Title: Long-term air pollution exposure and markers of cardiometabolic health in the National Longitudinal Study of Adolescent to Adult Health (Add Health) Study

Authors: Mercedes A. Bravo,¹ Fang Fang,² Dana B. Hancock,² Eric O. Johnson,²,³ Kathleen Mullan Harris⁴,⁵

Author affiliations: ¹ Global Health Institute, School of Medicine, Duke University, Durham, NC, USA; ² GenOmics, Bioinformatics, and Translational Research Center, RTI International, Research Triangle Park, NC, USA; ³ Fellow Program, RTI International, Research Triangle Park, NC, USA; ⁴ Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; ⁵ Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

ABSTRACT

Background: Air pollution exposure is associated with cardiovascular morbidity and mortality. Although exposure to air pollution early in life may represent a critical window for development of cardiovascular disease risk factors, few studies have examined associations of long-term air pollution exposure with markers of cardiovascular and cardiometabolic health in young adults.

Objectives: By combining health data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) with air pollution data from the Fused Air Quality Surface using Downscaling (FAQSD) archive, we: (1) calculated multi-year estimates of exposure to ozone (O₃) and particulate matter with an aerodynamic diameter ≤2.5µm (PM₂.₅) for Add Health participants; and (2) estimated associations between air pollution exposures and multiple markers of cardiometabolic health.

Methods: Add Health is a nationally representative longitudinal cohort study of over 20,000 adolescents aged 12–19 in the United States (US) in 1994–95 (Wave I). Participants have been followed through adolescence and into adulthood with five in-home interviews. Estimated daily concentrations of O₃ and PM₂.₅ at census tracts were obtained from the FAQSD archive and used to generate tract-level annual averages of O₃ and PM₂.₅ concentrations. We estimated associations between average O₃ and PM₂.₅ exposures from 2002–07 and markers of cardiometabolic health measured at Wave IV (2008–09), including hypertension, hyperlipidemia, body mass index (BMI), diabetes, C-reactive protein, and metabolic syndrome.

Results: The final sample size was 11,259 individual participants. The average age of participants at Wave IV was 28.4 years (range: 24–34 years). In models adjusting for age, race/ethnicity, and sex, long-term O₃ exposure (2002–07) was associated with elevated odds of hypertension, with an odds ratio (OR) of 1.015 (95% confidence interval [CI]: 1.011, 1.029); obesity (1.022 [1.004, 1.040]); diabetes (1.032 [1.009,1.054]); and metabolic syndrome (1.028 [1.014, 1.041]); PM₂.₅ exposure (2002–07) was associated with elevated odds of hypertension (1.022 [1.001, 1.045]).

Conclusion: Findings suggest that long-term ambient air pollution exposure, particularly O₃ exposure, is associated with cardiometabolic health in early adulthood.

Key words: air pollution, cardiometabolic health, young adult health, longitudinal, multi-year, long-term] air pollution exposure, early life exposure

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
INTRODUCTION
An extensive literature has demonstrated that exposure to ambient air pollution has wide-ranging harmful effects on health.\(^1\) Exposures to ozone (O\(_3\)) and particulate matter with an aerodynamic diameter ≤2.5\(\mu\)m (PM\(_{2.5}\)) are linked to cardiovascular and respiratory-related mortality,\(^2\) morbidity,\(^5,8\) and hospital admissions,\(^9,10\) even at concentrations below the National Ambient Air Quality Standards.\(^11,12\) A growing body of research has investigated the relationship between air pollution exposure, particularly traffic-related air pollution exposure, and risk factors for early markers of cardiovascular disease. Evidence increasingly suggests that exposure to air pollutants is associated with risk factors for early indicators of cardiovascular disease that may develop years or decades prior to clinical manifestations of more severe disease.\(^13\)

Specifically, chronic exposure to air pollution early in the life course may directly affect development of major risk factors for cardiovascular disease, including obesity, hypertension, and metabolic disorders. Studies have documented associations between residential proximity to roads and arterial stiffness,\(^14\) an early-stage indicator of atherosclerosis, and the underlying cause of most cardiovascular disease. There is also evidence to suggest that long-term exposure to air pollution, including PM\(_{2.5}\) and O\(_3\), is associated with accelerated atherosclerosis,\(^15\) elevated blood pressure and increased mean arterial pressure,\(^16\) obesity,\(^17,18\) and diabetes.\(^9,20\) Much of the work on air pollution and cardiovascular disease risk has focused on middle-aged and older adults, e.g., ≥45 years, the demographic age group at highest risk for morbidity and mortality associated with cardiovascular disease.

