Complex methods for complex data: key considerations for interpretable and actionable results in exposome research: accompanying material

Marta Ponzano, Ran Rotem, Andrea Bellavia

Contents:

1. **Integrating social constructs in exposome research: illustrative simulation study**
   - ***A.*** *Overview of the simulation study.*
   - ***B.*** *Simulation details.*
   - ***C.*** *Quantification of the health disparity: stratification for the non-modifiable social construct.*
   - ***D.*** *Incorporation of the social construct: benefits of including proximal factors.*
   - ***E.*** *Simulation code.*

Integrating social constructs: illustrative simulation study

In this section we present an illustrative example on simulated data to show how to handle non-modifiable social constructs. Specifically, this example will consider the variable "race", which is a social construct due to complex historical and societal factors rooted in systemic racism.[1–6] We will explore the impact of environmental racism on health disparities, specifically of racial environmental health disparities in Monoethyl phthalate (MEP) on birth weight. First, we unpacked what is meant by race and racism, clarifying the diagram of interest. Second, a dataset was simulated based on this diagram using previously published papers. Third, we quantified the environmental health disparity of interest. Finally, we incorporated information on race in assessing the overall exposure-outcome association with different sets of adjustment. Simulation details as well as Stata code is available.

PART 1.A.

Overview of the simulation study: the impact of Environmental Racism on Low Birth Weight

As an example, we can consider the role of environmental racism on the 2-fold increased risk of delivering a low birth weight (BW) infant among non-Hispanic black individuals compared to non-Hispanic white individuals.[7,8] Numerous studies have evaluated a host of social and environmental factors for this striking disparity. Among these are the consideration of environmental chemicals, such as phthalates. In the U.S., diethyl phthalate is one of the most commonly used phthalate parent compounds, which is mostly used in personal care products. Monoethyl phthalate (MEP) is the biologically-active metabolite of diethyl phthalate.[9–11] Research has documented the impact of Eurocentric beauty standards, contributing to racial disparities in targeted marketing and overabundance of availability phthalate-containing personal care products in communities of color. Studies have shown the following associations: a) higher concentrations of certain phthalates and lower birth weight; b) higher use of certain personal care products (i.e. hair oils and perfumes) with higher concentrations of MEP; c) higher use of hair oils with lower birth weight.[12–21] Recognizing the importance of both individual and community drivers of phthalate-associated personal care product use and chemical exposures, a conceptual diagram should denote not only individual usage, but also targeted marketing and neighborhood accessibility of safer products. Figure 1 shows a proposed conceptual diagram for the racial environmental health disparities under study.

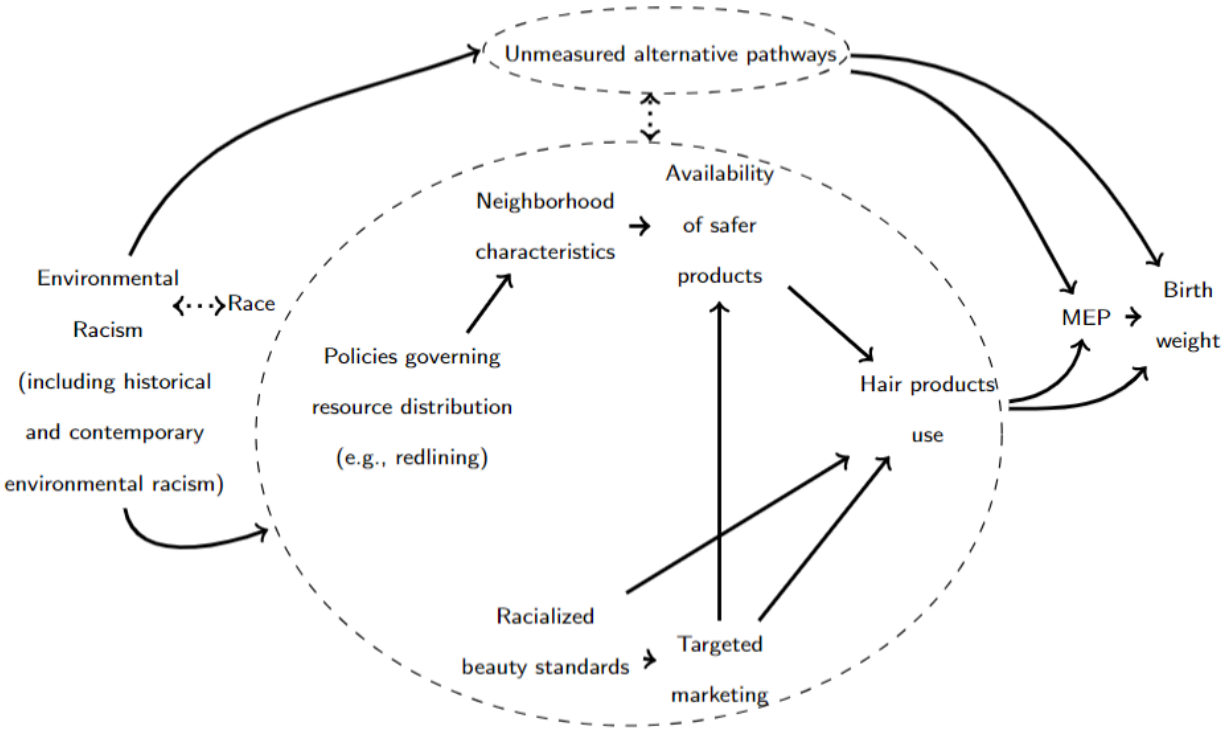

Figure 1. C*onceptual diagram for racial environmental health disparities in Monoethyl phthalate (MEP), birth weight, and the contribution of racism-induced differences in hair products usage. Based on study-*

specific characteristics, some factors will be unavailable or unknown (unmeasured factors). Dotted lines signify associations and not causal effects.

