Is residential proximity to polluted sites during pregnancy associated with preterm birth or low birth weight? Results from an integrated exposure database in North Carolina (2003–2015)

Corinna Keeler [1]✉, Thomas J. Luben [2], Nina Forestieri [3], Andrew F. Olshan [1] and Tania A. Desrosiers [1]

BACKGROUND: Preterm birth (PTB) and term low birth weight (LBW) have been associated with pollution and other environmental exposures, but the relationship between these adverse outcomes and specific characteristics of polluted sites is not well studied.
OBJECTIVES: We conducted a retrospective cohort study to examine relationships between residential proximity to polluted sites in North Carolina (NC) and PTB and LBW. We further stratified exposure to polluted sites by route of contaminant emissions and specific contaminants released at each site.
METHODS: We created an integrated exposure geodatabase of polluted sites in NC from 2002 to 2015 including all landfills, Superfund sites, and industrial sites. Using birth certificates, we assembled a cohort of 1,494,651 singleton births in NC from 2003 to 2015. We geocoded the gestational parent residential address on the birth certificate, and defined exposure to polluted sites as residence within one mile of a site. We used log-binomial regression models to estimate adjusted risk ratios (aRR) and 95% confidence intervals (CI). Binomial models were used to estimate adjusted risk differences (aRD) per 10,000 births and 95% CIs for associations between exposure to polluted sites and PTB or LBW.
RESULTS: We observed weak associations between residential proximity to polluted sites and PTB [aRR(95% CI): 1.07(1.06,1.09); aRD(95% CI): 61(48,74)] and LBW [aRR(95% CI): 1.09(1.06,1.12); aRD(95% CI): 24(17,31)]. Secondary analyses showed increased risk of both PTB and LBW among births exposed to sites characterized by water emissions, air emissions, and land impoundment. In analyses of specific contaminants, increased risk of PTB was associated with proximity to sites containing arsenic, benzene, cadmium, lead, mercury, and polycyclic aromatic hydrocarbons. LBW was associated with exposure to arsenic, benzene, cadmium, lead, and mercury.
SIGNIFICANCE: This study provides evidence for potential reproductive health effects of polluted sites, and underscores the importance of accounting for heterogeneity between polluted sites when considering these exposures.

IMPACT STATEMENT: We documented an overall increased risk of both PTB and LBW in births with gestational exposure to polluted sites using a harmonized geodatabase of three site types, and further examined exposures stratified by site characteristics (route of emission, specific contaminants present). We observed increased risk of both PTB and LBW among births exposed to sites with water emissions or air emissions, across site types. Increased risk of PTB was associated with gestational proximity to sites containing arsenic, benzene, cadmium, lead, mercury, and polycyclic aromatic hydrocarbons; increased risk of LBW was associated with exposure to arsenic, benzene, cadmium, lead, and mercury.

Keywords: Preterm birth; Low birth weight; Adverse birth outcomes; Polluted sites; Superfund; Industrial sites

Journal of Exposure Science & Environmental Epidemiology (2023) 33:229–236; https://doi.org/10.1038/s41370-022-00475-0

INTRODUCTION

Preterm birth (PTB) and low birth weight (LBW) are leading causes of neonatal morbidity and mortality in the United States, and are both associated with increased disability and increased medical expenses across the life course [1, 2]. The rate of both PTB and LBW has increased steadily in the United States each year from 2014 to 2019, reaching a PTB rate of 10.23% of births in 2019 and a LBW rate of 8.31% of all births in 2019 [3].

Both PTB and LBW have been linked to various prenatal environmental exposures including air pollution [4–9] and groundwater contamination [10–12]. These outcomes have also been examined in the context of gestational residential exposure to

[1]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [2]Center for Public Health and Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Durham, NC, USA. [3]Birth Defects Monitoring Program, State Center for Health Statistics, Division of Public Health, North Carolina Department of Health and Human Services, Raleigh, NC, USA. ✉email: cykeeler@live.unc.edu

Received: 13 April 2022 Revised: 30 August 2022 Accepted: 31 August 2022
Published online: 13 September 2022

polluted sites [13], which can include landfills, hazardous waste sites, industrial sites, incinerators, and other sites of point-source contaminant emissions. Polluted sites include sites that have been formally classified as hazardous waste sites by a government or a regulatory mechanism, such as sites on the Superfund registry in the United States, as well as other sites with known pollutants that do not appear on official registries of hazardous waste sites but nevertheless are the source of known environmental contamination. Previous epidemiologic research regarding relationships between adverse birth outcomes and polluted sites has largely focused on one specific type of facility, such as landfills [14–17], active industrial sites [18, 19], or hazardous waste sites [20, 21], and these studies provide inconsistent evidence about the both the direction and magnitude of the association between polluted sites and PTB and LBW. A recent systematic review found that within each category of polluted sites (landfills, industrial sites, and hazardous waste sites) there were studies which documented no association, as well as studies that identified positive associations, between proximity-based exposure to those sites and adverse birth outcomes [13].

Polluted sites across categories and facility registries may have shared characteristics that are relevant to their potential impact on human health, such as the specific contaminants emitted at each site; the route of emission into water, air, or land-based releases; and the industrial sector with which the facility is associated [22]. By integrating data on a variety of types of polluted sites, it is possible to examine relationships between adverse health outcomes and the specific characteristics of a site that could underpin exposure pathways or health mechanisms. For example, by integrating data about shared characteristics of polluted sites across multiple registries, it is possible to examine the associations between PTB and residence near any point source of a known harmful substance like arsenic [23], regardless of whether that source is a landfill, a Superfund site, or an industrial site, which increases statistical power for examining specific environmental exposures.

In this study, we developed an exposure database that harmonized spatial and tabular data between three different types of polluted sites. We leveraged a large administrative cohort of births in North Carolina with geocoded addresses to examine relationships between PTB and LBW and residential proximity to different types of sites, different routes of emission of point-source pollutants, and different specific contaminants.