However, exposure to air pollution during early developmental stages (childhood, adolescence, and early adulthood) may represent critical windows for development of risk factors for cardiovascular disease. For example, a 2021 study of children in Los Angeles, CA, observed that children with higher exposure to traffic-related air pollution had greater annual change in carotid artery intima-media thickness, a measure indicative of subclinical atherosclerosis.\(^21\) Jerrett et al. examined traffic pollution around children’s residences and observed that higher levels of traffic-related air pollution were associated with higher body mass index (BMI) in children 10-18 years, particularly in females.\(^22\) Li et al. studied adults aged 18 to 29 and found that high concentrations of some air pollutants were associated with impairments in high-density lipoprotein (HDL) functionality.\(^23\) and McGuinn et al. found that long-term PM\(_{2.5}\) exposure was associated with lipoprotein increases in adults.\(^24\) Studies of air pollution exposure and cardiometabolic outcomes in children, adolescents, or young adults are few, often have shorter follow up periods, and tend to be based on small or non-representative samples in localized geographic areas.\(^23,25,26\)

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a nationally representative cohort of 20,745 individuals in the United States (US). Add Health participants were enrolled as adolescents aged 12-19 years in 1994–95 and have been followed through adolescence and into adulthood until, most recently, 2016–18.\(^27\) Add Health and its collection of individual-level measures of cardiometabolic health offer an important and unique opportunity to further our understanding of potential adverse health effects of air pollution exposures that occur in early adulthood. Markers of cardiometabolic health can provide insight into underlying biological processes in premorbid disease pathways and may indicate biological dysfunction before clinical manifestation of more severe disease.

In this study, we used publicly available daily estimates of ambient concentrations of two criteria air pollutants, PM\(_{2.5}\) and O\(_3\), to generate annual average exposure estimates, and merged these to Add Health participants based on census tract(s) of residence across the multiple waves of Add Health. We then: (1) evaluated whether multi-year averages of air pollutant exposure prior to Wave IV, which occurred in 2008–09, differed by individual-level characteristics (e.g., race/ethnicity, sex); and (2) estimated associations between air pollution exposure and six markers of cardiometabolic health measured at Wave IV: hypertension, hyperlipidemia, obesity, diabetes, C-reactive protein, and a summary measure of metabolic syndrome. This is the first study of long-term air pollution exposure in Add Health and, to the
best of our knowledge, one of a limited number of studies examining air pollution exposures and cardiometabolic health in a cohort of young adults.26

**METHODS**

**Add Health cohort**

We used data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), an ongoing, nationally representative cohort of adolescents in grades 7–12 during the 1994–95 school year. Add Health used a multistage, stratified, school-based, cluster sampling design to select a probability sample of more than 20,000 adolescents from school rosters for in-home interviews at Wave I (WI) in 1994–95 (ages 12–19). The cohort has been followed up with four subsequent interviews: WII in 1996 (ages 13–20); WIII in 2001–02 (ages 18–26); WIV in 2008–09 (ages 24–32); and WV in 2016–18 (ages 33–43).27,28 Response rates have ranged from 72% to 90% across waves, and non-response bias has been minimal.29,30 Details on the Add Health study objectives, design, and sampling strategy are provided elsewhere.27,28

For the exposure measures, the present paper uses Add Health participant data from Waves I, III, and IV with recently constructed and merged air pollution data, as described later. For the cardiometabolic health markers, we used the biological and clinical data collected at Wave IV (response rate=80.3%), including systolic and diastolic blood pressure; lipid panels; measured height and weight used to calculate BMI; indicators of diabetes (e.g., fasting or non-fasting glucose and/or glycated hemoglobin [HbA1c]); and concentrations of C-reactive protein as a measure of inflammation.31 We examined six key markers of cardiometabolic health at Wave IV: hypertension, hyperlipidemia, obesity, diabetes, and C-reactive protein concentration, and a summary measure of metabolic syndrome. Following clinical guidelines to identify the high-risk category, all cardiometabolic health markers were coded as binary variables (0/1) and defined as follows.