Conceptualizing and unpacking the meaning of "race" in this context allows a more defined and thoughtful identification of the analytical steps required for addressing the question of the potential contribution of phthalate metabolite concentrations in the low-birth-weight disparity. In the next section we detail these analytical steps in simulated data.

PART 1.B.

Simulation details

We simulated a gradient of seven settings ranging from a situation where most of the environmental health disparity is due to the measured factors that contribute to the environmental exposure disparity through racism (Scenario 1), to a setting where the disparity is largely due to unmeasured factors (Scenario 7).

We simulated a sample of 10,000 subjects in order to reproduce the example of MEP (environmental exposure) and birth weight (health outcome) based on associations from the literature[12,19,21,22,23] . Race was simulated based on a Bernoulli variable with parameter 0.19. The use of hair products containing endocrine disrupting chemicals (EDCs) was a Bernoulli variable with parameters of 0.494 and 0.077 respectively for subjects with race B and A. The other unmeasured factors involved in the association under study were quantified as a normal continuous variable with standard deviation 3 and with means $\mu 1$ and $\mu 0$ of 4 and 0 respectively for subjects with race B and A. MEP urinary concentration was simulated as a normal variable with standard deviation 12 and with mean defined as $\beta 0+\beta 1$(use of hair products containing EDCs)+ $\beta 2$(other unmeasured factors)+ $\beta 3$(interaction between hair products and other unmeasured factors), with $\beta 0$, $\beta 1$, $\beta 2$ and $\beta 3$ respectively set to 61, 20, 0.1 and 0.5. Birth weight was simulated as a normal variable with standard deviation 450 and with mean defined as $\alpha 0+ \alpha 1(2^{nd}$ quartile of MEP) $+\alpha 2(3^{rd}$ quartile of MEP) $+\alpha 3(4^{th}$ quartile of MEP) $+\alpha 4$(use of hair products containing EDCs) $+\alpha 5$(interaction between hair products use and MEP) + $\alpha 6$(other unmeasured factors)+ $\alpha 7$(interaction between hair products use and other unmeasured factors)+ $\alpha 8$(interaction between other unmeasured factors and MEP), with $\alpha 0$, $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 4$, $\alpha 5$, $\alpha 6$, $\alpha 7$, $\alpha 8$ respectively set to 3490, -34.6, -200.2, -72.1, -54, -2, -2, -0.5, -0.005. This scenario was replicated 1000 times and average estimates were calculated. Six other additional scenarios were simulated by making the following modifications: Scenario 2): $\mu 1$ set to 6; Scenario 3): $\alpha 6$ set to -5; Scenario 4): $\beta 2$ set to 1.5; Scenario 5): $\alpha 6$ set to -5 and $\beta 2$ set to 1.5; Scenario 6): $\alpha 6$ set to -5, $\beta 2$ set to 1.5 and $\beta 3$ set to 1; Scenario 7): $\alpha 6$ set to -5, $\beta 2$ set to 1.5, $\beta 3$ set to 1, $\alpha 7$ set to -2 and $\alpha 8$ set to -0.05.

PART 1.C.

Quantification of the health disparity: stratification for the non-modifiable social construct

Taking into account the need for solution-oriented research in environmental health disparities, the following core steps can be taken to study, describe, and estimate the magnitude of an environmental health disparity: 1) investigate differences in the racial distribution of the environmental exposure; 2) investigate differences in the racial distribution of the health outcome; 3) evaluate the effects of the environmental exposure on the health outcome over subpopulations of interest. Figure 2 displays the results of these steps in our simulated data considering Scenario 1. First, the historically marginalized group (race B) has a higher distribution of the environmental factor (MEP concentration). Second, the historically marginalized group has a greater proportion of individuals with the adverse health outcome (lower BW). Third, the association between MEP on BW is associated with a 23% greater reduction in the outcome among individuals in race B compared to race A (-6.12 g versus -4.97 g). Table 1 summarizes the findings in all the 7 simulated scenarios.