METHODS

Study population

We constructed a retrospective cohort of all singleton births in NC from 2003 to 2015 ($N = 1,545,579$) using birth certificate records provided for this study by the NC Department of Health and Human Services' Division of Public Health. Observations with missing information for birth date ($N = 1$, 0.00%), birth weight ($N = 628$, 0.04%), or estimated gestational age ($N = 853$, 0.05%) were excluded. To account for fixed cohort bias [24], we censored the cohort to only include births that could have achieved 22 weeks gestation by January 1, 2003 (conceived after July 31, 2002) and births that could have achieved 44 weeks gestation by December 31, 2015 (conceived before February 26, 2015).

Gestational parent address at delivery as recorded on the birth certificate was geocoded using ArcGIS 10.8, and births with gestational parent addresses that were missing or unable to be geocoded ($N = 106$) or were located outside of North Carolina ($N = 1587$) were excluded. After exclusions, the final analytical dataset contained 1,494,651 observations.

Outcome ascertainment

We defined preterm birth as delivery before the completion of 37 weeks gestation, and defined term low birth weight as birth weight less than 2500 g for births with a gestational age of 37 weeks or greater [1, 2]. Birth weight and clinical estimate of gestational age at delivery were obtained from the birth certificate.

Exposure data

We created a spatial database of polluted sites in NC that included three distinct categories of sites: toxic industrial sites, Superfund sites, and landfills. For each category, a site was included if it was on a list or registry in at least one of the study period years (2002–2015), along with the year(s) it was present on the registry. (While outcome data encompassed births from 2003 to 2015, polluted sites from 2002 were included to assess gestational exposure for births in 2003.) The spatial distribution of all polluted sites included in our exposure database is shown in Fig. 1. Of the 100 counties in North Carolina, 93 contained at least one polluted site and 71 contained 5 or more sites.

Industrial sites. Data about industrial sites were obtained from the United States Environmental Protection Agency (EPA) Toxic Release Inventory (TRI) program. The TRI program requires industrial facilities that manage, manufacture, release, or use chemicals known to be toxic, carcinogenic, or otherwise harmful to human health to submit annual reports about chemical releases including the amount of each chemical released and the route of emission; the EPA makes these reports publicly available including the auxiliary information about quantity and type of chemical releases. We obtained the TRI annual data for all facilities in the United States from 2002 to 2015 from the EPA website, and identified all facilities in NC or within 5 miles of NC for inclusion in our exposure database ($N = 1454$). We summarized additional relevant information for each facility-year observation including total toxic chemical emissions for each site for each year, whether or not there were toxic releases into the air and/or water for each site for each year, and the specific chemicals that were reported at each site for each year.

Superfund sites. Superfund sites are highly contaminated sites that contain materials known to present significant harms to human health [25]. Unlike TRI facilities, Superfund sites are usually no longer operational or actively generating new toxic substances, but the concentration or type

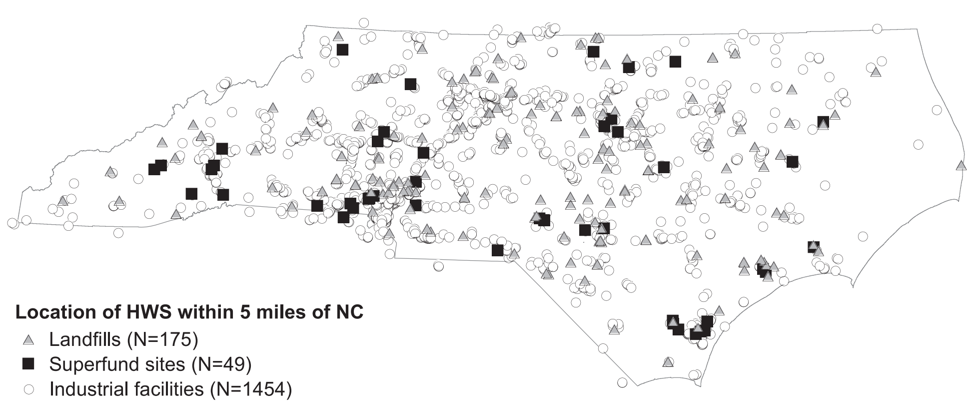

Fig. 1 Map showing spatial location of all polluted sites (HWS) in NC or within 5 miles of NC. The map includes all sites that were present on a public EPA registry of industrial facilities, a public EPA registry of Superfund sites, or a public NCDEQ registry of landfills, during at least one year from 2002–2015.

of toxic material is of such public health concern that the site is prioritized for federal cleanup funding and remediation management through the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) [22]. These federally recognized sites are documented on a registry called the National Priority List (NPL). We obtained the spatial location of all NPL Superfund sites located in NC or within five miles of NC for inclusion in the exposure database ($N = 49$) from the EPA Envirofacts database website. We retained information about the years each site was on the NPL, cleanup activities at the site and (where relevant) removal from the NPL, the specific contaminants that were documented at each site, and the contaminated areas or media at each site. We recoded the contaminated areas or media variable to represent route of contaminant emission.

Landfills. We obtained spatial data about landfill locations from the NC Department of Environmental Quality (NCDEQ). NCDEQ maintains a public online database of all actively permitted landfills for each year, along with their spatial location and the types of materials sequestered there (i.e., industrial waste, residential waste, etc.). We identified landfills with active permits any time from 2002 to 2015 ($N = 175$). Data were not available regarding the specific chemicals, substances, or contaminants present at landfills.

Integration of ATSDR data. To identify specific contaminants for further study, we used data from the Agency for Toxic Substances and Disease Registry (ATSDR), a division of the United States Centers for Disease Control (CDC). The ATSDR maintains a list of all substances present at Superfund sites, and creates a biennial Substance Priority List (SPL) which ranks nearly 300 substances according to the threat they pose to human health [26]. We obtained the SPL rankings from the ATSDR for the entire study period (2003–2015), and identified any substance that was in the top ten of the SPL at any point during the study period for further analysis. The same ten substances appeared in the top ten of the ranked SPL throughout the entire study period: arsenic, benzene, benzo(a)pyrene, benzo(b)fluoranthene, cadmium, lead, mercury, polychlorinated biphenyls, polycyclic aromatic hydrocarbons (PAH), and vinyl chloride. Benzo(a) pyrene and benzo(b)fluoranthene are both species of polycyclic aromatic hydrocarbons, so these two substances were combined into the PAH category [27]. We then identified each time that any of these SPL top ten substances were documented at either the industrial facilities or Superfund sites in our exposure database, using a matching algorithm combining the substance name and/or its unique Chemical Abstract Service registry number to flag the polluted sites in the database that contained each of the identified chemicals.