Hypertension was indicated if participants met at least one of the following criteria: having stage 1 or stage 2 hypertension according to the guidelines from the Seventh Report of the Joint National Committee on Prevention Detection, Evaluation, and Treatment of High Blood Pressure (JNC7): systolic blood pressure of 140–159 mm Hg or diastolic blood pressure of 80–89 mm Hg for stage 1 hypertension; systolic blood pressure of ≥160 mm Hg or diastolic blood pressure of ≥100 mm Hg for stage 2 hypertension;32 or self-reported history of having physician-diagnosed high blood pressure; or self-reported anti-hypertensive medication use.

Hyperlipidemia was indicated if participants had a self-reported history of physician-diagnosed high cholesterol or triglycerides; or self-reported anti-hyperlipidemic medication use in the past 4 weeks.

Participants’ BMI was calculated based on height and weight measured at Wave IV as kg/m². Respondents were considered obese if BMI was ≥30.

Diabetes was indicated if participants met at least one of the following criteria: either fasting glucose ≥126 mg/dL or non-fasting glucose ≥200 mg/dL or glycated hemoglobin (HbA1c) ≥6.5%; or self-reported history of having physician-diagnosed diabetes except during pregnancy; or self-reported anti-diabetic medication use.

Inflammation was measured by high-sensitivity C-reactive protein (hsCRP) and coded as a binary indicator of high inflammation if hsCRP was > 3mg/L. C-reactive protein is a widely measured biomarker of systemic inflammation, a well-established risk factor for the development and progression of cardiometabolic diseases.33,34
Metabolic syndrome represents a cluster of conditions that increases the risk of heart disease, stroke, and diabetes. Metabolic syndrome was indicated if participants had three or more of the following conditions: (1) blood pressure > 130/85 mm Hg, or self-reported physician-diagnosed hypertension, or self-reported anti-hypertensive medication usage; (2) HbA1c ≥ 5.7%; (3) membership in the lowest two high-density lipoprotein (HDL) cholesterol deciles for women or the lowest three HDL deciles for men; (4) membership in the top three deciles of triglycerides; and (5) waist circumference ≥ 88 cm for women or ≥ 102 cm for men. This definition of metabolic syndrome has been used in previous studies of Add Health participants.

Analytic Samples/Inclusion Criteria
We created an analytical dataset from the Add Health study sample at Wave I (n=20,745). We included only Wave I individuals who participated in Wave III and Wave IV; were geocoded; resided in the continental US (individuals residing in Alaska or Hawaii were excluded because air pollution data was not available for these states); and had non-missing data on key covariates, i.e., age, race/ethnicity, and sex (Figure 1). We also included only US-born participants, due to health differences by nativity status, and women who were not pregnant at Wave IV, as anthropometric and physiological markers differ for pregnant vs. non-pregnant women.

With these inclusion criteria, a maximum of 11,259 participants remained for statistical analyses. Sample sizes varied slightly across analyses depending on the health outcome, due to missing health measurements for some markers of cardiometabolic health (Figure 1).

Air pollution exposure
Air pollution exposure estimates were generated using the publicly available Fused Air Quality Surface using Downscaling (FAQSD) files [https://www.epa.gov/hesc/rsig-related-downloadable-data-files]. The FAQSD files include daily predictions of 24-hour average PM$_{2.5}$ concentrations and 8-hour maximum O$_3$ concentrations at 2010 US Census tract centroids. The FAQSD files were generated using a statistical model ("downscaler") that relates monitoring data and gridded output from the Community Multiscale Air Quality (CMAQ) model using a linear regression model with additive and multiplicative bias coefficients that can vary in space and time (see Supplemental Information for additional details).

FAQSD files included predictions of daily 24-hour average PM$_{2.5}$ concentrations and 8-hour maximum O$_3$ concentrations at census tracts for 2002–07. We averaged daily values to generate annual average concentration estimates of tract-level 8-hour maximum O$_3$ and 24-hour average PM$_{2.5}$ for each year. Add Health participants were then matched to air pollution values based on their census tract(s) of residence and year. Residential location information on participants was available at each wave, but not for between-wave years, while air pollution data were obtained for 2002–07 (Figure 2). As illustrated in Figure 2, annual average PM$_{2.5}$ and O$_3$ concentrations from 2002–05 were attached to Add Health participants based on their census tract of residence at Wave III (2001–02) and annual average PM$_{2.5}$ and O$_3$ concentrations from 2006–09 were attached to Add Health participants based on their census tract of residence at Wave IV (2008–09). Because we did not have information on when individuals moved in between-wave years, for those who changed residences we assumed they moved close to the “midpoint” of years between waves. Thus, we assumed that Add Health participants resided at their Wave III address between 2002–05, and at their Wave IV address between 2006–09.