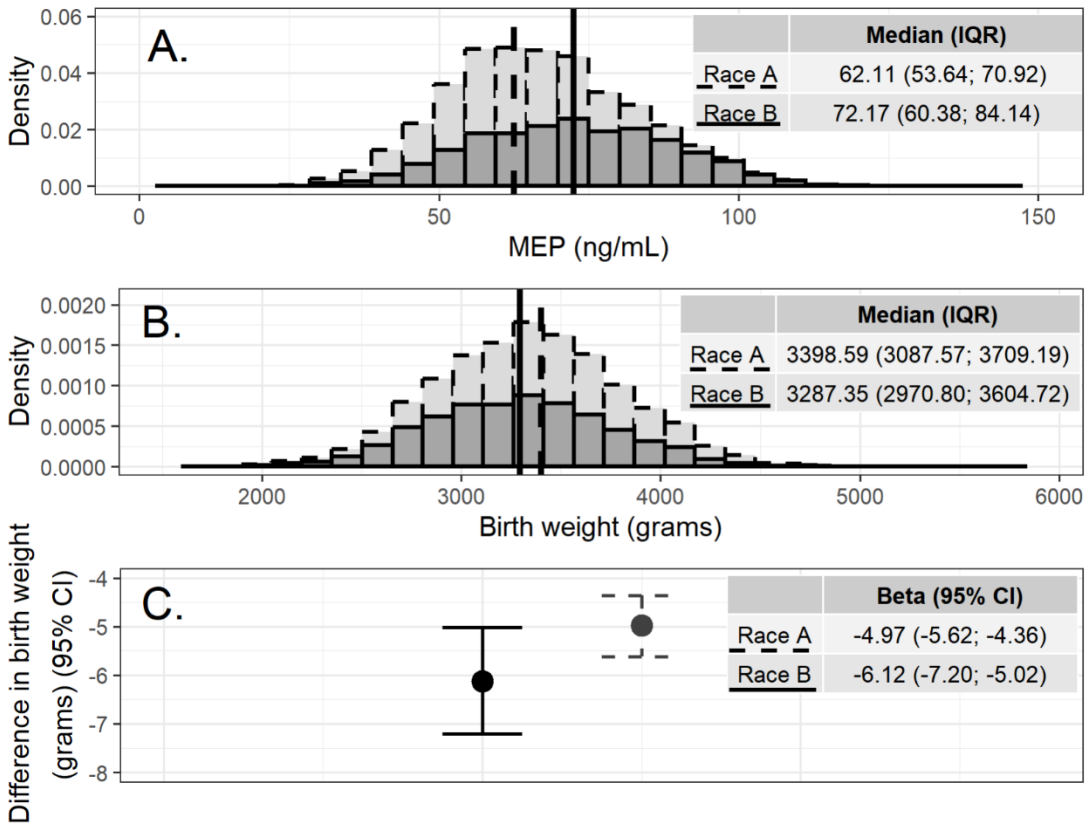

Figure 2. *Analytical steps to describe and quantify an environmental health disparity in a simulated dataset: quantify difference in exposure's distribution (A); quantify difference in the outcome distributions (B); and evaluate strata-specific effects of the exposure on the outcome (C).*

Table 1. MEP (ng/mL) distribution, Birth weight (grams) distribution and strata-specific effects of MEP on the BW under the 7 different simulation scenarios.

Scenario 1		
MEP distribution, *Median (IQR)*	Overall	63.52 (54.55; 73.33)
	Race A	62.11 (53.64; 70.92)
	Race B	72.17 (60.38; 84.14)
BW distribution, *Median (IQR)*	Overall	3377.92 (3064.24; 3691.00)
	Race A	3398.59 (3087.57; 3709.19)
	Race B	3287.35 (2970.80; 3604.72)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.70 (-6.22; -5.17)
	Race A	-4.97 (-5.62; -4.36)
	Race B	-6.12 (-7.20; -5.02)
Scenario 2		
MEP distribution, *Median (IQR)*	Overall	63.57 (54.58; 73.47)
	Race A	62.11 (53.64; 70.92)
	Race B	72.85 (60.71; 85.20)
BW distribution, *Median (IQR)*	Overall	3376.73 (3062.80; 3689.98)
	Race A	3398.53 (3087.48; 3709.13)
	Race B	3281.29 (2964.57; 3598.63)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.77 (-6.28; -5.23)
	Race A	-4.99 (-5.64; -4.39)
	Race B	-6.11 (-7.17; -5.05)
Scenario 3		
MEP distribution, Median (IQR)	Overall	63.52 (54.55; 73.33)
	Race A	62.11 (53.64; 70.92)
	Race B	72.17 (60.38; 84.14)
BW distribution, Median (IQR)	Overall	3375.72 (3061.47; 3689.19)
	Race A	3398.61 (3087.44; 3709.37)
	Race B	3275.42 (2958.64; 3592.92)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.81 (-6.33; -5.28)
	Race A	-4.99 (-5.64; -4.38)
	Race B	-6.15 (-7.22; -5.05)
Scenario 4		
MEP distribution, Median (IQR)	Overall	64.32 (54.65; 75.01)
	Race A	62.14 (53.12; 71.46)
	Race B	77.58 (65.56; 90.05)
BW distribution, Median (IQR)	Overall	3376.84 (3062.64; 3690.24)
	Race A	3399.97 (3088.60; 3710.89)
	Race B	3275.70 (2960.06; 3592.01)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.48 (-5.92; -4.98)
	Race A	-4.93 (-5.52; -4.36)
	Race B	-5.37 (-6.38; -4.35)
Scenario 5		
MEP distribution,	Overall	64.32 (54.65; 75.01)