Exposure assignment. For our primary analyses, each birth was categorized as exposed to a polluted site(s) if the gestational parent residential address on the birth certificate was located within 1 mile (1.6 km) of any site on any polluted sites registry (TRI registry, Superfund NPL registry, or landfill permitting registry) during the year corresponding to the start of pregnancy (ie, year of conception). Date of conception were derived using clinical estimate of gestational age on the birth certificate. The 1-mile exposure threshold was selected due to the strong literature precedent for measuring residential proximity to polluted sites based on a 1-mile buffer [18, 28–30].

In sub-analyses, we evaluated potential exposure to polluted sites with stratified exposure categories based on the specific characteristics of the sites. We assigned exposure to industrial sites, Superfund sites, and landfills individually; exposure to land-based pollutants, air-based pollutants, and water-based pollutants individually; and exposure to sites containing ATSDR priority substances individually. These separate exposure categories were each defined according to the same criteria described above: residential address within 1 mile of polluted sites within each category, with matching between year of conception and presence of site on the registry in that year's reporting. To enable a sensitivity analysis examining associations across a range of distance thresholds, we categorized exposure using varying distance thresholds ranging from 0.5 miles to 3 miles: the main exposure of 1 mile (1.6 km), and other distance thresholds of 0.5 miles (0.8 km), 1.5 miles (2.41 km), 2 miles (3.2 km), and 3 miles (4.8 km). All exposure assignments were performed in R version 4.0.3.

Analytical approach

We calculated both relative and absolute measures of association between potential exposures to polluted sites and risk of PTB and term LBW, separately. We used log-binomial regression models to estimate crude and adjusted risk ratios (RR) and 95% confidence intervals (CI). We used binomial models with an identity link to estimate crude and adjusted risk differences (RD) and 95% CIs; risk differences for each outcome were expressed as the number of estimated additional cases per 10,000 live births, and interpreted as positive or negative compared to 0 births on an absolute scale. Relative and absolute models were estimated separately for both outcomes and for each exposure category. The reference group for all analyses was births with no exposure to any polluted site. For analyses of LBW, the analytic sample was restricted to term births only ($N = 1,362,980$).

We selected covariates for adjustment as informed by a directed acyclic graph (Supplementary Fig. 1). The adjustment set included gestational parent education, gestational parent age, gestational parent Medicaid status at delivery, and gestational parent smoking. Gestational parent Medicaid status and education were the most suitable proxies for socioeconomic status available from the birth certificate. We did not adjust for gestational parent race-ethnicity, which was collected as a combined 6-category variable on the birth certificate (White, Black, Hispanic, Asian/Asian-American, American Indian, and Other) through birthing person self-identification, because we wanted to avoid attenuation of any true variation by race-ethnicity in the association between residential proximity to polluted sites and adverse birth outcomes [31]; instead, we identified race-ethnicity as a possible effect measure modifier and conducted analyses of effect measure modification by race-ethnicity for both PTB and LBW and the primary exposure of residential address within one mile of any polluted site.

Additional sensitivity analyses examined associations between each of the exposure categories (site type: Superfund, industrial, or landfill; route of emission: air, water, or land; specific ATSDR contaminants present at sites) and PTB and LBW, across a range of distance thresholds (0.5–3.0 miles). Finally, we conducted a sensitivity analysis of the associations between specific contaminants and PTB and LBW by restricting the exposure classification to sites which reported the presence of exactly one SPL top ten contaminant throughout the entire study period, in order to account for possible confounding by contaminant co-occurrence.

All statistical analyses were conducted in R version 4.0.3. This study protocol was approved by the Institutional Review Boards of the University of North Carolina at Chapel Hill UNC IRB Number: 09-0828) and the NC Division of Public Health. The EPA's Human Subjects Research Officer also reviewed and approved this work (HSR-001254).

RESULTS

Study population

Among all singleton births in North Carolina from 2003 to 2015, 131,671 (8.8%) were classified as preterm. Among term births, 35,946 (2.6%) had a birth weight less than 2500 g. Based on a 1-mile threshold, 261,856 (17.5%) of births in the analytic dataset were in proximity to at least one polluted site based on gestational parent residential address. Both PTB and LBW were observed more frequently among births with proximity to polluted sites: 9.5% of births with exposure to polluted sites were preterm compared to 8.7% of births without exposure to polluted sites. Among term births, 3.0% of births with potential exposure to polluted sites were LBW compared to 2.6% of non-exposed births (Table 1). Table 1 describes the study population characteristics and the distribution of the covariates among exposed and unexposed births. Supplementary Table 1 presents the study population characteristics and distribution of covariates further stratified by type of polluted site.

Associations between residential proximity to polluted sites and PTB

We observed an increased risk of PTB among births with gestational parent residence within one mile of a polluted site [adjusted RR [aRR] (95% CI): 1.07 (1.06,1.09)] (Table 2). In our analyses of absolute risk, residential proximity to any polluted site was associated with an adjusted risk difference of 61 cases of PTB per 10,000 births (95% CI: 48, 74).

We conducted sub-analyses examining specific types and characteristics of polluted sites (Table 2). We observed evidence of greater risk of PTB associated with proximity to industrial sites

Table 1. Distribution of preterm births, term births classified as low birth weight, and selected characteristics from birth certificates among singleton births in North Carolina (2003–2015), stratified by gestational parent (GP) residential proximity to polluted sites.

