Journal of Exposure Science and Environmental Epidemiology (2017) 27, 513–520

www.nature.com/jes

ORIGINAL ARTICLE

Measurement error in mobile source air pollution exposure estimates due to residential mobility during pregnancy

Audrey Flak Pennington[1], Matthew J. Strickland[2], Mitchel Klein[1], Xinxin Zhai[3], Armistead G. Russell[3], Craig Hansen[4,5] and Lyndsey A. Darrow[2]

Prenatal air pollution exposure is frequently estimated using maternal residential location at the time of delivery as a proxy for residence during pregnancy. We describe residential mobility during pregnancy among 19,951 children from the Kaiser Air Pollution and Pediatric Asthma Study, quantify measurement error in spatially resolved estimates of prenatal exposure to mobile source fine particulate matter ($PM_{2.5}$) due to ignoring this mobility, and simulate the impact of this error on estimates of epidemiologic associations. Two exposure estimates were compared, one calculated using complete residential histories during pregnancy (weighted average based on time spent at each address) and the second calculated using only residence at birth. Estimates were computed using annual averages of primary $PM_{2.5}$ from traffic emissions modeled using a Research LINE-source dispersion model for near-surface releases (RLINE) at 250 m resolution. In this cohort, 18.6% of children were born to mothers who moved at least once during pregnancy. Mobile source $PM_{2.5}$ exposure estimates calculated using complete residential histories during pregnancy and only residence at birth were highly correlated ($r_S > 0.9$). Simulations indicated that ignoring residential mobility resulted in modest bias of epidemiologic associations toward the null, but varied by maternal characteristics and prenatal exposure windows of interest (ranging from −2% to −10% bias).

Journal of Exposure Science and Environmental Epidemiology (2017) **27,** 513–520; doi:10.1038/jes.2016.66; published online 14 December 2016

Keywords: empirical/statistical models; epidemiology; particulate matter

INTRODUCTION

Residential mobility is common during pregnancy; in the United States, it is estimated that between 11% and 32% of pregnant women change residences at least once between conception and delivery.[1–7] Understanding mobility patterns during pregnancy is important for the design and interpretation of studies examining spatially varying environmental exposures during pregnancy. Although several studies have examined prenatal residential movement, there is room for further understanding of this topic. Limitations of previous research, noted in a review article by Bell and Belanger[8] in 2012, include the use of retrospectively collected and incomplete residence data, lack of information on detailed relocation information by demographic factors, and the use of populations that limit generalizability of results.

Prenatal air pollution exposure is frequently estimated using maternal residential location at the time of delivery as a proxy for residence during the entire gestational period.[9–12] This practice of not accounting for residential mobility (usually because of the lack of longitudinal residence information) can result in exposure measurement error and has the potential to bias resulting estimates of health effects. Previous studies examining residential mobility found relatively high agreement between prenatal air pollution exposure estimates calculated using this method and estimates using complete residential history data.[1,3,13] One of these studies found little impact of not accounting for this mobility on effect estimates.[1] However, the geographic resolution of assigned air pollution exposure in these studies varied substantially, ranging from 1 to 19,968 square km, and the spatial resolution of pollutant concentrations is a major determinant of the impact of residential mobility on assigned exposure. For example, if most residential changes during pregnancy involve moves < 5 km, and air pollution exposure is assigned at a 10 km resolution, residential mobility will likely have little impact on assigned pollution concentrations. A recent study by Brokamp *et al.*[14] examined the impact of residential mobility on estimates of traffic-related air pollution at a high spatial resolution in childhood, but not during pregnancy. There is no literature reporting the impact of residential mobility during pregnancy at a spatial resolution that would capture fine-scale variation in pollution from mobile sources.

Therefore, we describe residential mobility during pregnancy using prospectively collected residential history data from a large cohort of Health Maintenance Organization (HMO) members in the Southeastern United States. We also quantify measurement error attributable to using maternal residence only at the time of delivery to estimate average prenatal exposure to primary fine particulate matter ($PM_{2.5}$) from mobile sources, modeled at a 250 m grid resolution, and simulate the impacts of this error on estimates of epidemiologic associations by pregnancy trimester.