Median (IQR)	Race A	62.14 (53.12; 71.46)
	Race B	77.58 (65.56; 90.05)
BW distribution, Median (IQR)	Overall	3374.57 (3059.70; 3688.69)
	Race A	3399.97 (3088.26; 3711.29)
	Race B	3263.62 (2947.79; 3580.27)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.76 (-6.21; -5.27)
	Race A	-5.14 (-5.73; -4.57)
	Race B	-5.53 (-6.54; -4.52)
Scenario 6		
MEP distribution, Median (IQR)	Overall	64.31 (54.64; 75.08)
	Race A	62.11 (53.10; 71.42)
	Race B	78.21 (65.67; 91.77)
BW distribution, Median (IQR)	Overall	3374.21 (3059.23; 3688.40)
	Race A	3399.81 (3088.09; 3711.14)
	Race B	3262.34 (2946.20; 3579.09)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-5.72 (-6.16; -5.24)
	Race A	-5.14 (-5.72; -4.58)
	Race B	-5.48 (-6.44; -4.51)
Scenario 7		
MEP distribution, Median (IQR)	Overall	64.31 (54.64; 75.08)
	Race A	62.11 (53.10; 71.42)
	Race B	78.21 (65.67; 91.77)
BW distribution, Median (IQR)	Overall	3370.47 (3054.25; 3685.84)
	Race A	3399.20 (3087.07; 3710.96)
	Race B	3244.36 (2926.56; 3562.80)
Estimates of the MEP - BW associations, β (95% CI)	Overall	-6.15 (-6.59; -5.68)
	Race A	-5.36 (-5.94; -4.80)
	Race B	-6.03 (-7.02; -5.06)

PART 1.D.

Incorporation of the social construct: benefits of including proximal factors.

Another common way the variable "race" is used in environmental epidemiology is as an adjustment covariate, often specified as a confounder, for assessing the independent association between the environmental exposure on the health outcome in the overall population. Specifically, a confounder is a variable that is associated with the exposure and the outcome and is not a part of the causal pathway. As such, including "race" as an adjustment covariate does not imply unrealistic deterministic assumptions for its effect on either exposure or outcome. Adjusting for "race" can therefore improve precision as it blocks alternative pathways that can lead from exposure to outcome. Nevertheless, from the conceptual diagram in Figure 1, "race" is associated with both exposure and outcome only through the factors that are the real determinants of the disparity. As such, if these factors were available, including these covariates as confounders as an explanation of "race", would provide similar results with improved interpretation. Of note, we are not stating to replace race, but rather to understand its meaning analytically by unpacking it and including these variables in the model to better understand the hypothesized meaning.

In Figure 3, we show how the estimate of the exposure-outcome association in the simulated data changes with different sets of adjustment, as compared to the real effect (adjusting for both measured and unmeasured factors). We present results for two potential scenarios: one where the environmental health disparity is explained by mostly measured factors (Scenario 1) and one where it is explained by both measured and unmeasured factors (Scenario 7). In both cases, adjusting for race alone severely underestimates the total effect – meaning that both the estimate and interpretation of the associations are impacted by the ability to specify the drivers of environmental racial inequities for the specific research question. Adjusting for measured factors known to be driving race and racism's effect on the environmental exposure disparity is key even when unmeasured factors considerably contribute to the disparity. In other words, thoughtful consideration and appropriate specification of what race means and how racism is operating might substantially reduce bias through reducing measurement error in this critically important variable. Further inclusion of race in the adjustment can provide additional, albeit slight, improvement to the precision of the estimate. It is therefore acceptable to incorporate race as an adjustment covariate in regression models as long as it is interpreted as a confounder (that is, clearly specifying that it is not involved in the causal pathways of interest) and it is expected to increase estimates precision. These results were consistent when including a potential interaction between race and measures or unmeasured factors. Additionally, results from all the intermediate scenarios are reported in Table 2.