	GP residence ≤ 1 mile from polluted site	GP residence > 1 mile from polluted site	Total
Total *N*	261,856 (17.5%)	1,232,795 (82.5%)	1,494,651
Term *N*	236,891 (17.4%)	1,126,089 (82.6%)	1,362,980
PTB (% of total *N*)	24,965 (9.5%)	106,706 (8.7%)	131,671 (8.8%)
Term LBW (% of term *N*)	7150 (3.0%)	28,796 (2.6%)	35,946 (2.6%)
GP education[a]			
Less than HS	67,764 (28.0%)	216,389 (19.1%)	284,153 (20.7%)
Completed HS	70,740 (29.2%)	290,268 (25.7%)	361,008 (26.3%)
More than HS	103,762 (42.8%)	623,562 (55.2%)	727,324 (53.0%)
GP age (years)[a]			
10–19	32,904 (12.6%)	119,392 (9.7%)	152,296 (10.2%)
20–24	78,463 (30.0%)	314,956 (25.5%)	393,419 (26.3%)
25–29	71,906 (27.5%)	342,809 (27.8%)	414,715 (27.7%)
30–34	51,405 (19.6%)	291,023 (23.6%)	342,428 (22.9%)
35–39	22,395 (8.6%)	136,071 (11.0%)	158,466 (10.6%)
40–60	4776 (1.8%)	28,519 (2.3%)	33,295 (2.2%)
GP Medicaid[a]			
No	101,260 (38.7%)	629,323 (51.0%)	730,583 (48.9%)
Yes	160,596 (61.3%)	603,472 (49.0%)	764,068 (51.1%)
GP smoking[a]			
No	213,874 (88.3%)	1,008,958 (89.3%)	1,222,832 (89.1%)
Yes	28,468 (11.7%)	121,116 (10.7%)	149,584 (10.9%)
GP race[a]			
NH White	110,963 (42.4%)	726,754 (59.0%)	837,717 (56.0%)
NH Black	83,862 (32.0%)	266,517 (21.6%)	350,379 (23.4%)
Hispanic/Latinx	53,875 (20.6%)	179,879 (14.6%)	233,754 (15.6%)
NH Asian	9564 (3.7%)	40,122 (3.3%)	49,686 (3.3%)
NH American Indian	3066 (1.2%)	17,320 (1.4%)	20,386 (1.4%)
Other/unknown	526 (0.2%)	2203 (0.2%)	2729 (0.2%)

GP gestational parent, *HS* high school, *LBW* low birth weight, *NH* non-hispanic, *PTB* preterm birth.

[a]For all covariates, counts and percentages are reported from the full analytic dataset ($N = 1{,}494{,}651$), which includes both preterm and term births.

[aRR (95% CI): 1.07 (1.06, 1.09)] or landfills [aRR (95% CI): 1.06 (1.02,1.11)], while we did not observe site-specific associations between PTB and Superfund sites [aRR (95% CI): 1.03 (0.98,1.08)]. In the analysis of route of contaminant emission, there were positive associations between PTB and air emissions [aRR (95% CI): 1.07 (1.05,1.09)], water emissions [aRR (95% CI): 1.09 (1.05,1.12)], and land emissions [aRR (95% CI): 1.07 (1.06,1.09)]. The greatest risk difference we observed was for births with potential exposure to polluted sites with water emissions [aRD ((95% CI): 77 cases of PTB per 10,000 live births (51, 103)).

Associations between residential proximity to polluted sites and LBW

Residential proximity to any polluted site was associated with a crude risk ratio of 1.18 (95% CI: 1.15, 1.21) and an adjusted risk ratio of 1.09 (95% CI: 1.06, 1.12) for LBW (Table 3). In our analyses of absolute risk, residential proximity to any polluted site was associated with an adjusted risk difference of 24 cases of LBW per 10, 000 term live births (95% CI: 17, 31).

In analyses stratified by exposure characteristics, we observed increased risk of LBW associated with proximity to industrial facilities [aRR (95% CI): 1.11 (1.08,1.14)], and all routes of contaminant emission including air, water, and land emissions. We did not observe an association between LBW and either Superfund sites [aRR (95% CI): 1.07 (0.98,1.17)] or landfills [aRR (95% CI): 0.96 (0.87,1.06)]. Full results for LBW and all exposure categories are included in Table 3.

Associations between specific contaminants and PTB and LBW

We evaluated associations between both PTB and LBW and sites containing contaminants in the top ten of the ranked ATSDR Substance Priority List (SPL) (Table 4). Residential proximity to any site containing, processing, or emitting any SPL top-ten contaminant was associated with both PTB [aRR (95% CI): 1.07(1.05,1.09)] and LBW [aRR (95% CI): 1.06 (1.02,1.11)]. Proximity to any site containing a SPL top-ten substance was associated with an increased absolute risk of 63 cases of PTB per 10,000 live births (95% CI: 45,81) and an increased absolute risk of 18 cases of LBW per 10,000 term live births (95% CI: 8,28).

In analyses of specific contaminants and PTB, we observed particularly notable effects for proximity to sites containing PAH [aRR (95% CI): 1.14 (1.09,1.18)], benzene [aRR (95% CI): 1.11 (1.06,1.16)], cadmium [aRR (95% CI): 1.12 (1.05,1.19)], arsenic [aRR (95% CI): 1.10 (1.04,1.17)], and mercury [aRR (95% CI): 1.09 (1.03,1.15)] (Table 4). Proximity to sites containing lead was also associated with increased risk of PTB [aRR (95% CI): 1.05 (1.03,107)] (Table 4).

In analyses of specific contaminants and LBW, exposure to sites containing cadmium was associated with increased risk of LBW

Table 2. Associations between gestational parent residential address within 1 mile (1.6 km) of a polluted site, site characteristics, and PTB (preterm birth) in North Carolina (2003–2015).