[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [2]School of Community Health Sciences, University of Nevada Reno, Reno, Nevada, USA; [3]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA; [4]Kaiser Permanente Georgia Center for Clinical and Outcomes Research, Atlanta, Georgia, USA and [5]South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia. Correspondence: Dr. AF Pennington, Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Mailstop 1518-002-2BB, Atlanta, GA 30332-4201, USA. Tel.: +1 404 712 6841. Fax: +1 404 727 8744.
E-mail: aflak@emory.edu
Received 7 July 2016; revised 9 September 2016; accepted 25 October 2016; published online 14 December 2016

METHODS

The Kaiser Air Pollution and Pediatric Asthma Study (KAPPA) is a historical birth cohort of 24,608 children born between 2000 and 2010 and enrolled in Kaiser Permanente Georgia (KPGA) HMO for at least the first year of life. Emory University and KPGA institutional review boards approved this study. This analysis was completed among a subset of 19,951 children from the KAPPA Study. Children were excluded from the analysis if they were not linked to mothers who were also enrolled in KPGA ($n = 2817$), if their mothers did not have residential history information available for pregnancy ($n = 758$), or if mothers resided outside the metropolitan Atlanta region for which air quality estimates were available at any point during pregnancy ($n = 631$). Because the KAPPA study was originally developed to examine the effect of exposure to air pollution in the first year of life, we also excluded 451 children without estimates of residential air pollution exposure during the first year of life. This was an administrative decision for consistency with future publication on this cohort. Residential mobility during pregnancy in this cohort was defined using data from KPGA medical records and Georgia birth certificates. For each pregnancy, conception date was estimated using gestational week information from the birth certificate. For the 2909 children without gestational age data, the start of the prenatal period was defined as 38 weeks before the date of birth (assuming a full-term gestational age of 40 weeks). For both calculations it was assumed that conception occurred at day 14, per obstetric convention. All children with prenatal residence information were included in our analyses, including those who had siblings in the cohort or for whom residence data were not contiguous. We completed sensitivity analyses excluding 1468 children whose mothers had ≥ 90 days of missing residence data during pregnancy.

We describe patterns of prenatal residential mobility among this cohort by calculating the percent of children born to mothers who changed residential locations during pregnancy. We classified mobility by season, pregnancy trimester, individual characteristics, and neighborhood socioeconomic status (SES). Neighborhood SES was determined at census block group spatial resolution using maternal residence at the time of delivery and novel demographic clusters created by Georgia Department of Public Health. These clusters classify block groups using variables from the 2010 US Census on factors such as age, income, housing, and employment.[15] Among women who changed residence during pregnancy, we examined the number of, and distances between, residential locations, and we compared air pollution concentrations between residences at conception and birth.

Average annual concentrations of $PM_{2.5}$ contributed by primary mobile sources were modeled at 250 m spatial resolution for years 2002 to 2010. A Research LINE-source dispersion model for near-surface releases (RLINE) was used to model hourly concentrations of mobile source contributed $PM_{2.5}$ using data on mobile source emissions and meteorology as inputs.[16] These estimates were then averaged to create one estimate for each year that were used in our analyses. Emissions inputs for 2010 were created by Atlanta Regional Commission's Atlanta Roadside Emissions Exposure Study (AREES) using data on traffic patterns and composition, mobile emissions, and meteorology for the 20 county metropolitan Atlanta area.[17] As the road network has not changed substantially over the study period, these 2010 data were used as an input for the 2002–2010 RLINE models, scaling each year by annual average emissions. Meteorological data were available from the meteorological processors of the American Meteorological Society (AMS) and US Environmental Protection Agency (EPA) Regulatory Model (AERMOD),[18] for 2002 to 2010 at hourly resolution for a monitor at Hartsfield-Jackson Atlanta International Airport, assumed to represent the whole spatial domain. Because RLINE results were found to overestimate spatial gradients compared with observations, estimates were calibrated to observation-based mobile source impacts from three stationary air pollution monitors in metropolitan Atlanta estimated by a chemical mass balance model.[19] Additional information about the air pollution modeling for this work is included in Supplementary Information online.

