEDICINE FOR THE PREDICTION OF AN INDIVIDUAL'S RISK OF PTB

Currently, providers' ability to predict which women will have PTB and which women will have a healthy, full-term delivery is poor. An ideal test in the prenatal period would enable obstetricians to identify which women will deliver early, and once identified, providers would then offer an effective therapy to prevent PTB. Imperfect interventions[38,39] currently are: women with prior PTB (at high risk of a subsequent PTB) are offered progesterone,[40] and if found to have a short cervix, they are additionally offered a cerclage[41] and women without a prior spontaneous PTB (sPTB) but with a short cervix are offered progesterone.[42,43] For medically indicated PTB due to conditions such as preeclampsia or intrauterine growth restriction, providers may prescribe aspirin, but it is only mildly protective. The bottom line is that we still do understand the pathogenesis of PTB and thus are limited in strategies to improve birth outcomes.

Precision medicine research has the potential to improve the landscape of pregnancy care. The goal of these studies is, first, to find biomarkers that identify *individual* women who are at particularly high risk of PTB and, second, to identify therapeutic targets to prevent it. Yet, to date, no panel of biomarkers informs an individual woman and her physician of her percentage of chance of PTB. However, such advances may be possible. In a small study, Ngo et al. studied circulating cell-free RNA using RNA sequencing among women at risk for sPTB [training dataset of women with preterm contractions ($n = 15$) and validation cohort with prior sPTB ($n = 23$)] and found that the model performed moderately well.[44] The model misclassified 1 out of the 5 sPTB deliveries, and 3 of the 18 full-term deliveries (area under the receiver operating characteristic curve 0.81). A recent study analyzed the cervicovaginal microbiome and a protein involved in innate immunity, beta defensin 2, in 549 women (107 sPTB and 432 term) at three time points in pregnancy. This study showed that a microbiome lacking *lactobacillus* dominance (community state type IV-B) early in pregnancy is associated with sPTB; 32% of women with sPTB compared to 21% of women with term birth had community state type IV-B ($p = 0.017$).[17] This study also showed that beta defensin 2 levels were lower among women, specifically black women, who went on to have sPTB compared to women with term births. This highlights the complex balance between the microbiome and the immune system to modulate PTB risk. Hence, combining low beta defensin 2 with low relative abundance of *lactobacillus* appears to increase sPTB risk. While black women were more likely to have a higher-risk microbiome with respect to sPTB, they also had higher levels of beta defensin 2 that mitigated that risk. The authors propose that these findings could lead to "innovative therapeutic opportunities to prevent sPTB including combination of microbiome-based therapeutics and immune modulators." However, while associations between the microbiome and immune status are interesting, with respect to prediction, we may require more data (larger studies with more combinations of variables to identify the most informative markers) so as to avoid false positives and false negatives that are common in epidemiologic frameworks that assess risk as opposed to prediction. Nonetheless, if the microbiome and immune status were therapeutic targets that would be useful among black women, such approaches could potentially reduce disparities. These two studies highlight the two promises of precision medicine, respectively: (1) to predict and (2) to identify molecular targets that could be used to prevent an individual from having sPTB.

LIMITATIONS OF PRECISION MEDICINE METHODS TO UNDERSTAND POPULATION'S RISK OF PTB—AND HOW TO IMPROVE THEM

Genomic approaches

Even though omic approaches are not yet ready to predict an *individual* woman's risk of PTB, omic data may shed light on mechanisms by which risk factors lead to differential risk of PTB among *populations* and thus racial disparities. Genome-wide association studies (GWAS) have been limited by several factors. The first is under representation of minority women. A striking example was a large study of 43,568 women to examine genetic variants associated with sPTB.[45] In this study, six single-nucleotide polymorphisms (SNPs) were associated with sPTB that replicated in a combined dataset of three Nordic cohorts ($n = 8643$). However, data were obtained from 23andMe®, a company using customers' DNA to provide individuals with genomic analysis of ancestry and other genetic traits for a fee, and restricted to only participants with >97% European genetic ancestry. Restriction to these women could have been because of under representation of other ancestral groups using 23andMe®, the availability to replicate with a European cohort, or because of a desire to analyze a more homogenous population for the purpose of isolating effects of SNPs, but reasons are not stated. Nonetheless, data restricted to one ancestral group cannot be used to determine the extent to which disparities in PTB are due to genetic variation between ancestral groups (or racial groups that do track, though incompletely, with self-identified race).