We examined air pollution exposure and the key markers of cardiometabolic health at Wave IV in the analytical sample of 11,259 individuals. We used FAQSD output to estimate average PM$_{2.5}$ and O$_3$ exposures from 2002–07, which represents the long-term average air pollution exposure prior to collection of cardiometabolic health measures at Wave IV in 2008–09. Because FAQSD data are not available prior to 2002, we are unable to estimate air pollution exposure prior to 2002 using this particular data source.
Written informed consent was obtained at all waves and protocols approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

**Statistical analysis**

We calculated descriptive statistics of the study population, as well as average air pollution exposures by individual-level characteristics (e.g., age, sex, race/ethnicity). We tested for differences in 2002–07 O₃ and PM₂.₅ concentrations by sex and racial/ethnic group, applying post-hoc tests for multiple pairwise comparisons and adjusting p-values for multiple comparisons using Bonferroni. To examine associations between air pollution exposures and markers of cardiometabolic health among young adults, we fit logistic regression generalized estimating equations (GEEs) to estimate the association between each air pollutant (O₃, PM₂.₅) and health outcome. Models adjusted for Add Health participant’s age (years), sex, and race/ethnicity (non-Hispanic Black, Non-Hispanic White, Hispanic, Other):

\[
\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{air pollutant}_{i, 2002-07} + \beta_2 \text{sex}_i + \beta_3 \text{race/ethnicity}_i + \beta_4 \text{age}_i
\]

where \( \pi_i \) is the probability of experiencing the outcome of interest for subject \( i \). The coefficient \( \beta_i \) represents the estimated association between the health outcome and air pollution exposure averaged over 2002–07; the coefficients \( \beta_2, \beta_3, \) and \( \beta_4 \) represent the associations of sex, race/ethnicity, and Wave IV age, respectively, with the outcome of interest. Models were fit separately for each pollutant and health outcome. An exchangeable correlation matrix was specified on the primary sampling unit ID in Add Health. We fit unweighted multivariable models with covariates adjusting for unequal selection into the sample and accounting for the clustered school design using GEEs.

A two-sided \( \alpha=0.05 \) was used to signify statistical significance. Statistical analyses were conducted using R statistical software. The `gee` package was used for fitting GEE models.

**Sensitivity analysis**

As a sensitivity analysis, we fit adjusted models that estimated associations between air pollution exposure in 2006–07 and markers of cardiometabolic health measured at Wave IV. Analysis of 2002–07 air pollution exposure requires that individuals participated in both Waves III and IV, because their residential address at each wave is required to assign air pollution exposure. In contrast, analysis of 2006–07 air pollution exposure only requires that individuals participated in Wave IV, because only a Wave IV address is used to assign air pollution exposure estimates for 2006 and 2007 (e.g., Figure 2). The 2006–07 exposure period represents lagged exposure in the years just prior to Wave IV. We conducted this sensitivity analysis to explore whether results were consistent: (1) across exposure averaging intervals, e.g., 2002–07 and 2006–07; and (2) using the slightly larger sample that is available for the 2006–07 analysis. We used the same set of restriction criteria previously applied to the sample of Wave III and IV participants, resulting in an analytical sample size of \( n=13,867 \) individuals for the sensitivity analysis (SM, Figure S1).

**RESULTS**

**Descriptive statistics**

Descriptive statistics of the study sample are provided in Table 1. At Wave IV, Add Health participants were, on average, 28 years old, and just over half the sample was female. Non-Hispanic White participants comprised over half of the sample (~66%). The most common cardiometabolic health outcomes were high inflammation and obesity, followed by hypertension and metabolic syndrome.

**Air pollution exposure**

Air pollution exposure estimates are summarized in Table 2 overall, and by sex and race/ethnicity. Average O₃ and PM₂.₅ exposures were 38.3 ppb and 12.7 \( \mu \text{g/m}^3 \), respectively, over the 2002–07 period. T-
tests with unequal variances were used to test for differences in O₃ and PM₂.₅ concentrations by sex and the Kruskal-Wallis test was used to evaluate differences in 2002–07 air pollutant concentrations by racial/ethnic group.⁵¹ Air pollutant concentrations did not differ by sex, but there were differences between in PM₂.₅ and O₃ by racial/ethnic group (p<0.0001).