The modeled air pollution and residential history information from KPGA administrative records were used to estimate average mobile source primary $PM_{2.5}$ exposure during pregnancy. The 2010 $PM_{2.5}$ estimates are shown in Figure 1; the spatial pattern was nearly identical for years 2002–2009, although there was a temporal trend, with concentrations decreasing over time. Given the consistency of the spatial patterns of mobile source pollution and because prenatal periods for children in our cohort began in 1999 and the earliest available $PM_{2.5}$ estimates were for 2002, 2002 data were used to estimate prenatal exposures in 1999–2001. For each pregnancy the following exposures were calculated for the entire gestational period and each pregnancy trimester: (1) exposure calculated from the annual average concentrations using complete residential histories as a weighted average based on time residing at each address, and (2) exposure calculated from the annual average concentrations using only maternal residence at the time of delivery (commonly implemented in studies without available residential histories in the prenatal period). For brevity we will refer to the first estimate accounting for mobility as the "complete exposure" and the second estimate not accounting for mobility as the "naive exposure". Differences between complete and naive exposure estimates are solely because of spatial differences in pollution. For example, for a pregnancy that started in 2003 and ended in 2004, both exposure estimates take into account pollution from 2003 and 2004. The only difference in estimates is that the complete exposure is a weighted average of all residential locations for the time period, whereas the naive exposure only uses the location at the time of delivery.

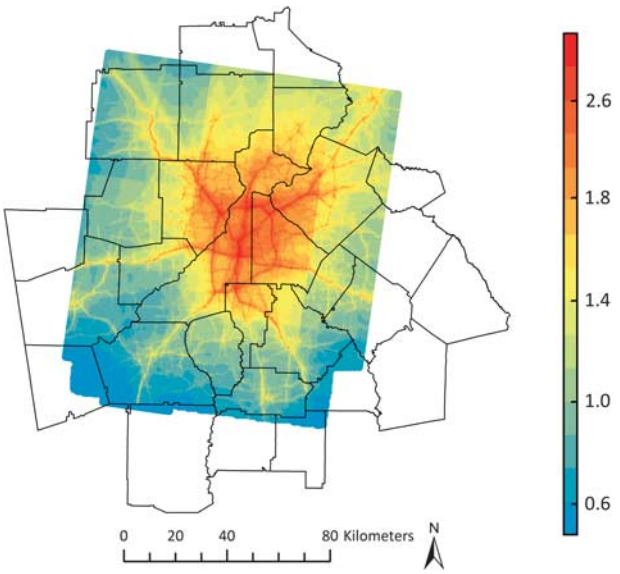

Figure 1. 2010 RLINE-modeled primary mobile source $PM_{2.5}$ ($\mu g/m^3$).

We simulated the impact of not accounting for residential mobility in this HMO cohort when estimating exposure on an expected association between prenatal $PM_{2.5}$ and a hypothetical disease. We assessed whether the magnitude of bias varied for different specified effects: risk differences of 0.01, 0.05, and 0.10 and risk ratios of 1.05, 1.1, and 1.2 for an increase of 1 $\mu g/m^3$ of prenatal $PM_{2.5}$ exposure from primary mobile source emissions. Simulations were performed using the following steps. (1) Calculate probability of disease for each child in our sample using a baseline risk of 10%, the specified effect of exposure, and the child's prenatal $PM_{2.5}$ exposure estimate ("complete exposure"). (2) Randomly generate outcome status (yes/no) for each child by using the probability from step 1 to represent a binomial probability parameter. (3) Run two binomial linear regression models predicting the outcome generated in step 2, one using the complete prenatal exposure as the predictor (to ensure it yielded results close to the specified effect) and the other using the naive exposure as the predictor. (4) Repeat steps 2 and 3 100,000 times. (5) Summarize results of each set of 100,000 simulations using the median of the resulting parameter estimates and estimate the bias of the effect due to using naive exposure estimates (calculated for risk differences as $(RD_{naive} - RD_{specified})/RD_{specified}$ and for risk ratios as bias of excess risk $((RR_{naive} - 1) - (RR_{specified} - 1))/(RR_{specified} - 1)$). We chose 100,000 iterations for the simulation so that replicating the process would produce essentially the same results. Additional simulations were completed focusing on trimester-specific exposure and stratifying by race and other maternal and child factors. Trimester-specific exposure estimates were calculated taking into account the trimester start and stop dates; the complete estimate was a time-weighted average of the annual average concentrations at all residences during the trimester and the naive estimate used only residence at the time of delivery. Analyses were completed in SAS 9.3 (SAS Institute, Cary, NC, USA) and R 3.1,[20] maps were created in ArcMap 10.1 by ESRI. Simulation code is available from the authors upon request.