In studies with more diversity in ancestral groups, often there remains residual under representation of participants of non-European descent. For example, in a study of 791 family trios with 270 PTBs, there were just 14 (5%) PTBs among women of African Ancestry. This is too few to determine whether ancestral genetic variation contributes to population differences in PTB risk.[46] In another large GWAS of sPTB, 1349 cases were ancestry-matched to full-term controls (based on genomic fetal signals).[47] The investigators then performed stratified analyses within ancestral groups to identify genes within groups associated with sPTB. They found only two SNPs in intergenic regions, one within the ancestral group from Africa and one within the ancestral group from the Americas that were associated with sPTB. However, this approach cannot determine the extent to which ancestral genetic patterns are responsible for racial differences in sPTB. Notably, these SNPs failed to replicate in several independent cohorts[47] and were not reported in prior GWAS analyses of PTB.[45,46] In summary, the contribution of genetics to PTB remains an open

question and is even less well established as an explanation for racial disparities.

Epigenomic approaches

In various disease states, epigenomic signals may result from exposures that change epigenomic biomarkers. Importantly, it is thought that epigenomic changes are potentially more modifiable than genetic sequences. However, it can be a mistake to assume that these epigenomic changes are on the causal pathway to disease. We believe that one epigenetic mechanism, DNA methylation, provides an illustrative cautionary tale. It is an example of how epigenetic assays may not only help understand physiology but also to mistake DNA methylation as in the causal pathway linking exposure to health outcome. Most human epigenetic studies are observational, but one blinded, randomized cross-over study of air pollution (an exposure associated with PTB[48]) evaluated blood pressure and DNA methylation in peripheral blood.[49] Fifteen healthy participants participated in the trial in which they were exposed to odorless concentrated ambient particles (akin to levels in highly polluted cities) or medical air. Exposure to the concentrated particles resulted in significantly lower DNA methylation of repetitive element DNA (*Alu*) as well as the *TLR4* gene. Although not an epigenome-wide study, short-term exposure to pollution caused measurable changes in DNA methylation and air pollution caused higher systolic blood pressure. The investigators also found an association (not necessarily causal) between lower DNA methylation and higher blood pressure. Hence, this elegant hypothesis-generating study demonstrated causality of air pollution and DNA methylation changes (as well as air pollution and changes in blood pressure). Yet, it does not determine whether the observed DNA methylation changes are causally responsible for differences in blood pressure. In other words, DNA methylation might be the biologic process that connects air pollution to elevated blood pressure; or alternatively, there could be other, unmeasured mechanisms along the causal pathway that confound the relationship between DNA methylation and blood pressure (Fig. 2). In the latter case, DNA methylation would be a marker of air pollution but not a therapeutic target to lower blood pressure. One mechanistic hypothesis is that air pollution could cause differences in leukocyte demargination from the bone marrow and different leukocyte populations in the peripheral blood, which by definition have differential DNA methylation. If the mechanism were activation of adrenergic pathways or hormonal, this could lead to higher blood pressure independent of the changes in DNA methylation. This illustrates that even in well-controlled human studies of exposure, biomarker, and outcome, it can be difficult to demonstrate that biomarkers are on the causal pathway. This is particularly relevant with respect to sPTB where peripherally circulating biomarkers may or may not be relevant to reproductive organs. Nonetheless, even if biomarkers are not on the causal pathway, they can still serve as markers of exposures or risk of outcomes, but they may not be useful targets of therapeutic agents.

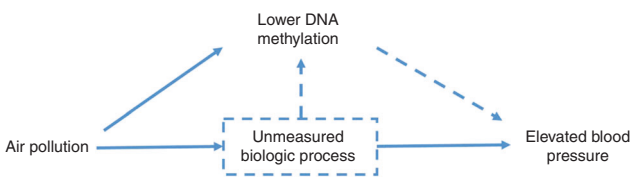

Fig. 2 Example of how an exposure (air pollution) might both cause a change to a biomarker (DNA methylation) and an outcome (blood pressure) yet not be on the causal pathway. Here an unmeasured biologic process acts as a mediator of the pollution–blood pressure relationship and as a confounder in the methylation–blood pressure relationship. Solid lines established, dashed lines hypothesized