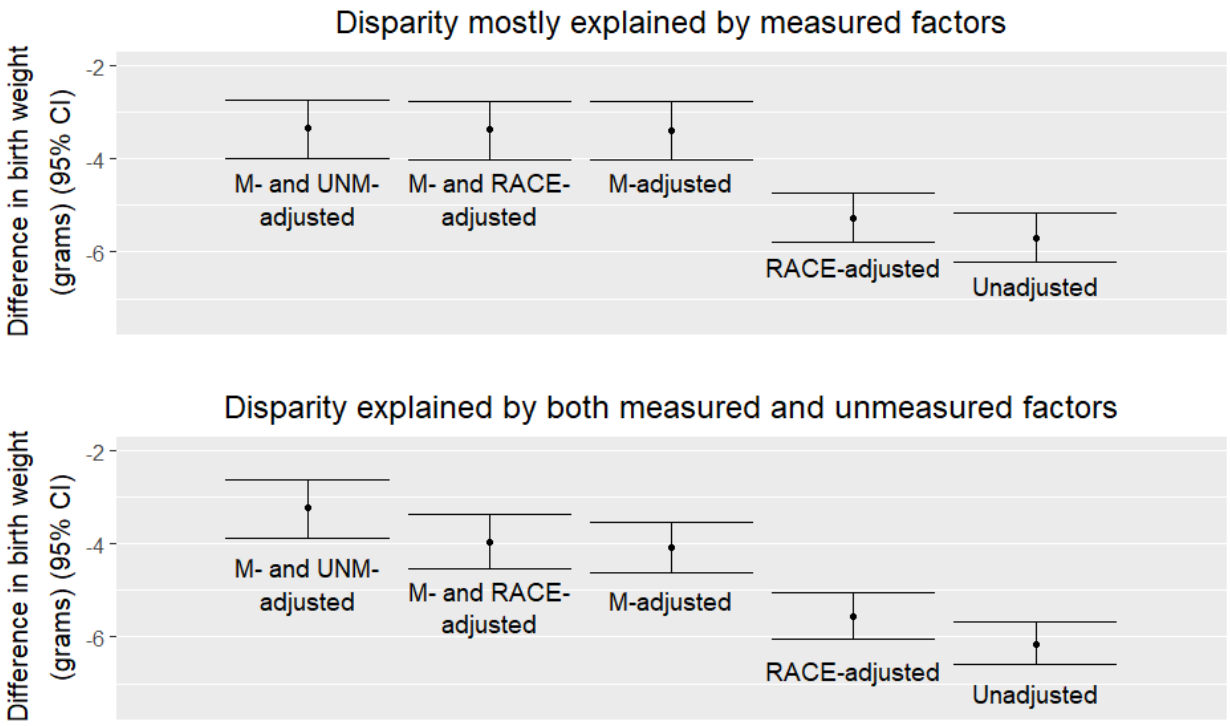

Figure 3. *Estimates of the overall effect of MEP on BW under several adjustment sets with and without the inclusion of the variable race. M-Measured factors; UNM-Unmeasured factors.*

Table 2. Overall estimates of the MEP (ng/mL) -BW (grams) associations (β (95% CI)) under the 7 different simulation scenarios. M-Measured factors; UNM-Unmeasured factors

Scenario 1	
M- and UNM- adjusted	-3.35 (-3.98; -2.75)
Unadjusted	-5.70 (-6.22; -5.17)
M-adjusted	-3.38 (-4.03; -2.77)
M- and RACE-adjusted	-3.37 (-4.02; -2.77)
RACE-adjusted	-5.28 (-5.80; -4.73)
Scenario 2	
M- and UNM- adjusted	-3.36 (-4.01; -2.77)
Unadjusted	-5.77 (-6.28; -5.23)
M-adjusted	-3.40 (-4.04; -2.81)
M- and RACE-adjusted	-3.39 (-4.03; -2.79)
RACE-adjusted	-5.30 (-5.83; -4.76)
Scenario 3	
M- and UNM- adjusted	-3.35 (-3.98; -2.75)
Unadjusted	-5.81 (-6.33; -5.28)
M-adjusted	-3.42 (-4.08; -2.82)
M- and RACE-adjusted	-3.40 (-4.06; -2.81)
RACE-adjusted	-5.30 (-5.83; -4.76)
Scenario 4	
M- and UNM- adjusted	-3.15 (-3.80; -2.55)

Unadjusted	-5.48 (-5.92; -4.98)
M-adjusted	-3.36 (-3.92; -2.79)
M- and RACE-adjusted	-3.34 (-3.93; -2.75)
RACE-adjusted	-5.04 (-5.55; -4.52)
Scenario 5	
M- and UNM- adjusted	-3.15 (-3.80; -2.55)
Unadjusted	-5.76 (-6.21; -5.27)
M-adjusted	-3.66 (-4.23; -3.10)
M- and RACE-adjusted	-3.60 (-4.19; -3.01)
RACE-adjusted	-5.24 (-5.75; -4.72)
Scenario 6	
M- and UNM- adjusted	-3.16 (-3.80; -2.56)
Unadjusted	-5.72 (-6.16; -5.24)
M-adjusted	-3.67 (-4.23; -3.11)
M- and RACE-adjusted	-3.61 (-4.20; -3.02)
RACE-adjusted	-5.23 (-5.73; -4.73)
Scenario 7	
M- and UNM- adjusted	-3.22 (-3.87; -2.63)
Unadjusted	-6.15 (-6.59; -5.68)
M-adjusted	-4.08 (-4.63; -3.53)
M- and RACE-adjusted	-3.95 (-4.54; -3.37)
RACE-adjusted	-5.55 (-6.04; -5.04)

PART 1.E.