Non-Hispanic Black and non-Hispanic White groups had the highest O₃ exposure (38.6 ppb); Other racial/ethnic groups (e.g., Asian, Native American, “other” race) had the lowest O₃ exposure (36.1 ppb). For O₃ exposure, Dunn’s test indicated that the non-Hispanic White group had different O₃ exposures compared to Hispanic (p < 0.0001) and Other groups (p < 0.0001). The non-Hispanic Black group also had different O₃ exposures compared to Hispanic (p < 0.0001) and Other groups (p < 0.0001). With respect to PM₂.₅, the non-Hispanic Black group had the highest PM₂.₅ exposure (13.6 µg/m³), while the non-Hispanic White group had the lowest PM₂.₅ exposure (12.2 µg/m³). Dunn’s tests indicated that the non-Hispanic Black group had different PM₂.₅ exposures compared to non-Hispanic White (p < 0.0001) and Hispanic groups (p < 0.0001). The non-Hispanic White group had different PM₂.₅ exposures compared to Hispanic (p < 0.0001) and Other groups (p < 0.0001). PM₂.₅ exposure in the Hispanic group differed from that in the Other group as well (p=0.00112).

**Associations of air pollution exposure and health outcomes**

Odds ratios (ORs) for associations of each air pollutant (O₃ and PM₂.₅) with each health outcome are presented in Table 3. In GEEs adjusting for age, race/ethnicity, and sex, 2002–07 O₃ exposure was associated with elevated odds of hypertension, with an OR of 1.015 (95% confidence interval [CI]: 1.011, 1.034); obesity (1.022 [1.004, 1.040]); diabetes (1.032 [1.009,1.054]); and metabolic syndrome (1.028 [1.014, 1.041]) (Table 3). In models adjusting for age, race/ethnicity, and sex, 2002–07 PM₂.₅ exposure was associated with elevated odds of hypertension (1.022 [1.001, 1.045]).

**Sensitivity analysis**

Summary statistics of the analytic sample examined for 2006–07 exposure are provided in SM Table S1. Exposure to O₃ in 2006–07 was associated with elevated odds of obesity (1.014 [1.003, 1.025]); diabetes (1.024 [1.004, 1.045]); inflammation (1.012 [1.003, 1.022]); and metabolic syndrome (1.013 [1.001, 1.026]) (SM Table S2). Statistically significant associations were not observed between 2006–07 PM₂.₅ exposure and health outcomes, although there was a marginal association of PM₂.₅ with hypertension (p=0.0518).

**DISCUSSION**

Add Health is a longstanding, nationally representative cohort of individuals with rich longitudinal data and a follow up period of nearly 25 years. Here, for the first time, we leverage the wealth of data available in this cohort to evaluate health outcomes in early adulthood that may be associated with long-term exposure to air pollution. Specifically, we: (1) examined whether multi-year exposures to O₃ and PM₂.₅ differed by individual-level characteristics; and (2) estimated associations between air pollution exposure during the cohort’s early to mid-20s and multiple markers of cardiometabolic health measured when the cohort entered their late 20s.

We observed associations of 2002–07 O₃ exposure with elevated odds of hypertension, obesity, diabetes, and metabolic syndrome. Results were similar for a shorter 2-year lagged period (2006–07) O₃ exposure, which was associated with elevated odds of obesity, diabetes, metabolic syndrome, and inflammation. Long-term exposure to PM₂.₅ (2002–07) was associated with elevated odds of hypertension, and no statistically significant associations were observed between 2006–07 PM₂.₅ exposure and health outcomes examined in our study.
Associations were in the expected direction, i.e., higher air pollution levels were associated with elevated odds of cardiometabolic outcomes, and we observed stronger evidence for associations between health outcomes and O₃ compared to PM₂.₅.⁵² This is notable because the current body of evidence for cardiovascular and cardiometabolic health outcomes is strongest for PM₂.₅ exposure, although associations between O₃ and cardiometabolic health have also been observed.⁵³ For example, a 2021 meta-analysis of ambient air pollution exposure and blood pressure in children and adolescents concluded that short-term and long-term exposures to air pollution, specifically PM₂.₅, may increase blood pressure, but the authors stated that conclusions could not be drawn for O₃ due to the limited number of studies.⁵⁴ Our results suggest that investment in additional studies of O₃ is warranted. Additionally, future work should evaluate whether more recent exposures or long-term, cumulative exposures are more relevant to health outcomes.