When considering whether DNA methylation differences represent biologic pathways by which racial disparities occur, there are additional challenges to consider. Consider for a moment a theoretical molecular epidemiologic study of race, DNA methylation, and PTB. One might find in this study that DNA methylation differs by race. If similar to national data,[4] this theoretical study would likely find black women have extreme PTB (<28 weeks of gestation) rates that are three times as high as white women. Assume lastly, DNA methylation is found to be associated with PTB. One might conclude that differential DNA methylation is causally related to racial disparities in PTB, but such a conclusion would be premature. There are potential alternative explanations. The first is that DNA methylation is not causally related to PTB but is simply a passive marker of SNP frequencies at potential methylation sites (CpG sites in which cytosines are followed by guanines in the DNA sequence) that vary by continental ancestry, which may be somewhat correlated with race (genetic confounding), but that some other mechanism is contributing to differences in PTB (akin to Fig. 2 but substituting black race for air pollution and PTB for blood pressure). The second alternative explanation is that a diffrerent co-exposure, an unmeasured confounder (such as air pollution exposure, which is higher, on average, among black families[33,34,50]) leads to differential DNA methylation that is causally involved in the pathophysiology of PTB (classical confounding). Neither self-identified race nor the ancestral DNA patterns that are more common among black women would have been the causative etiology of differences in DNA methylation. Race was perhaps confounded by socioeconomic position that led to a complex set of exposures including air pollution, which was the true cause of DNA methylation differences. In other words, doing expensive omic assays does not avoid the traps of traditional epidemiologic studies such as confounding.

Concern about confounding in molecular epidemiologic studies of PTB is not new. The most striking example of a biomarker that is strongly associated with sPTB but is not a useful therapeutic target to reduce sPTB is a Nugent score[51] positive for bacterial vaginosis (BV). Trials of antibiotics therapy are only successful at treating BV but do not reduce the risk of sPTB,[49] suggesting that BV may not be causally related to sPTB. Making the leap that since BV is more common among black women[52] and that BV is associated with sPTB,[53] that BV plays a causal role in racial disparities sPTB, may be misguided given the lack of efficacy of treating BV in preventing sPTB. Simply treating BV will not reduce racial disparities in sPTB.

Potential solutions to omic studies of racial disparities in PTB

Whether a single biomarker for a condition like BV or an omic signature, such as the epigenome or microbiome, is measured, to determine the extent to which a biomarker might explain a race–PTB association, causal mediation analyses are needed.[54] While rarely employed in birth outcome studies, studies of aging populations have demonstrated that DNA methylation may mediate the smoking–lung function relationship.[55] Causal mediation methods can quantify the direct and indirect (through a proposed mediator) effects between an exposure and an outcome. Such approaches will be necessary to determine the extent to which omic assay results that differ by race explain disparities in PTB risk. In order to perform formal causal mediation analyses, sufficient sample sizes with variation in the distribution of the biomarker and in the incidence of the outcome across participants of all racial groups is required. With respect to the recent cervicovaginal microbiome study demonstrating associations of the microbiome with both race and sPTB, the study design included frequency matching of sPTB cases to term controls by race, eliminating the racial disparity in sPTB in the analytic dataset.[17] This approach was important for the primary goal of analyzing the main effect of the microbiome and sPTB but precludes a mediation analysis of the microbiome as a mediator

between race and sPTB. Race-stratified analyses in which a biomarker is associated with PTB among African Americans allows for the possibility that the mediator is involved in racial disparities. However, both causal mediation and race-stratified analyses do not completely avoid the pitfalls of observational studies; confounding can still occur. Interventional animal studies and eventually human intervention trials would be required to demonstrate that a mediator were causally involved in sPTB before concluding that differences by race are involved in racial disparities in sPTB. Such trials should target modifying the omic signatures, and even more likely combinations of omic signatures that may work in concert, and should have as a primary outcome sPTB rates. Lastly, while there are high rates of sPTB among African Americans, the vast majority of women deliver full-term, healthy infants. Identifying omic signatures that mediate resilience in spite of experiences of discrimination or physical toxic environmental exposures may help to identify therapeutic interventions to bolster biologic resilience.