Table 1. KAPPA cohort characteristics, impact of residential mobility on prenatal mobile source $PM_{2.5}$ exposure estimates.

	Children in cohort, n (% of total)	*Children whose mothers moved in pregnancy, n (% of row)*	*Median move distance*	*Spearman's correlation between complete and naive $PM_{2.5}$ exposure estimates*[a]	
				All children	*Children whose mothers moved in pregnancy*
Cohort	19,951	3709 (18.6)	10 km	0.96	0.80
Maternal race					
Black	7157 (35.9)	1609 (22.5)	10 km	0.94	0.76
White	8757 (43.9)	1295 (14.8)	9 km	0.97	0.82
Other[b]	2186 (11.0)	364 (16.7)	8 km	0.97	0.80
Unknown Race	1851 (9.3)	441 (23.8)	8 km	0.95	0.82
Maternal education					
< 12th grade	280 (1.4)	59 (21.1)	6 km	0.97	0.89
High School/GED	2524 (12.7)	480 (19.0)	9 km	0.97	0.82
Some college or more	13,113 (65.7)	2265 (17.3)	10 km	0.96	0.79
Missing	4034 (20.2)	905 (22.4)	9 km	0.95	0.78
Maternal age					
< 25	1763 (8.8)	536 (30.4)	9 km	0.93	0.77
25– < 30	5759 (28.9)	1259 (21.9)	10 km	0.95	0.80
30– < 35	7364 (36.9)	1245 (16.9)	10 km	0.96	0.80
35– < 40	4153 (20.8)	579 (13.9)	8 km	0.97	0.78
≥ 40	912 (4.6)	90 (9.9)	6 km	0.99	0.94
Maternal marital status					
Married	15,279 (76.6)	2517 (16.5)	10 km	0.97	0.81
Not married	1762 (8.8)	499 (28.3)	10 km	0.93	0.74
Missing	2910 (14.6)	693 (23.8)	9 km	0.94	0.77
Maternal neighborhood socioeconomic status(SES)[c]					
Highest SES	12,569 (63.0)	2118 (16.9)	11 km	0.95	0.75
Urban/suburban	1950 (9.8)	416 (21.3)	6 km	0.91	0.60
Rural, average to low SES	963 (4.8)	169 (17.6)	9 km	0.96	0.79
Lowest SES	4468 (22.4)	1006 (22.5)	8 km	0.95	0.80
Child birth year					
2000	2054 (10.3)	403 (19.6)	9 km	0.96	0.79
2001	1977 (9.9)	442 (22.4)	9 km	0.94	0.75
2002	1946 (9.8)	386 (19.8)	9 km	0.96	0.81
2003	1929 (9.7)	403 (20.9)	10 km	0.95	0.80
2004	1871 (9.4)	336 (18.0)	12 km	0.95	0.75
2005	1741 (8.7)	324 (18.6)	9 km	0.96	0.77
2006	1935 (9.7)	323 (16.7)	11 km	0.96	0.77
2007	1919 (9.6)	333 (17.4)	10 km	0.96	0.78
2008	1835 (9.2)	313 (17.1)	9 km	0.97	0.85
2009	1403 (7.0)	219 (15.6)	9 km	0.97	0.84
2010	1341 (6.7)	227 (16.9)	9 km	0.97	0.81

Among each characteristic, *P*-values for Pearson's χ^2 tests for proportion who move were < 0.01. [a]Complete exposure estimates are calculated as a weighted average of time spent at each residence during the prenatal period, and naive exposure estimates are calculated assuming residence at birth applied to entire prenatal period. [b]Includes Asian, American Indian, Alaska Native, Native Hawaiian, or other Pacific Islander, and mothers identifying with more than one racial group. [c]Neighborhood socioeconomic status was classified at census block group spatial resolution using maternal residence at time of delivery and described using demographic clusters created by the Georgia Department of Public Health.[15]