Although the magnitudes of the associations between air pollution exposure and health outcomes estimated here were small, air pollution exposure is ubiquitous, and elevated odds ratios applied to an entire population translate into substantial impacts on public health. Moreover, air pollution exposure is modifiable, and the public health risk it poses can be mitigated through regulation, e.g., more stringent ambient air quality standards. To this point, Brauer et al (2021) recently argued that because air pollution exposure is pervasive, reducing pollutant concentrations presents a powerful and critical opportunity to *equitably* reduce cardiovascular disease burden.¹³ The European Society of Cardiology also recognized air pollution exposure as a major modifiable risk factor relevant to the prevention of cardiovascular disease and emphasized the need for research regarding the role of air pollution in relation to early markers of cardiovascular health such as hypertension.⁵⁵

Multiple cohort studies in the US have explored associations between air pollution exposures and cardiovascular and/or cardiometabolic health outcomes. To date, most have focused on middle aged and older adults with minimum ages of 45–50 years and/or are non-representative or small, geographically localized samples. Examples include the Multi-Ethnic Study of Atherosclerosis (MESA),⁵⁵ Atherosclerosis Risk in Communities (ARIC),⁷ the Framingham Heart Study,⁷³ the Jackson Heart Study,⁵⁹ the Women’s Health Initiative (WHI),⁶⁰ and the Study of Women’s Health Across the Nation (SWAN),⁶¹ among others. In these cohorts, associations have been observed between exposures to PM₂.₅ and/or O₃ (e.g., ≥1 year) and: fatal and non-fatal cardiovascular events;⁶⁰ diabetes;⁶² impaired renal function;⁶³ systemic inflammation;⁶⁵ and arterial injury and subclinical markers of arterial disease that are predictive of coronary heart disease and stroke in individuals without cardiovascular disease.⁷

The present study of the Add Health cohort is one of few that examines long-term air pollution exposures and cardiometabolic health outcomes in a national cohort of young adults from communities across the US. It adds to a growing body of literature on air pollution exposure and cardiometabolic health in children and young adults. For example, a study of 158 individuals in Southern California aged 17–22 years, found that one-year NO₂ exposure was associated with higher fasting serum lipid measures (e.g., total cholesterol) and one-month O₃ exposure was associated with higher triglyceride levels and lower HDL levels.⁶⁴ Results of another study in Southern California that examined 173 young adults aged 18–23 years suggested that long-term (e.g., ≥1 year) air pollution exposure may contribute to the metabolic dysfunction in youth through lipolysis and altered fatty acid metabolism.⁶⁵ A third study from California observed that higher annual average PM₂.₅ exposure over 3 years of follow up was associated with more rapid declines in insulin sensitivity over time and lower insulin sensitivity, independent of adiposity, in overweight and obese children aged 8–15 years at enrollment.⁶⁶ A longitudinal study of approximately 3,300 children in multiple Southern California communities enrolled at 9-10 years of age and followed until age 18 found that traffic density was associated with BMI at age 18, particularly in females.⁶⁷ Although we did not observe associations between PM₂.₅, O₃, and cholesterol specifically, our results add to the existing literature indicating long term air pollution exposure is associated with multiple measures of cardiometabolic health, even in young adults.
In addition to associations between air pollution exposure and health, we also observed differences in PM$_{2.5}$ and O$_3$ exposure by race/ethnicity. Non-Hispanic Black participants (and non-Hispanic White participants) had the highest O$_3$ exposure, while Other race group participants had the lowest O$_3$ exposure. Non-Hispanic Black participants also had the highest PM$_{2.5}$ exposure, while non-Hispanic White participants had the lowest PM$_{2.5}$ exposures. These findings are consistent with literature examining racial/ethnic disparities in air pollution exposure, which have generally found that non-Hispanic Blacks have higher exposures to air pollutants, especially PM$_{2.5}$.\textsuperscript{68-71} This disparity may reflect, among other things, documented patterns of siting hazardous waste sites, polluting industrial facilities, and other undesirable activities or contaminated land-use types disproportionately in communities with lower socioeconomic status and a higher proportion of minoritized racial/ethnic groups.\textsuperscript{71,72} For this reason, it has been argued that reducing air pollution exposure is one potential pathway for reducing racial/ethnic disparities in cardiometabolic health.\textsuperscript{13}

This study has several limitations. Add Health collects current residential address at each wave but does not collect residential histories. Thus, we do not have between-wave residential address information. Average air pollution exposures were estimated assuming that individuals resided at their Wave III residential address between 2002–05, and Wave IV residential address between 2006–07. Additionally, the FAQSD data attached to the Add Health data and used to estimate air pollution exposures was only available at the census tract level and are not available prior to 2002. Air pollution estimates continue to evolve with increasingly higher resolution and made publicly available. Future studies of high-resolution estimates, including data prior to 2002, are needed in Add Health and other cohorts.