	N exposed PTB/*N* exposed term births	Risk ratio (RR)		Risk difference (RD) per 10,000	
		Crude RR (95% CI)	Adjusted[a] RR (95% CI)	Crude RD (95% CI)	Adjusted[a] RD (95% CI)
Any polluted site	24965/236891	1.10 (1.09, 1.12)	1.07 (1.06, 1.09)	88 (76, 101)	61 (48, 74)
Type of site					
Industrial facility	22338/210329	1.11 (1.09, 1.12)	1.07 (1.06, 1.09)	95 (82, 108)	64 (51, 77)
Superfund site	1878/19611	1.01 (0.97, 1.05)	1.03 (0.98, 1.08)	8 (−30, 47)	30 (−9, 70)
Landfill	1823/17985	1.06 (1.02, 1.11)	1.06 (1.02, 1.11)	55 (15, 96)	46 (5, 88)
Route of emission					
Air	16701/158187	1.10 (1.09, 1.12)	1.07 (1.05, 1.09)	90 (75, 105)	61 (46, 77)
Water	5345/50676	1.10 (1.07, 1.13)	1.09 (1.05, 1.12)	89 (64, 114)	77 (51, 103)
Land	23241/220064	1.10 (1.09, 1.12)	1.07 (1.06, 1.09)	90 (77, 103)	62 (49, 75)
ATSDR SPL contaminant[b]	11204/106582	1.10 (1.08, 1.12)	1.07 (1.05, 1.09)	84 (68, 104)	63 (45, 81)

[a]Covariates for all adjusted models were gestational parent (GP) age, GP smoking, GP Medicaid status, and GP education.
[b]This category contains all sites that reported the presence of any substance in the top 10 of the ranked Agency for Toxic Substances and Disease Registry (ATSDR) Substance Priority List (SPL) at any point during the study period (2003–2015).

[aRR (95% CI): 1.23(1.09,1.39)]; sites containing mercury, lead, arsenic, and benzene showed slightly smaller effect estimates but were also positively associated with increased risk of LBW (Table 4).

There were 1502 polluted sites that reported information about the presence of specific contaminants, allowing them to be included in this sub-analysis. Of these, 788 reported the presence of one or more of the ATSDR SPL top ten contaminants (Supplementary Fig. 2); among those sites, 613 reported the presence of exactly one of the selected contaminants across the study period. We conducted a sensitivity analysis examining associations between PTB and LBW and the eight contaminants identified for this study, using this subset of 613 sites that did not have co-occurrence of SPL top-ten contaminants (Supplementary Table 2). Our main findings were consistent with this narrower exposure definition; in the sensitivity analyses focused on sites with only one contaminant, we observed positive associations between the same six contaminants and PTB and the same five contaminants and LBW.

Additional analyses

We conducted analyses of potential effect measure modification by race-ethnicity (Supplementary Table 3). We observed positive associations in stratified analyses of association between risk of PTB and proximity to polluted sites for the NH Black and Hispanic/Latinx strata, although the magnitude of the effects in relative analyses was small and confidence intervals for these two groups overlapped with other strata (NH Black aRR: 1.04 (1.02,1.07), Hispanic/Latinx aRR: 1.05 (1.01,1.09); for full results see Supplementary Table 3). In stratified analyses of absolute risk, residential proximity to polluted sites was associated with an adjusted risk difference of 52 cases of PTB per 10,000 live births (95% CI: 25,78) within the NH Black strata and 40 cases of PTB per 10,000 live births (95% CI: 13,66) within the Hispanic/Latinx strata. We did not observe meaningful differences in the association between exposure to polluted site and risk of LBW when stratifying aRRs by race-ethnicity.

We examined sensitivity of our main findings to the distance threshold used to define the residential address proximity exposure (Supplementary Table 4). The findings presented in Tables 2 and 3 were generally consistent across distance thresholds, though small numbers at the smallest distance threshold included in this analysis (0.5 miles/0.8 km) contributed to larger uncertainty in the effect estimates across exposure categories at that distance threshold.

DISCUSSION

We observed increased risk of PTB and LBW associated with gestational parent residential proximity to polluted sites, among singleton births in North Carolina from 2003 to 2015. We conducted additional analyses to investigate whether specific characteristics of polluted sites may be associated with both PTB and LBW. In analyses categorizing contaminant emission into air, water, and land across all site types, we found that air emissions, water emissions, and land emissions are each associated with increased absolute and relative risk of PTB and LBW. In analyses of categories of polluted sites, we observed that increased risk of PTB was associated with proximity to both industrial sites and landfills but not with Superfund sites. Term LBW was associated with only industrial sites. Finally, we examined relationships between both outcomes and the presence of toxic substances ranked in the top ten of the ATSDR SPL during the study period (regardless of site type). We observed increased risk of both PTB and LBW associated with each of these contaminants of increased public health concern, especially arsenic, benzene, cadmium, and polycyclic aromatic hydrocarbons.

We report increased risk of PTB and LBW in births with residential proximity to industrial sites. These findings and the magnitude of effect that we observed are consistent with other studies that have identified increased rates of LBW and PTB outcomes near industrial facilities [19, 32]. On the other hand, we did not observe an association between exposure to Superfund sites and either PTB or LBW (Tables 2, 3). Previous studies provide some evidence of increased risk of LBW near Superfund sites in New York State [21], and increased risk of PTB near a hazardous waste site in Nova Scotia (which is in Canada and therefore not part of the United States-based Superfund program) [20]; however, in general there is not a large body of evidence about relationships between these birth outcomes and hazardous waste sites. The sites on the Superfund registry are by definition highly contaminated but many are currently inactive or no longer operational; because the CERCLA legislation went into effect in 1980, many of these sites have been inactive since before 1980 [22]. One possible explanation for the different effects by site category observed in this study is that many Superfund sites

Table 3. Associations between gestational parent (GP) residential address within 1 mile (1.6 km) of a polluted site, site characteristics, and LBW (low birth weight) among term births in North Carolina (2003–2015).