RESULTS

In this HMO cohort, 18.6% of children were born to mothers who changed residence at least once during pregnancy (Table 1). Women of black race were more likely to move during pregnancy than women of white race (22.5% *vs* 14.8%). Mobility decreased with increasing maternal age and education. For example, 21.1% of mothers who did not complete high school moved during pregnancy compared with 17.3% of mothers who attended at least some college. Across levels of neighborhood SES, mothers with the lowest SES had the most mobility and mothers with the highest SES had the least mobility (Table 1). The distance moved ranged from < 1 to 106 km, with a mean move distance of 13 km and a median of 10 km. Median move distance varied among cohort subgroups, ranging from 6 to 12 km (Table 1). Although the majority of children whose mothers moved during pregnancy only moved once (84.1%), the number of moves during pregnancy ranged from 0 to 8 (Table 2). Compared with mothers who did not move during pregnancy, the 591 mothers who moved twice or more were more likely to be of black race (52.3% *vs* 34.2%) and less likely to have attended at least some college (56.9% *vs* 66.8%) or live in a neighborhood classified as having the highest SES (50.6% *vs* 64.4%) (all Pearson's χ^2 test *P*-values < 0.01). Moves were equally likely to occur in each pregnancy trimester (Table 2). Examining moves by season, moves were slightly more likely to

Table 2. Residential mobility during pregnancy by trimester and season.

	n (% of 19,951 pregnancies)
Number of moves during pregnancy	
0	16,242 (81.4)
1	3118 (15.6)
2	469 (2.4)
3+	122 (0.6)
Mobility by pregnancy trimester	
Moved in first trimester	1396 (7.0)
Moved in second trimester	1414 (7.1)
Moved in third trimester	1407 (7.1)
Mobility by season	
Moved in winter	1001 (5.0)
Moved in spring	1026 (5.1)
Moved in summer	1165 (5.8)
Moved in fall	1050 (5.3)

The mobility by pregnancy trimester and season sections count the number of children whose mothers moved during each trimester and season. The totals of these two sections are not equivalent because of the event of multiple moves by one mother occurring during the same trimester or season.

occur during summer months than in winter, spring, or fall months. Results were similar in a sensitivity analysis excluding children whose mothers had ≥ 90 days of missing residence data during pregnancy.

The spatial distribution of primary $PM_{2.5}$ closely mirrored the road network, with concentrations highest inside the I-285 highway encircling metropolitan Atlanta and decreasing with increasing distance from the city center (Figure 1). Figure 2 presents the distribution of $PM_{2.5}$ exposure estimates during the full pregnancy accounting for mobility ("complete exposure"); exposure estimates ranged from 0.49 to 5.59 $\mu g/m^3$ with a mean exposure of 1.77 $\mu g/m^3$. This represents exposure solely to primary mobile source $PM_{2.5}$ and does not include exposure to secondary $PM_{2.5}$ such as sulfates. A change of 1 $\mu g/m^3$, the quantity we used for scaling risk differences and risk ratios in the simulation, represents a change from the 10th to the 89th percentile of the exposure distribution. Average prenatal $PM_{2.5}$ exposure estimates calculated without accounting for mobility ("naive exposure") were at most 2.32 $\mu g/m^3$ different than the complete exposure estimates, with a mean difference of 0.03 $\mu g/m^3$ (e.g., equivalent of a change from the 50th to the 53rd percentile of the exposure distribution). Spearman's correlation coefficients between complete and naive exposure estimates were 0.96 for estimates of exposure during the entire pregnancy, and 0.92, 0.95, and 0.99 for first, second, and third trimester exposure, respectively. Because residential mobility varied by demographic characteristics, correlation between exposure estimates also varied in our sample from 0.91 to 0.99 among all children and 0.60 to 0.94 among children whose mothers moved during pregnancy (Table 1). In order to assess whether mothers who moved during pregnancy moved to higher or lower pollution areas, we examined differences in decile of $PM_{2.5}$ exposure at conception and birth residential locations (Table 3). For this table, exposure deciles were based on only children whose mothers moved during pregnancy, and hence all rows and columns sum to 10%. In general, we found that mothers who moved resided in similar deciles of exposure at the two time points.