However, this study has important strengths. This is the first study of air pollution exposures and health outcomes in the Add Health cohort. This cohort represents a large national sample of young people that includes individuals from across the US, in both urban and rural environments, and is a longstanding cohort with rich individual- and contextual-level longitudinal data. Although associations between air pollution exposure and cardiovascular and cardiometabolic health have been evaluated in other study cohorts, prior results were derived from overwhelmingly middle aged and older cohorts and/or were based on small or local samples. Examining health outcomes earlier in the life course may provide more insight into the relevance of environmental exposures in developing early markers of cardiometabolic health (or disease) and thus more opportunities for early, targeted, and/or preventative or mitigative intervention.

This study adds to the body of evidence that exposures of air pollutants may be key risk factors for cardiovascular disease, including hypertension, high BMI, diabetes, inflammation, and metabolic syndrome in early life.\textsuperscript{26,51} Biological mechanisms through which air pollution influences development of cardiovascular and cardiometabolic disease, or risk factors for disease have been postulated, including increased systemic inflammation,\textsuperscript{73,74} platelet activation in the bloodstream,\textsuperscript{75,76} autonomic nervous system changes,\textsuperscript{77,78} alterations in the vascular cell type composition,\textsuperscript{74} and alterations in the gut microbiome,\textsuperscript{79} but overall remain poorly understood. Future research would benefit from bringing together longitudinal studies representing a broad lifespan (childhood, adolescent, and adulthood) with repeated biological assessments, precisely measured environmental exposures capturing critical development periods, and clinical follow-up to better understand if and how long-term air pollution exposures shape cardiovascular and cardiometabolic disease risks across the life course.

ACKNOWLEDGEMENTS

**Funding sources.** This research uses data from Add Health, funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with
cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. Support was provided by the RTI University Scholars program (Harris), Whitehead Scholars program at Duke School of Medicine (Bravo), and the RTI Fellow program (Johnson).

Disclosures: None to declare.

Conflict of interest. None to declare.

REFERENCES


27. Harris KM. *The Add Health Study: Design and Accomplishments.* Carolina Population Center at the University of North Carolina at Chapel Hill; 2013.


### TABLES

#### Table 1. Distributions of participants’ characteristics for those who participated in Wave III and Wave IV (n=11,259)*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean value (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at exam, Wave IV, years</td>
<td>28.4 (1.78)</td>
</tr>
<tr>
<td>O₃ exposure (ppb), 2002-2007</td>
<td>38.3 (3.67)</td>
</tr>
<tr>
<td>PM₂.₅ exposure (µg/m³), 2002-2007</td>
<td>12.7 (2.47)</td>
</tr>
<tr>
<td>Women, %</td>
<td>5,961 (52.9)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>2,550 (22.6)</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>6,881 (61.1)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1,281 (11.4)</td>
</tr>
<tr>
<td>Other</td>
<td>547 (4.88)</td>
</tr>
<tr>
<td>Health outcomes</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>2,840 (26.1)</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>921 (8.18)</td>
</tr>
<tr>
<td>Obese (BMI &gt; 30)</td>
<td>4,192 (37.8)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>811 (7.20)</td>
</tr>
<tr>
<td>High inflammation</td>
<td>4,074 (38.7)</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td>2,331 (20.7)</td>
</tr>
</tbody>
</table>

*Summary statistics are presented for the analytic sample size of Wave III and Wave IV participants (n=11,259).