	N exposed term LBW/N exposed non-LBW among term births	Risk ratio (RR)		Risk difference (RD) per 10,000	
		Crude RR (95% CI)	Adjusted[a] RR (95% CI)	Crude RD (95% CI)	Adjusted[a] RD (95% CI)
Any polluted site	7150/229741	1.18 (1.15, 1.21)	1.09 (1.06, 1.12)	46 (39, 54)	24 (17, 31)
Type of site					
Industrial facility	6496/203833	1.21 (1.18, 1.24)	1.11 (1.08, 1.14)	53 (45, 61)	28 (20, 35)
Superfund site	537/19074	1.07 (0.98, 1.16)	1.07 (0.98, 1.17)	18 (−5, 41)	19 (−3, 42)
Landfill	446/17539	0.97 (0.88, 1.06)	0.96 (0.87, 1.06)	−7 (−30, 26)	0 (−22, 22)
Route of emission					
Air	4880/153307	1.21 (1.17, 1.24)	1.11 (1.08, 1.15)	53 (44, 62)	30 (21, 38)
Water	1539/49137	1.19 (1.13, 1.25)	1.11 (1.06, 1.17)	48 (33, 63)	29 (14, 44)
Land	6746/213318	1.20 (1.17, 1.23)	1.10 (1.07, 1.13)	51 (43, 59)	27 (19, 34)
ATSDR SPL contaminant[b]	3141/103441	1.15 (1.11, 1.20)	1.06 (1.02, 1.11)	39 (29, 50)	18 (8, 28)

[a]Covariates for all adjusted models were gestational parent (GP) age, GP smoking, GP Medicaid status, and GP education.
[b]This category contains all sites that reported the presence of any substance in the top 10 of the ranked Agency for Toxic Substances and Disease Registry (ATSDR) Substance Priority List (SPL) at any point during the study period (2003–2015).

included in this exposure database are highly contaminated but had been inactive for 30 or more years at the start of the study period (2003), while industrial sites on the TRI registry are currently active facilities that are therefore subject to the TRI reporting standards. Further research could examine the role of time since site activity in order to identify a possible decay or decrease in the effect of exposure to a certain site and adverse birth outcomes; another line of related research could examine birth outcomes before and after site cleanup.

A strength of this study is the harmonized exposure database that we developed to incorporate spatial, temporal, and tabular information about polluted sites in NC. We combined data from an EPA database of industrial sites, an EPA database of Superfund sites, an NCDEQ database of landfills, and a CDC ATSDR database of substance toxicity. The exposure database developed as part of this project includes data for each facility for each year, allowing for temporal matching between year of conception and exposure to polluted sites that were operational or actively a site of concern during that year, thereby mitigating some exposure misclassification due to misspecification of timing of exposure relative to pregnancy. Furthermore, we identified site characteristics such as the route of pollutant emission and the presence of specific contaminants that could be relevant to the potential mechanisms of how exposure to polluted sites adversely affects human health, and recoded these variables to be consistent between different registries of polluted sites in the United States, allowing for exposure categorization by relevant exposure characteristics instead of simply relying on which registry a site appeared.

While the harmonized exposure database allowed us to examine these dimensions of polluted site exposure, it also had limitations. Spatial locations for polluted sites were determined using latitude-longitude coordinates provided by each registry, which we converted to spatial point locations for use in the exposure assignment. Therefore, our exposure assignment does not account for facility size, or use boundaries or parcel lines to determine the spatial area covered by a specific polluted site. This could lead to exposure misclassification around larger sites, if a given site is large enough that the one-mile distance threshold used for defining exposure to that site does not accurately capture all the births with true residential address within a mile of the site (births could have be missed if they were less than a mile from the boundary of the site area, but more than a mile from the point location identified by the spatial coordinates). However, this exposure misclassification would likely undercount the exposed births, leading to attenuated estimates of association. Additionally, residential proximity to polluted sites during pregnancy is merely a proxy for actual gestational exposure to a polluted site or a toxic substance. We were not able to directly measure exposure or account for how different people living in proximity to polluted sites may have experienced different levels and types of exposures based on their daily mobility and activity patterns, which can affect an individual person's exposure environment [33, 34]. Finally, our data only included information on gestational parent residential address at the time of delivery and we were unable to account for possible residential relocation during pregnancy, which could have affected an individual's actual gestational exposure due to residential proximity to polluted sites.

Our outcome dataset was a retrospective administrative birth cohort developed from birth certificates, which offered both strengths and limitations. This large cohort captured all births in North Carolina from 2003 to 2015, providing sufficient geographic coverage to examine risk of PTB and term LBW within 1 mile of a polluted site. However, by virtue of this study design, we were limited to information on the birth certificate for ascertaining outcomes and covariates. We defined PTB in this study as less than 37 weeks completed gestation based on the clinical estimate of gestational age recorded on the birth certificate, which may not represent the true gestational age, particularly in lower birth weight infants [35]. However, studies that have validated clinical gestational age on the birth certificate against medical records have found relatively high agreement between methods of assessing gestational age [36, 37]. Additionally, we were unable to distinguish between spontaneous preterm birth, which includes both spontaneous preterm labor and premature rupture of membranes, and planned preterm delivery occurring when the mother or fetus meets indicating criteria, based on data available on the birth certificate. One estimate suggests that 30–35% of preterm deliveries are planned due to a medical indication [38], so the inability to ascertain spontaneous preterm birth only could lead to substantial misclassification of PTB in this study.

We leveraged the harmonized exposure database to examine exposure to specific substances known to be harmful to human health to better understand the relationships between these specific contaminants and PTB and LBW. We identified

Table 4. Adjusted effect estimates of association between adverse birth outcomes and gestational parent (GP) residence within 1 mile of a polluted site containing a substance in the top ten of the ranked ATSDR Substance Priority List, for both PTB and LBW.

Contaminant	PTB		LBW	
	Adjusted[a] RR (95% CI)	*Adjusted[a] RD per 10,000 (95% CI)*	*Adusted[a] RR (95% CI)*	*Adjusted[a] RD per 10,000 (95% CI)*
Any SPL top-10 contaminant[b]	1.07 (1.05, 1.09)	63 (45, 81)	1.06 (1.02, 1.11)	18 (8, 28)
Arsenic	1.10 (1.04, 1.17)	96 (41, 152)	1.17 (1.05, 1.31)	43 (11, 75)
Benzene	1.11 (1.06, 1.16)	95 (49, 141)	1.12 (1.02, 1.23)	34 (8, 61)
Cadmium	1.12 (1.05, 1.19)	99 (37, 161)	1.23 (1.09, 1.39)	53 (16, 90)
Lead	1.05 (1.03, 1.07)	46 (26, 67)	1.07 (1.02, 1.11)	21 (9, 32)
Mercury	1.09 (1.03, 1.15)	78 (27, 129)	1.11 (1.00, 1.23)	26 (−3, 55)
PAH[c]	1.14 (1.09, 1.18)	120 (83, 159)	1.05 (0.97, 1.13)	15 (−5, 36)
PCB[d]	1.05 (0.93, 1.19)	42 (−61, 145)	1.17 (0.91, 1.51)	16 (−39, 72)
Vinyl chloride	1.04 (0.97, 1.11)	32 (−22, 87)	1.06 (0.93, 1.20)	10 (−21, 40)