CONCLUSION

Omic assays common in "precision medicine" and "precision public health" have the potential to risk-stratify individuals for PTB and to elucidate mechanisms by which exposures cause PTB. If biomarkers do represent pathways by which PTB occurs, they may be appropriate therapeutic targets to prevent PTB. Such biomedical interventions do not solve social and environmental injustices that may be the underlying causes of disparities but are likely simpler. Studies that attempt to explain racial disparities in PTB with omic assays need to have adequate representation and sufficient sampling of women across racial groups. However, observational omic studies are affected by all of the same challenges (confounding, bias) that affect observational epidemiologic studies more generally. In order for biomarkers from omic studies to be useful at the *individual* level, they must perform well as predictors (as opposed to simply being statistically significantly associated with exposures or outcome) for clinical intervention. In order for omic assays to be helpful in reducing racial disparities in PTB at the *population* level, they either need to be good biomarkers of exposures that are causal and track by race (thus motivating policies to reduce these exposures) or they need to be specifically modifiable among black women so as to reduce the excess risk of PTB among black women. Lastly, omic assays, when their results differ by race, should not be assumed to be innate biologic, genetic differences between races. While they may be subtly affected by SNPs that differ by ancestry, differences in omic signatures may also be due to differential exposures to toxic exposures by race due to longstanding inequity in the United States. Determining why they differ and whether there are ways to mitigate their effects represents a new frontier in tackling racial disparities in birth outcomes.