Table 4 presents the results of simulations on the expected bias caused by exposure measurement error due to not accounting for residential mobility during pregnancy. Overall, the magnitude of the bias of the association between prenatal $PM_{2.5}$ exposure and a hypothetical outcome was modest and resulted in effect estimates

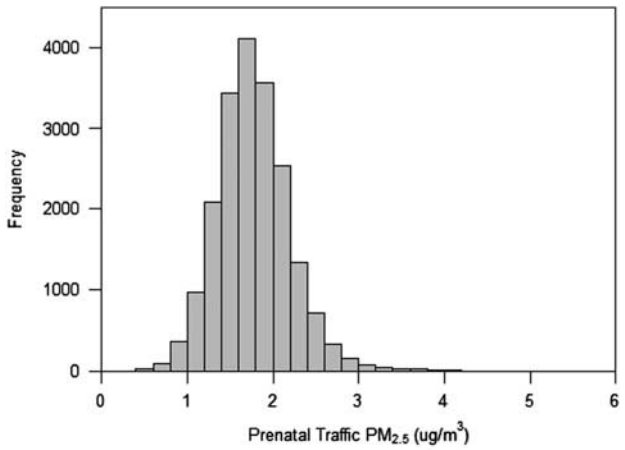

Figure 2. Distribution of prenatal mobile source $PM_{2.5}$ exposure accounting for complete residential history ($n = 19,951$).

closer to the null than the specified effects. For example, examining $PM_{2.5}$ exposure during the entire pregnancy for all children, with a specified risk difference of 0.05, the median risk difference for 100,000 simulations using the complete exposure was 0.0500, and the median risk difference for the naive exposure was 0.0476. When increasing the specified risk difference to 0.10, the complete exposure resulted in a median risk difference of 0.1000 and the naive exposure resulted in a median risk difference of 0.0952. Figure 3a displays the risk differences resulting from the 200,000 binomial linear regression models completed with a specified risk difference of 0.10 (100,000 for complete exposure (gray) and 100,000 for naive exposure (blue)). The distributions of risk differences are similar, with the one resulting from naive exposure shifted closer to the null value of 0. Increasing the specified risk difference resulted in an increase of the absolute difference between median estimates from naive and complete exposure, ranging from 0.0005 for a risk difference of 0.01 to 0.0048 for a risk difference of 0.10, but did not impact the percent bias. Patterns were similar using the risk ratio as the measure of association of interest (Table 4 and Figure 3b). When stratifying by race, the magnitude of the bias was larger among children born to black mothers than children born to white mothers because of their differential rates of residential mobility (−8% to − 10% bias *vs* − 3% to − 4% bias depending on specified effect). Similarly, when examining trimester-specific exposures, bias was greatest for first trimester associations because of the greater cumulative residential mobility between the start of the trimester and delivery. For all results in Table 4, the underestimation of the risk difference or risk ratio because of not accounting for residential mobility ranged from − 2% to a − 10% bias in the median effect estimate.

To further explore the variability in bias attributable to residential mobility, we completed the simulation for two additional subgroups of the cohort: (1) children born to black mothers who were < 30 years old, living in neighborhoods classified as having the lowest SES ($n = 1157$), and (2) children born to white mothers who were > 30 years old, living in neighborhoods classified as having the highest SES ($n = 4028$). These subgroups were chosen because of their contrasting mobility rates during pregnancy, 30.7% and 12.0%, respectively. A specified risk difference of 0.10 was used for both groups. In the high mobility group, the median estimated risk difference resulting from the complete exposure was 0.0999 and the median risk difference resulting from the naive exposure was 0.0890 (−11% bias). In the low mobility group, the median risk differences were 0.1000 when using the complete exposure and 0.0974 when using the naive exposure (−3% bias). The discrepancy in results between the two groups was larger when examining first

Table 3. Comparison of mobile source $PM_{2.5}$ exposure decile at the residential location at conception and birth, among children whose mothers changed residences during pregnancy ($n = 3709$).

% of 3,709 children		Decile of exposure at residential location at conception									
		1	2	3	4	5	6	7	8	9	10
Decile of exposure at residential location at birth	1	4.9%	1.4%	0.9%	0.8%	0.5%	0.4%	0.2%	0.2%	0.2%	0.4%
	2	1.8%	2.4%	1.4%	1.3%	0.8%	0.7%	0.4%	0.5%	0.4%	0.4%
	3	1.0%	2.6%	1.6%	0.9%	0.8%	0.8%	0.6%	0.8%	0.5%	0.4%
	4	0.6%	1.2%	2.1%	1.2%	0.8%	1.0%	0.9%	0.8%	0.9%	0.5%
	5	0.3%	0.7%	1.4%	1.7%	1.3%	1.2%	1.1%	0.9%	0.8%	0.6%
	6	0.5%	0.4%	0.8%	1.7%	2.2%	1.1%	1.0%	0.7%	0.8%	0.8%
	7	0.2%	0.5%	0.5%	1.0%	1.6%	1.5%	1.4%	1.1%	1.1%	1.2%
	8	0.2%	0.3%	0.5%	0.5%	0.8%	1.8%	1.7%	1.8%	1.2%	1.2%
	9	0.3%	0.2%	0.2%	0.4%	0.6%	0.9%	1.8%	2.0%	2.1%	1.5%
	10	0.1%	0.3%	0.5%	0.5%	0.6%	0.7%	0.9%	1.3%	2.1%	3.0%