#### Table 2. Distributions of O₃ and PM₂.₅ exposure by demographic characteristics

<table>
<thead>
<tr>
<th></th>
<th>2002-2007 (n=11,259)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ozone</td>
<td>PM₂.₅</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Concentration (ppb)</td>
<td>p-value</td>
<td>Concentration (µg/m³)</td>
<td>p-value</td>
</tr>
<tr>
<td>Overall</td>
<td>38.3 (3.67)</td>
<td>0.426</td>
<td>12.7 (2.47)</td>
<td>0.982</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>38.3 (3.67)</td>
<td>&lt; 0.0001</td>
<td>12.7 (2.49)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Male</td>
<td>38.3 (3.67)</td>
<td></td>
<td>12.7 (2.49)</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>38.6 (3.35)</td>
<td></td>
<td>13.6 (2.05)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>38.6 (3.45)</td>
<td></td>
<td>12.2 (2.27)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>37.0 (4.18)</td>
<td></td>
<td>12.8 (3.28)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>36.1 (4.85)</td>
<td></td>
<td>13.4 (3.06)</td>
<td></td>
</tr>
</tbody>
</table>

*t-tests with unequal variances were used to test for differences in O₃ and PM₂.₅ exposure by sex. The Kruskal Wallis test was used to test for differences in O₃ and PM₂.₅ by racial/ethnic group. P-values in the table correspond to those obtained from t-tests (sex) and Kruskal Wallis tests (race/ethnicity). Additionally, a Dunn’s test was used for pairwise comparisons of air pollutant concentrations for different racial/ethnic groups.
Table 3. Wave IV health outcomes and air pollution exposure (2002-2007) in adjusted GEE logistic regression models

<table>
<thead>
<tr>
<th>Health outcome (n)</th>
<th>Ozone</th>
<th>PM$_{2.5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Hypertension (n=10,883)</td>
<td><strong>1.015 (1.011, 1.029)</strong></td>
<td><strong>0.0342</strong></td>
</tr>
<tr>
<td>Hyperlipidemia (n=11,259)</td>
<td>1.009 (0.991, 1.028)</td>
<td>0.310</td>
</tr>
<tr>
<td>Obese (n=11,101)</td>
<td><strong>1.022 (1.004, 1.040)</strong></td>
<td><strong>0.0192</strong></td>
</tr>
<tr>
<td>Diabetes (n=11,259)</td>
<td><strong>1.032 (1.009, 1.054)</strong></td>
<td><strong>0.00502</strong></td>
</tr>
<tr>
<td>Inflammation (n=10,514)</td>
<td>1.012 (0.999, 1.025)</td>
<td>0.0627</td>
</tr>
<tr>
<td>Metabolic syndrome (n=9,518)</td>
<td><strong>1.028 (1.014, 1.041)</strong></td>
<td><strong>0.0000770</strong></td>
</tr>
</tbody>
</table>

*a* GEEs were fit as logistic regression models with a binary health outcome, Primary Sampling Unit School Identifier (PSUSCID) as the cluster ID variable, and an exchangeable correlation matrix. Models adjusted for 2002-07 pollutant concentration, age at Wave IV (years), sex (female or male), and race/ethnicity (non-Hispanic Black, non-Hispanic White, Hispanic, and Other).
FIGURES

Figure 1. Sample size restrictions flow chart

1. Add Health
   N=20,745
2. Wave III and Wave IV
   N=13,034
3. Geocoded
   N=12,790
4. Air pollution
   N=12,609
5. Non-missing
   race/ethnicity, age, sex
   N=12,609
6. US born
   N=11,682
7. Not pregnant (Wave IV)
   N=11,259

- C-reactive protein
  N=10,514
- Hypertension
  N=10,883
- Hyperlipidemia
  N=11,259
- Metabolic syndrome
  N=9,518
- BMI
  N=11,101
- Diabetes
  N=11,259
Figure 2. Residential address collection, availability of air pollution data, and assignment of air pollution data to Add Health cohort members

<table>
<thead>
<tr>
<th>Collection of residential address in Add Health cohort</th>
<th>Availability of FAQSD air pollution data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2002</td>
</tr>
<tr>
<td></td>
<td>2003</td>
</tr>
<tr>
<td></td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td>2005</td>
</tr>
<tr>
<td></td>
<td>2006</td>
</tr>
<tr>
<td></td>
<td>2007</td>
</tr>
<tr>
<td></td>
<td>2009</td>
</tr>
</tbody>
</table>

Wave III residential address used to assign air pollution exposure (census tract) for 2002-2005

Wave IV residential address used to assign air pollution exposure (census tract) for 2006-2009

*FAQSD air pollution data were not available in 2001, which is shown in gray. Air pollution data from 2008 and 2009 were merged with Add Health participant data as shown here, but air pollution data from 2008-2009 were not used in the present study.*