Simulation STATA code.

```
clear all
scalar n_iter=1000
scalar  BW_MEP_unadjusted_1000     =0
scalar  BW_MEP_products_other_1000 =0
scalar  BW_MEP_race_1000           =0
scalar  BW_MEP_products_1000       =0
scalar  BW_MEP_products_race_1000  =0
matrix define V_BW_MEP_unadjusted_1000       =J(1000, 1, 0)
matrix define V_BW_MEP_products_other_1000   =J(1000, 1, 0)
matrix define V_BW_MEP_race_1000             =J(1000, 1, 0)
matrix define V_BW_MEP_products_1000         =J(1000, 1, 0)
matrix define V_BW_MEP_products_race_1000    =J(1000, 1, 0)

scalar  BW_MEP_race1_1000          =0
scalar  BW_MEP_race0_1000          =0
matrix define V_BW_MEP_race1_1000   =J(1000, 1, 0)
matrix define V_BW_MEP_race0_1000   =J(1000, 1, 0)

scalar  _MEP_median_overall_1000            =0
scalar  _MEP_quartile1_overall_1000         =0
scalar  _MEP_quartile3_overall_1000         =0
scalar  _BW_median_overall_1000             =0
scalar  _BW_quartile1_overall_1000          =0
scalar  _BW_quartile3_overall_1000          =0
scalar  _MEP_median_race1_1000              =0
scalar  _MEP_quartile1_race1_1000           =0
scalar  _MEP_quartile3_race1_1000           =0
scalar  _BW_median_race1_1000               =0
scalar  _BW_quartile1_race1_1000            =0
scalar  _BW_quartile3_race1_1000            =0

scalar  _MEP_median_race0_1000             =0
```

```
scalar  _MEP_quartile1_race0_1000          =0
scalar  _MEP_quartile3_race0_1000          =0
scalar  _BW_median_race0_1000              =0
scalar  _BW_quartile1_race0_1000           =0
scalar  _BW_quartile3_race0_1000           =0

forvalues iter = 1/1000 {

      set more off
      set obs 10000
      set seed   `iter'

      gen race =.
      replace race=rbinomial(1,0.19)

      gen products=.
      replace products=rbinomial(1,0.494) if race==1
      replace products=rbinomial(1,0.077) if race==0

      gen other=.
      replace other=rnormal(4, 3) if  race==1
      replace other=rnormal(0, 3) if  race==0
      gen int_products_other=products*other

      scalar  beta_MEP               =61
      scalar  beta_products_MEP      =20
      scalar  beta_other_MEP         =0.1
      scalar  beta_other_products_MEP =0.5

      gen MEP=rnormal(beta_MEP+ ///
                      beta_products_MEP*products+   ///
                              beta_other_MEP*other+ ///
                              beta_other_products_MEP*int_products_other, 12)

      gen int_other_MEP   =other*MEP
```

```
gen int_products_MEP=products*MEP

xtile MEP_QUART = MEP, nq(4)
tab MEP_QUART, generate(MEP_QUART)

scalar beta_BW                =3490
scalar beta_MEP_QUART2_BW     =-34.6
scalar beta_MEP_QUART3_BW     =-200.2
scalar beta_MEP_QUART4_BW     =-72.1
scalar beta_products_BW       =-54
scalar beta_products_MEP_BW   =-2
scalar beta_other_BW          =-2
scalar beta_other_products_BW =-0.5
scalar beta_other_MEP_BW      =-0.005

gen BW=rnormal(beta_BW+ ///
                      beta_MEP_QUART2_BW*MEP_QUART2+ ///
                      beta_MEP_QUART3_BW*MEP_QUART3+ ///
                      beta_MEP_QUART4_BW*MEP_QUART4+ ///
                      beta_products_BW*products+ ///
                      beta_products_MEP_BW*int_products_MEP + ///
                      beta_other_BW*other + ///
                      beta_other_products_BW*int_products_other + ///
                      beta_other_MEP_BW*int_other_MEP   ///
                      , 450)

  regress BW MEP
  matrix b=e(b)
  matrix V_BW_MEP_unadjusted_1000[`iter',1]=(b[1,1])
  scalar  BW_MEP_unadjusted_1000=BW_MEP_unadjusted_1000+b[1,1]

  regress BW MEP products  other
  matrix b=e(b)
  matrix V_BW_MEP_products_other_1000[`iter',1]=(b[1,1])
  scalar  BW_MEP_products_other_1000=BW_MEP_products_other_1000+b[1,1]
```

```
regress BW MEP race
 matrix b=e(b)
 matrix V_BW_MEP_race_1000[`iter',1]=(b[1,1])
 scalar  BW_MEP_race_1000=BW_MEP_race_1000+b[1,1]

 regress BW MEP products
 matrix b=e(b)
 matrix V_BW_MEP_products_1000[`iter',1]=(b[1,1])
 scalar  BW_MEP_products_1000=BW_MEP_products_1000+b[1,1]

 regress BW MEP products race
 matrix b=e(b)
 matrix V_BW_MEP_products_race_1000[`iter',1]=(b[1,1])
 scalar  BW_MEP_products_race_1000=BW_MEP_products_race_1000+b[1,1]