Color intensity increases with percent (lightest: 0.5-0.9%, middle: 1-1.9%, darkest: ≥2%). Outlined boxes denote children whose exposure decile was the same at their conception and birth addresses. Deciles were based on 3709 children whose mothers moved during pregnancy, and hence all columns and rows sum to 10%. Conception residence decile cutpoints ($\mu g/m^3$): 1.38, 1.56, 1.69, 1.81, 1.93, 2.04, 2.15, 2.31, 2.53, and 4.32. Birth residence decile cutpoints ($\mu g/m^3$): 1.27, 1.43, 1.55, 1.66, 1.76, 1.87, 1.99, 2.13, 2.35, and 4.08.

Table 4. Results of simulation modeling the effect of prenatal mobile source $PM_{2.5}$ using exposure accounting for and not accounting for residential mobility.

Risk differences

	Percent mobility[a]	*Specified RD = 0.01* RD_C/RD_N/% Bias	*Specified RD = 0.05* RD_C/RD_N/% Bias	*Specified RD = 0.10* RD_C/RD_N/% Bias
Exposure in entire pregnancy				
All children	18.6	0.0100/0.0095/−5%	0.0500/0.0476/−5%	0.1000/0.0952/−5%
Children of black mothers	22.5	0.0100/0.0091/−9%	0.0501/0.0458/−8%	0.1001/0.0915/−9%
Children of white mothers	14.8	0.0100/0.0097/−3%	0.0500/0.0486/−3%	0.1001/0.0972/−3%
Exposure by trimester				
Trimester 1	18.6	0.0100/0.0092/−8%	0.0500/0.0463/−7%	0.1000/0.0927/−7%
Trimester 2	13.2	0.0100/0.0095/−5%	0.0500/0.0476/−5%	0.1000/0.0952/−5%
Trimester 3	7.1	0.0100/0.0098/−2%	0.0499/0.0490/−2%	0.1000/0.0981/−2%

Risk ratios

	Percent mobility[a]	*Specified RR = 1.05* RR_C/RR_N/% Bias	*Specified RR = 1.1* RR_C/RR_N/% Bias	*Specified RR = 1.2* RR_C/RR_N/% Bias
Exposure in entire pregnancy				
All children	18.6	1.0498/1.0471/−6%	1.0999/1.0944/−6%	1.1999/1.1880/−6%
Children of black mothers	22.5	1.0498/1.0448/−10%	1.0998/1.0899/−10%	1.2001/1.1791/−10%
Children of white mothers	14.8	1.0498/1.0482/−4%	1.0997/1.0967/−3%	1.2000/1.1930/−4%
Exposure by trimester				
Trimester 1	18.6	1.0498/1.0457/−9%	1.0999/1.0915/−9%	1.1998/1.1820/−9%
Trimester 2	13.2	1.0498/1.0471/−6%	1.0998/1.0943/−6%	1.1999/1.1880/−6%
Trimester 3	7.1	1.0497/1.0488/−2%	1.0998/1.0977/−2%	1.2000/1.1950/−3%

Abbreviations: RD, risk difference; RD_C, median risk difference calculated from complete exposure; RD_N, median risk difference from naive exposure; RR, risk ratio; RR_C, median risk ratio from complete exposure; RR_N, median risk ratio from naive exposure; % Bias, calculated as $(RD_N - RD_{specified})/RD_{specified}$ for risk differences and calculated as bias of excess risk for risk ratios $((RR_N - 1) - (RR_{specified} - 1))/(RR_{specified} - 1)$. Complete exposure estimates are calculated as a weighted average of time spent at each residence during the prenatal period, and naive exposure estimates are calculated assuming residence at birth applied to entire prenatal period. [a]For trimester-specific rates, calculated as cumulative mobility between start of trimester and delivery.

trimester exposure, with −19% bias in the high mobility group and −4% bias in the low mobility group (median risk differences from complete and naive exposure: high mobility group 0.1000 *vs* 0.0811; low mobility group 0.1001 *vs* 0.0960).