AUTHOR CONTRIBUTIONS

H.H.B. conceived of and wrote the manuscript. Each co-author contributed ideas, edited, and approved the final manuscript.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Wong, M. D., Shapiro, M. F., Boscardin, W. J. & Ettner, S. L. Contribution of major diseases to disparities in mortality. *N. Engl. J. Med.* **347**, 1585–1592 (2002).
2. MacDorman, M. F. & Mathews, T. J. Understanding racial and ethnic disparities in U.S. infant mortality rates. *NCHS Data Brief* 1–8 (2011).
3. Matthews, T. J., MacDorman, M. F. & Thoma, M. E. Infant mortality statistics from the 2013 period linked birth/infant death data set. *Natl Vital Stat. Rep.* **64**, 1–30 (2015).
4. Martin, J. A., Hamilton, B. E., Osterman, M. J. K., Driscoll, A. K. & Drake, P. Births: final data for 2016. *Natl Vital Stat. Rep.* **67**, 1–55 (2018).
5. Xu, J., Murphy, S. L., Kochanek, K. D., Bastian, B. & Arias, E. Deaths: final data for 2016. *Natl Vital Stat. Rep.* **67**, 1–76 (2018).
6. Anderson, J. G., Rogers, E. E. & Baer, R. J. et al. Racial and ethnic disparities in preterm infant mortality and severe morbidity: a population-based study. *Neonatology* **113**, 44–54 (2018).
7. Manrai, A. K., Ioannidis J. P. A. & Patel, C. J. Signals among signals: prioritizing non-genetic associations in massive datasets. *Am. J. Epidemiol.* **188**, 846–850 (2019).
8. Yudell, M., Roberts, D., DeSalle, R. & Tishkoff, S. Science and society. Taking race out of human genetics. *Science* **351**, 564–565 (2016).
9. Williams, D. R. Race and health: basic questions, emerging directions. *Ann. Epidemiol.* **7**, 322–333 (1997).
10. McKinnon, B., Yang, S., Kramer, M. S., Bushnik, T., Sheppard, A. J. & Kaufman, J. S. Comparison of black-white disparities in preterm birth between Canada and the United States. *CMAJ* **188**, E19–26 (2016).
11. David, R. J. & Collins, J. W. Jr. Differing birth weight among infants of U.S.-born blacks, African-born blacks, and U.S.-born whites. *N. Engl. J. Med.* **337**, 1209–1214 (1997).
12. Elo, I. T., Vang, Z. & Culhane, J. F. Variation in birth outcomes by mother's country of birth among non-Hispanic black women in the United States. *Matern. Child Health J.* **18**, 2371–2381 (2014).
13. Mittal, P., Romero, R. & Tarca, A. L. et al. Characterization of the myometrial transcriptome and biological pathways of spontaneous human labor at term. *J. Perinat. Med.* **38**, 617–643 (2010).
14. Bukowski, R., Sadovsky, Y. & Goodarzi, H. et al. Onset of human preterm and term birth is related to unique inflammatory transcriptome profiles at the maternal fetal interface. *PeerJ* **5**, e3685 (2017).
15. D'Silva, A. M., Hyett, J. A. & Coorssen, J. R. Proteomic analysis of first trimester maternal serum to identify candidate biomarkers potentially predictive of spontaneous preterm birth. *J. Proteomics* **178**, 31–42 (2018).
16. Hong, X., Sherwood, B. & Ladd-Acosta, C. et al. Genome-wide DNA methylation associations with spontaneous preterm birth in US blacks: findings in maternal and cord blood samples. *Epigenetics* **13**, 163–172 (2018).
17. Elovitz, M. A., Gajer, P. & Riis, V. et al. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. *Nat. Commun.* **10**, 1305 (2019).
18. Menon, R., Jones, J. & Gunst, P. R. et al. Amniotic fluid metabolomic analysis in spontaneous preterm birth. *Reprod. Sci.* **21**, 791–803 (2014).
19. Ghartey, J., Bastek, J. A., Brown, A. G., Anglim, L. & Elovitz, M. A. Women with preterm birth have a distinct cervicovaginal metabolome. *Am. J. Obstet. Gynecol.* **212**, 776 e1–e12 (2015).
20. Ghartey, J., Anglim, L., Romero, J., Brown, A. & Elovitz, M. A. Women with symptomatic preterm birth have a distinct cervicovaginal metabolome. *Am. J. Perinatol.* **34**, 1078–1083 (2017).
21. Ghaemi, M. S., DiGiulio, D. B. & Contrepois, K. et al. Multiomics modeling of the immunome, transcriptome, microbiome, proteome and metabolome adaptations during human pregnancy. *Bioinformatics* **35**, 95–103 (2019).
22. Class, Q. A., Rickert, M. E. & Oberg, A. S. et al. Within-family analysis of interpregnancy interval and adverse birth outcomes. *Obstet. Gynecol.* **130**, 1304–1311 (2017).
23. Nelson, M. C., Gordon-Larsen, P. & Adair, L. S. Are adolescents who were breast-fed less likely to be overweight? Analyses of sibling pairs to reduce confounding. *Epidemiology* **16**, 247–253 (2005).
24. Khoshnood, B., Lee, K. S., Wall, S., Hsieh, H. L. & Mittendorf, R. Short interpregnancy intervals and the risk of adverse birth outcomes among five racial/ethnic groups in the United States. *Am. J. Epidemiol.* **148**, 798–805 (1998).
25. Hogue, C. J., Menon, R., Dunlop, A. L. & Kramer, M. R. Racial disparities in preterm birth rates and short inter-pregnancy interval: an overview. *Acta Obstet. Gynecol. Scand.* **90**, 1317–1324 (2011).
26. Burris, H. H. et al. Racial disparities in preterm birth in USA: a biosensor of physical and social environmental exposures. *Arch. Dis. Child.* https://doi.org/10.1136/archdischild-2018-316486 (2019).
27. Clougherty, J. E., Levy, J. I. & Kubzansky, L. D. et al. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. *Environ. Health Perspect.* **115**, 1140–1146 (2007).
28. Padula, A. M., Huang, H. & Baer, R. J. et al. Environmental pollution and social factors as contributors to preterm birth in Fresno County. *Environ. Health* **17**, 70 (2018).