[a]All models were adjusted for gestational parent (GP) age, GP smoking, GP Medicaid status, and GP education.
[b]Contaminants of interest were identified as those in the top 10 of the ranked ATSDR Substance Priority List at any point during the study period (2003-2015).
[c]PAH: polycyclic aromatic hydrocarbons including benzo(a)pyrene, benzo(b)fluoranthene, and other polycyclic aromatic hydrocarbon compounds. These substances were listed separately on the SPL but are all under the umbrella category of PAHs and were collapsed into a single exposure category.
[d]*PCB* polychlorinated biphenyls.

contaminants for inclusion in our analysis based on their presence in the top ten of the ranked ATSDR SPL. Among the substances included in this analysis, increased risk of PTB was associated with residential proximity to sites containing arsenic, benzene, cadmium, lead, mercury, and PAH; similarly, increased risk of term LBW was associated with gestational parent residence near sites containing arsenic, benzene, cadmium, lead, and mercury. This study is the first that we are aware of to report associations between polluted sites containing many of these specific contaminants and PTB or LBW. Other studies have documented associations between metals, including lead and cadmium, and congenital abnormalities [39], and a study examining contaminated drinking water found relationships between benzene and PTB [12]. Our findings therefore add to the body of evidence regarding the adverse effect of toxic substances on fetal development and reproductive health, while suggesting that gestational exposure to these substances through residential proximity to polluted sites could be an under-explored source of exposure.

In conclusion, the complex etiology of PTB and LBW makes is difficult to design public health measures addressing these conditions [38, 40]. Many established risk factors for PTB and LBW are not intervenable: maternal age, maternal race, and genetics [40–42]. Other known risk factors for these adverse outcomes are individual behaviors during pregnancy that are difficult to address on a policy level [42]. However, environmental hazards such as air pollution, water contamination, and residential proximity to point-source pollutants are associated with PTB and LBW, and may be intervenable through environmental cleanup, site remediation, and regulation of certain contaminants. To that end, this study examined relationships between proximity to polluted sites and PTB and LBW, with an additional focus on specific characteristics of these sites that may be associated with elevated risk of these adverse birth outcomes. We incorporated industrial sites, Superfund sites, and landfills in a single exposure assessment and documented positive associations between both PTB and LBW and all polluted sites. We also observed increased risk of PTB and LBW due to exposure to different contaminant routes of emission and due to residence near sites containing specific substances in the top ten of the ranked ATSDR SPL. Given the prevalence of both PTB and LBW among births in the United States, and the fact that rates of these outcomes have been increasing in recent years, understanding the possible contributions of environmental exposures to PTB and LBW has important consequences for public health.

DISCLAIMER

The research described in this article has been reviewed by the Center for Public Health and Environmental Assessment, US EPA, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does the mention of trade names of commercial products constitute endorsement or recommendation for use.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. CDC. Preterm Birth [Internet]. 2020. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm. Accessed 17 Jun 2021.
2. CDC. Birthweight and Gestation [Internet]. 2021. https://www.cdc.gov/nchs/fastats/birthweight.htm. Accessed 17 Jun 2021.
3. Martin JA, Hamilton BE, Osterman MJK, Driscoll AK. Births: final data for 2019. Natl Vital- Stat Rep. 2021;70:1–51.
4. Liu Y, Xu J, Chen D, Sun P, Ma X. The association between air pollution and preterm birth and low birth weight in Guangdong, China. BMC Public Health. 2019;19:3.
5. Mekonnen ZK, Oehlert JW, Eskenazi B, Shaw GM, Balmes JR, Padula AM. The relationship between air pollutants and maternal socioeconomic factors on preterm birth in California urban counties. J Expo Sci Environ Epidemiol. 2021;31:503–13.
6. Alman BL, Stingone JA, Yazdy M, Botto LD, Desrosiers TA, Pruitt S, et al. Associations between PM2.5 and risk of preterm birth among liveborn infants. Ann Epidemiol. 2019;39:46–53. e2
7. Maisonet M, Bush TJ, Correa A, Jaakkola JJ. Relation between ambient air pollution and low birth weight in the Northeastern United States. Environ Health Perspect. 2001;109:351–6.
8. Ezziane Z. The impact of air pollution on low birth weight and infant mortality. Rev Environ Health. 2013;28:107–15.
9. Rogers JF, Dunlop AL. Air pollution and very low birth weight infants: a target population? Pediatrics. 2006;118:156–64.