 regress BW MEP if race==0
 matrix b=e(b)
 matrix V_BW_MEP_race0_1000[`iter',1]=(b[1,1])
 scalar  BW_MEP_race0_1000=BW_MEP_race0_1000+b[1,1]

 regress BW MEP if race==1
 matrix b=e(b)
 matrix V_BW_MEP_race1_1000[`iter',1]=(b[1,1])
 scalar  BW_MEP_race1_1000=BW_MEP_race1_1000+b[1,1]

 sum   MEP     , det
 scalar  _MEP_median_overall_1000=_MEP_median_overall_1000+r(p50)
 scalar  _MEP_quartile1_overall_1000=_MEP_quartile1_overall_1000+r(p25)
 scalar  _MEP_quartile3_overall_1000=_MEP_quartile3_overall_1000+r(p75)

 sum   BW     , det
 scalar  _BW_median_overall_1000=_BW_median_overall_1000+r(p50)
 scalar  _BW_quartile1_overall_1000=_BW_quartile1_overall_1000+r(p25)
 scalar  _BW_quartile3_overall_1000=_BW_quartile3_overall_1000+r(p75)

 sum   MEP if race==0 , det
```

```
        scalar  _MEP_median_race0_1000=_MEP_median_race0_1000+r(p50)
        scalar  _MEP_quartile1_race0_1000=_MEP_quartile1_race0_1000+r(p25)
        scalar  _MEP_quartile3_race0_1000=_MEP_quartile3_race0_1000+r(p75)

        sum   BW if race==0  , det
        scalar  _BW_median_race0_1000=_BW_median_race0_1000+r(p50)
        scalar  _BW_quartile1_race0_1000=_BW_quartile1_race0_1000+r(p25)
        scalar  _BW_quartile3_race0_1000=_BW_quartile3_race0_1000+r(p75)

        sum   MEP if race==1 , det
        scalar  _MEP_median_race1_1000=_MEP_median_race1_1000+r(p50)
        scalar  _MEP_quartile1_race1_1000=_MEP_quartile1_race1_1000+r(p25)
        scalar  _MEP_quartile3_race1_1000=_MEP_quartile3_race1_1000+r(p75)

        sum   BW if race==1 , det
        scalar  _BW_median_race1_1000=_BW_median_race1_1000+r(p50)
        scalar  _BW_quartile1_race1_1000=_BW_quartile1_race1_1000+r(p25)
        scalar  _BW_quartile3_race1_1000=_BW_quartile3_race1_1000+r(p75)

        drop race products other int* MEP* BW
       }

di BW_MEP_unadjusted_1000/n_iter
set obs 10000
matvsort V_BW_MEP_unadjusted_1000 V_BW_MEP_unadjusted_1000
di V_BW_MEP_unadjusted_1000[50, 1]
di V_BW_MEP_unadjusted_1000[950, 1]

di BW_MEP_products_other_1000/n_iter
set obs 10000
matvsort V_BW_MEP_products_other_1000 V_BW_MEP_products_other_1000
di V_BW_MEP_products_other_1000[50, 1]
di V_BW_MEP_products_other_1000[950, 1]

di BW_MEP_race_1000/n_iter
```

```
set obs 10000
matvsort V_BW_MEP_race_1000 V_BW_MEP_race_1000
di V_BW_MEP_race_1000[50, 1]
di V_BW_MEP_race_1000[950, 1]

di BW_MEP_products_1000/n_iter
set obs 10000
matvsort V_BW_MEP_products_1000 V_BW_MEP_products_1000
di V_BW_MEP_products_1000[50, 1]
di V_BW_MEP_products_1000[950, 1]

di BW_MEP_products_race_1000/n_iter
set obs 10000
matvsort V_BW_MEP_products_race_1000 V_BW_MEP_products_race_1000
di V_BW_MEP_products_race_1000[50,1]
di V_BW_MEP_products_race_1000[950,1]

di BW_MEP_race0_1000/n_iter
set obs 10000
matvsort V_BW_MEP_race0_1000 V_BW_MEP_race0_1000
di V_BW_MEP_race0_1000[50,1]
di V_BW_MEP_race0_1000[950,1]

di BW_MEP_race1_1000/n_iter
set obs 10000
matvsort V_BW_MEP_race1_1000 V_BW_MEP_race1_1000
di V_BW_MEP_race1_1000[50,1]
di V_BW_MEP_race1_1000[950,1]

di _MEP_median_overall_1000/n_iter
di _MEP_quartile1_overall_1000/n_iter
di _MEP_quartile3_overall_1000/n_iter

di _BW_median_overall_1000/n_iter
di _BW_quartile1_overall_1000/n_iter
```

```
di _BW_quartile3_overall_1000/n_iter

di _MEP_median_race0_1000/n_iter
di _MEP_quartile1_race0_1000/n_iter
di _MEP_quartile3_race0_1000/n_iter

di _BW_median_race0_1000/n_iter
di _BW_quartile1_race0_1000/n_iter
di _BW_quartile3_race0_1000/n_iter

di _MEP_median_race1_1000/n_iter
di _MEP_quartile1_race1_1000/n_iter
di _MEP_quartile3_race1_1000/n_iter

di _BW_median_race1_1000/n_iter
di _BW_quartile1_race1_1000/n_iter
di _BW_quartile3_race1_1000/n_iter

/*For scenarios 2-7, please consider the following modifications:
2. : replace other=rnormal(6, 3) if  race==1
3. : scalar  beta_other_BW=-5
4. : scalar  beta_other_MEP=1.5
5. : scalar  beta_other_BW=-5
     scalar  beta_other_MEP=1.5
6. : scalar  beta_other_BW=-5
     scalar  beta_other_MEP=1.5
     scalar  beta_other_products_MEP=1
7. : scalar  beta_other_BW=-5
     scalar  beta_other_MEP=1.5
     scalar  beta_other_products_MEP=1
     scalar  beta_other_products_BW=-2
     scalar  beta_other_MEP_BW=-0.05
*/
```