29. Cohen, S., Tyrrell, D. A. & Smith, A. P. Negative life events, perceived stress, negative affect, and susceptibility to the common cold. *J. Pers. Soc. Psychol.* **64**, 131–140 (1993).
30. Rosenberg, L., Palmer, J. R., Wise, L. A., Horton, N. J. & Corwin, M. J. Perceptions of racial discrimination and the risk of preterm birth. *Epidemiology* **13**, 646–652 (2002).
31. Copper, R. L., Goldenberg, R. L. & Das, A. et al. The preterm prediction study: maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am. J. Obstet. Gynecol.* **175**, 1286–1292 (1996).
32. Collins, J. W. Jr., David, R. J., Handler, A., Wall, S. & Andes, S. Very low birthweight in African American infants: the role of maternal exposure to interpersonal racial discrimination. *Am. J. Public Health* **94**, 2132–2138 (2004).
33. Pratt, G. C., Vadali, M. L., Kvale, D. L. & Ellickson, K. M. Traffic, air pollution, minority and socio-economic status: addressing inequities in exposure and risk. *Int. J. Environ. Res. Public Health* **12**, 5355–5372 (2015).
34. Mohai, P., Lantz, P. M., Morenoff, J., House, J. S. & Mero, R. P. Racial and socio-economic disparities in residential proximity to polluting industrial facilities: evidence from the Americans' Changing Lives Study. *Am. J. Public Health* **99** (Suppl 3), S649–56 (2009).
35. Mohai, P. & Saha, R. Racial inequality in the distribution of hazardous waste: a national-level reassessment. *Soc. Probl.* **54**, 343–370 (2007).
36. Brender, J. D., Maantay, J. A. & Chakraborty, J. Residential proximity to environmental hazards and adverse health outcomes. *Am. J. Public Health* **101**(Suppl 1), S37–52 (2011).
37. Casey, J. A., Karasek, D. & Ogburn, E. L. et al. Retirements of coal and oil power plants in california: association with reduced preterm birth among populations nearby. *Am. J. Epidemiol.* **187**, 1586–1594 (2018).
38. Alfirevic, Z., Stampalija, T. & Medley, N. Cervical stitch (cerclage) for preventing preterm birth in singleton pregnancy. *Cochrane Database Syst. Rev.* **6**, CD008991 (2017).
39. Dodd, J. M., Jones, L., Flenady, V., Cincotta, R. & Crowther, C. A. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. *Cochrane Database Syst. Rev.* CD004947 (2013).
40. Meis, P. J., Klebanoff, M. & Thom, E. et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. *N. Engl. J. Med.* **348**, 2379–2385 (2003).
41. Owen, J., Hankins, G. & Iams, J. D. et al. Multicenter randomized trial of cerclage for preterm birth prevention in high-risk women with shortened midtrimester cervical length. *Am. J. Obstet. Gynecol.* **201**, e1–8 (2009).
42. Hassan, S. S., Romero, R. & Vidyadhari, D. et al. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. *Ultrasound Obstet. Gynecol.* **38**, 18–31 (2011).
43. Fonseca, E. B., Celik, E., Parra, M., Singh, M. & Nicolaides, K. H. Fetal Medicine Foundation Second Trimester Screening Group. Progesterone and the risk of preterm birth among women with a short cervix. *N. Engl. J. Med.* **357**, 462–469 (2007).
44. Ngo, T. T. M., Moufarrej, M. N. & Rasmussen, M. H. et al. Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* **360**, 1133–1136 (2018).
45. Zhang, G., Feenstra, B. & Bacelis, J. et al. Genetic associations with gestational duration and spontaneous preterm birth. *N. Engl. J. Med.* **377**, 1156–1167 (2017).
46. Knijnenburg, T. A., Vockley, J. G. & Chambwe, N. et al. Genomic and molecular characterization of preterm birth. *Proc. Natl Acad. Sci.* **116**, 5819–5827 (2019).
47. Rappoport, N., Toung, J. & Hadley, D. et al. A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth. *Sci. Rep.* **8**, 226 (2018).
48. Stieb, D. M., Chen, L., Eshoul, M. & Judek, S. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ. Res.* **117**, 100–111 (2012).
49. Bellavia, A., Urch, B. & Speck, M. et al. DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled human exposure experiments. *J. Am. Heart Assoc.* **2**, e000212 (2013).
50. Nachman, K. E. & Parker, J. D. Exposures to fine particulate air pollution and respiratory outcomes in adults using two national datasets: a cross-sectional study. *Environ. Health* **11**, 25 (2012).
51. Nugent, R. P., Krohn, M. A. & Hillier, S. L. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J. Clin. Microbiol.* **29**, 297–301 (1991).
52. Allsworth, J. E. & Peipert, J. F. Prevalence of bacterial vaginosis: 2001-2004 National Health and Nutrition Examination Survey data. *Obstet. Gynecol.* **109**, 114–120 (2007).
53. Hillier, S. L., Nugent, R. P. & Eschenbach, D. A. et al. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. *N. Engl. J. Med.* **333**, 1737–1742 (1995).
54. VanderWeele, T. J. Mediation analysis: a practitioner's guide. *Annu. Rev. Public Health* **37**, 17–32 (2016).
55. Barfield, R., Shen, J. & Just, A. C. et al. Testing for the indirect effect under the null for genome-wide mediation analyses. *Genet. Epidemiol.* **41**, 824–833 (2017).