10. Ferguson KK, O'Neill MS, Meeker JD. Environmental contaminant exposures and preterm birth: a comprehensive review. J Toxicol Environ Health B Crit Rev. 2013;16:69–113.
11. Sonnenfeld N, Hertz-Picciotto I, Kaye WE. Tetrachloroethylene in drinking water and birth outcomes at the US Marine Corps Base at Camp Lejeune, North Carolina. Am J Epidemiol. 2001;154:902–8.
12. Ruckart PZ, Bove FJ, Maslia M. Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: a cross-sectional study. Environ Health. 2014;13:99.
13. Kihal-Talantikite W, Zmirou-Navier D, Padilla C, Deguen S. Systematic literature review of reproductive outcome associated with residential proximity to polluted sites. Int J Health Geogr. 2017;16:20.
14. Morgan OWC, Vrijheid M, Dolk H. Risk of low birth weight near EUROHAZCON hazardous waste landfill sites in England. Arch Environ Health. 2004;59:149–51.
15. Goldberg MS, Goulet L, Riberdy H, Bonvalot Y. Low birth weight and preterm births among infants born to women living near a municipal solid waste landfill site in Montreal, Quebec. Environ Res. 1995;69:37–50.
16. Elliott P, Briggs D, Morris S, de Hoogh C, Hurt C, Jensen TK, et al. Risk of adverse birth outcomes in populations living near landfill sites. BMJ. 2001;323:363–8.
17. Gilbreath S, Kass PH. Adverse birth outcomes associated with open dumpsites in Alaska Native Villages. Am J Epidemiol. 2006;164:518–28.
18. Brender JD, Zhan FB, Suarez L, Langlois PH, Moody K. Maternal residential proximity to waste sites and industrial facilities and oral clefts in offspring. J Occup Environ Med. 2006;48:565–72.
19. Castelló A, Río I, García-Pérez J, Fernández-Navarro P, Waller LA, Clennon JA, et al. Adverse birth outcomes in the vicinity of industrial installations in Spain 2004-2008. Environ Sci Pollut Res Int. 2013;20(Jul):4933–46.
20. Dodds L, Seviour R. Congenital anomalies and other birth outcomes among infants born to women living near a hazardous waste site in Sydney, Nova Scotia. Can J Public Health. 2001;92:331–4.
21. Baibergenova A, Kudyakov R, Zdeb M, Carpenter DO. Low birth weight and residential proximity to PCB-contaminated waste sites. Environ Health Perspect. 2003;111:1352–7.
22. Wagner T. The complete guide to the hazardous waste regulations: RCRA, TSCA, HMTA, OSHA, and Superfund. 3rd ed. New York: JWiley; 1999. xxiv+536.
23. Agency for Toxic Substances and Disease Registry. TOXICOLOGICAL PROFILE FOR ARSENIC. 2007;559.
24. Neophytou AM, Kioumourtzoglou MA, Goin DE, Darwin KC, Casey JA. Educational note: addressing special cases of bias that frequently occur in perinatal epidemiology. Int J Epidemiol. 2021;50:337–45.
25. US EPA O. What is Superfund? [Internet]. US EPA. 2017. https://www.epa.gov/superfund/what-superfund. Accessed 17 Jun 2021.
26. Centers for Disease Control. Substance Priority List | ATSDR [Internet]. 2020 Accesssed 16 Nov 2021. https://www.atsdr.cdc.gov/spl/index.html.
27. Agency for Toxic Substances and Disease Registry. Toxicological profile for polycyclic aromatic hydrocarbons. 1995;487.
28. Brender JD, Suarez L, Langlois PH, Steck M, Zhan FB, Moody K. Are maternal occupation and residential proximity to industrial sources of pollution related? J Occup Environ Med. 2008;50:834–9.
29. Langlois PH, Brender JD, Suarez L, Zhan FB, Mistry JH, Scheuerle A, et al. Maternal residential proximity to waste sites and industrial facilities and conotruncal heart defects in offspring. Paediatr Perinat Epidemiol. 2009;23:321–31.
30. Malik S, Schecter A, Caughy M, Fixler DE. Effect of proximity to hazardous waste sites on the development of congenital heart disease. Arch Environ Health: Int J. 2004;59:177–81.
31. Ward JB, Gartner DR, Keyes KM, Fliss MD, McClure ES, Robinson WR. How do we assess a racial disparity in health? Distribution, interaction, and interpretation in epidemiological studies. Ann Epidemiol. 2019;29:1–7.
32. Porter TR, Kent ST, Su W, Beck HM, Gohlke JM. Spatiotemporal association between birth outcomes and coke production and steel making facilities in Alabama, USA: a cross-sectional study. Environ Health. 2014;13:85.
33. Shafran-Nathan R, Yuval, Levy I, Broday DM. Exposure estimation errors to nitrogen oxides on a population scale due to daytime activity away from home. Sci Total Environ. 2017;580:1401–9.
34. Steinle S, Reis S, Sabel CE. Quantifying human exposure to air pollution–moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci Total Environ. 2013;443:184–93.
35. Stanley Emery E III, Eaton A, Grether JK, Nelson KB. Assessment of gestational age using birth certificate data compared with medical record data. Paediatr Perinat Epidemiol. 1997;11:313–21.
36. Callaghan WM, Dietz PM. Differences in birth weight for gestational age distributions according to the measures used to assign gestational age. Am J Epidemiol. 2010;171:826–36.
37. Dietz PM, Bombard JM, Hutchings YL, Gauthier JP, Gambatese MA, Ko JY, et al. Validation of obstetric estimate of gestational age on US birth certificates. Am J Obstet Gynecol. 2014;210:335.e1–335.e5.
38. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008;371:75–84.
39. Sanders AP, Desrosiers TA, Warren JL, Herring AH, Enright D, Olshan AF, et al. Association between arsenic, cadmium, manganese, and lead levels in private wells and birth defects prevalence in North Carolina: a semi-ecologic study. BMC Public Health. 2014;14:955.
40. Alam DS. Prevention of low birthweight. Nestle Nutr Workshop Ser Pediatr Program. 2009;63:209–21. discussion 221-225, 259–68
41. Ashton DM, Lawrence HC, Adams NL, Fleischman AR. Surgeon general's conference on the prevention of preterm birth. Obstet Gynecol. 2009;113:925–30.
42. Institute of Medicine. Preterm birth: causes, consequences, and prevention [Internet]. Washington, DC: The National Academies Press; 2007. https://www.nap.edu/catalog/11622/preterm-birth-causes-consequences-and-prevention.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the two EPA technical reviewers who read this manuscript and provided helpful feedback. CK was supported in part by training grant from the National Institute of Environmental Health Sciences [T32ES007018].

AUTHOR CONTRIBUTIONS

CK, TD, and AO conceived and designed the study. NF and TL developed the administrative cohort and specified relevant covariate and outcome variables. CK conducted the analysis. CK, TD, AO, and TL interpreted the results. CK prepared the draft manuscript. TD, AO, TL, and NF reviewed the manuscript and provided substantial feedback. All authors reviewed the results and approved the final version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICAL APPROVAL

This study protocol was approved by the Institutional Review Boards of the University of North Carolina at Chapel Hill UNC (IRB Number: 09-0828) and the NC Division of Public Health. The EPA's Human Subjects Research Officer also reviewed and approved this work (HSR-001254).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41370-022-00475-0.

Correspondence and requests for materials should be addressed to Corinna Keeler.