References

1. Boyd, R. W., Lindo, E. G., Weeks, L. D. & McLemore, M. R. On Racism: A New Standard For Publishing On Racial Health Inequities. *Health Aff. Forefr.* doi:10.1377/forefront.20200630.939347.
2. Gee, G. C. & Ford, C. L. STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions. *Bois Rev. Soc. Sci. Res. Race* **8**, 115–132 (2011).
3. Bailey, Z. D., Feldman, J. M. & Bassett, M. T. How Structural Racism Works — Racist Policies as a Root Cause of U.S. Racial Health Inequities. *N. Engl. J. Med.* **384**, 768–773 (2021).
4. Bailey, Z. D. *et al.* Structural racism and health inequities in the USA: evidence and interventions. *The Lancet* **389**, 1453–1463 (2017).
5. Ford, C. L. & Airhihenbuwa, C. O. The public health critical race methodology: Praxis for antiracism research. *Soc. Sci. Med.* **71**, 1390–1398 (2010).
6. Williams, D. R., Lawrence, J. A. & Davis, B. A. Racism and Health: Evidence and Needed Research. *Annu. Rev. Public Health* **40**, 105–125 (2019).
7. Ratnasiri, A. W. G. *et al.* Recent trends, risk factors, and disparities in low birth weight in California, 2005–2014: a retrospective study. *Matern. Health Neonatol. Perinatol.* **4**, 15 (2018).
8. Martin, J. A., Hamilton, B. E., Osterman, M. J. K. & Driscoll, A. K. Births: Final Data for 2018. *Natl. Vital Stat. Rep. Cent. Dis. Control Prev. Natl. Cent. Health Stat. Natl. Vital Stat. Syst.* **68**, 1–47 (2019).
9. Zota, A. R. & Shamasunder, B. The environmental injustice of beauty: framing chemical exposures from beauty products as a health disparities concern. *Am. J. Obstet. Gynecol.* **217**, 418.e1-418.e6 (2017).
10. Chan, M., Mita, C., Bellavia, A., Parker, M. & James-Todd, T. Racial/Ethnic Disparities in Pregnancy and Prenatal Exposure to Endocrine-Disrupting Chemicals Commonly Used in Personal Care Products. *Curr. Environ. Health Rep.* **8**, 98–112 (2021).
11. James-Todd, T. M., Chiu, Y.-H. & Zota, A. R. Racial/Ethnic Disparities in Environmental Endocrine Disrupting Chemicals and Women's Reproductive Health Outcomes: Epidemiological Examples Across the Life Course. *Curr. Epidemiol. Rep.* **3**, 161–180 (2016).

12. Rylander, L. & Källén, B. Reproductive outcomes among hairdressers. *Scand. J. Work. Environ. Health* **31**, 212–217 (2005).

13. Helm, J. S., Nishioka, M., Brody, J. G., Rudel, R. A. & Dodson, R. E. Measurement of endocrine disrupting and asthma-associated chemicals in hair products used by Black women. *Environ. Res.* **165**, 448–458 (2018).

14. Hauser, R. PHTHALATES AND HUMAN HEALTH. *Occup. Environ. Med.* **62**, 806–818 (2005).

15. Preston, E. V. *et al.* Socioeconomic and racial/ethnic differences in use of endocrine-disrupting chemical-associated personal care product categories among pregnant women. *Environ. Res.* **198**, 111212 (2021).

16. Preston, E. V. *et al.* Endocrine disrupting chemical-associated hair product use during pregnancy and gestational age at delivery: a pilot study. *Environ. Health* **20**, 1–9 (2021).

17. Chiu, Y.-H. *et al.* Evaluating effects of prenatal exposure to phthalate mixtures on birth weight: a comparison of three statistical approaches. *Environ. Int.* **113**, 231–239 (2018).

18. Koniecki, D., Wang, R., Moody, R. P. & Zhu, J. Phthalates in cosmetic and personal care products: Concentrations and possible dermal exposure. *Environ. Res.* **111**, 329–336 (2011).

19. James-Todd, T., Senie, R. & Terry, M. B. Racial/Ethnic Differences in Hormonally-Active Hair Product Use: A Plausible Risk Factor for Health Disparities. *J. Immigr. Minor. Health* **14**, 506–511 (2012).

20. Smarr, M. M. *et al.* Parental urinary biomarkers of preconception exposure to bisphenol A and phthalates in relation to birth outcomes. *Environ. Health* **14**, 73 (2015).

21. Fruh, V. *et al.* Urinary phthalate metabolite concentrations and personal care product use during pregnancy – Results of a pilot study. *Sci. Total Environ.* **835**, 155439 (2022).

22. U.S. Census Bureau QuickFacts: United States. https://www.census.gov/quickfacts/fact/table/US/LFE046221#qf-headnote-a].

23. Smarr, M. M. *et al.* Parental urinary biomarkers of preconception exposure to bisphenol A and phthalates in relation to birth outcomes. *Environ. Health* **14**, 73 (2015).