DISCUSSION

In this paper we explore residential mobility during pregnancy in an HMO cohort and (1) describe its impact on estimates of exposure to primary mobile source $PM_{2.5}$ and (2) estimate the

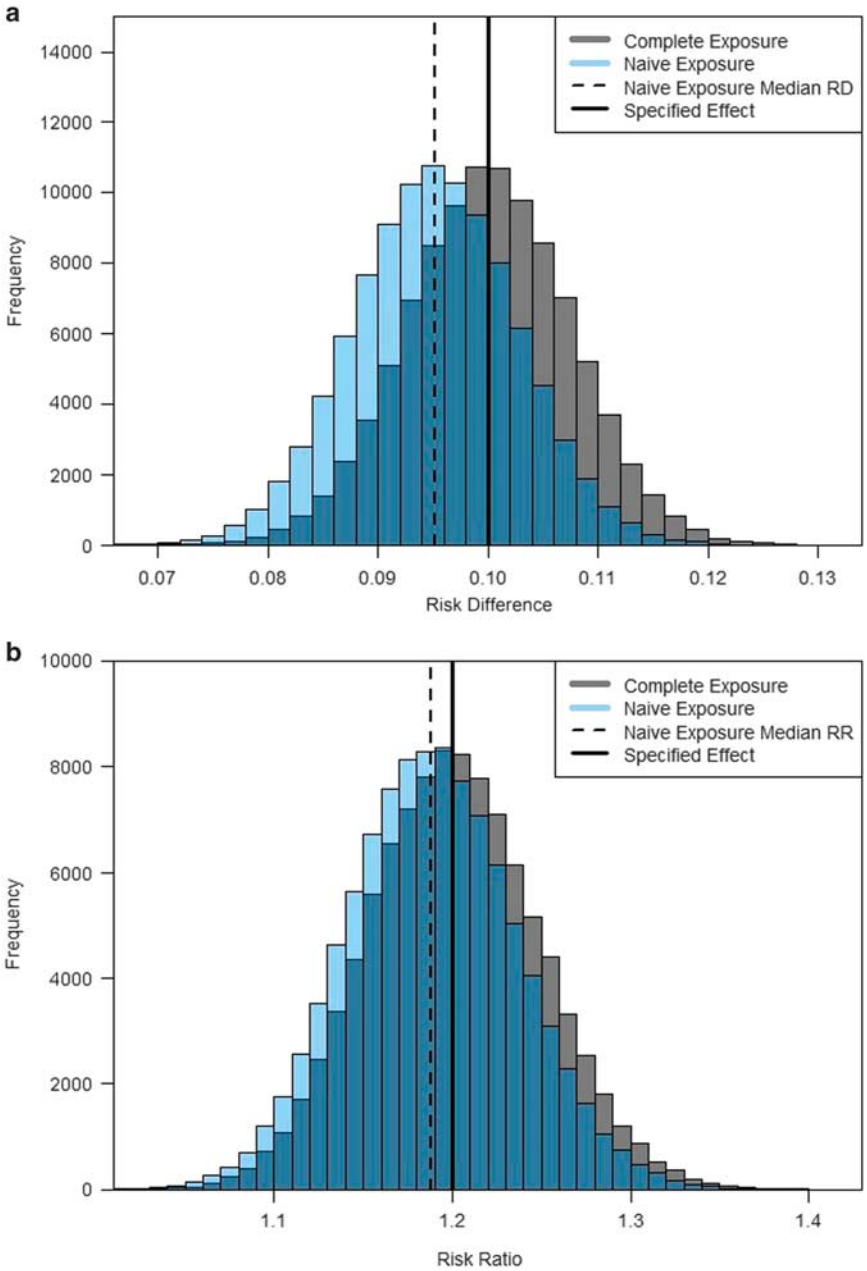

Figure 3. (**a**) Risk differences from simulation with 100,000 replications using complete and naive exposure estimates for all children (specified risk difference = 0.10). (**b**) Risk ratios from simulation with 100,000 replications using complete and naive exposure estimates for all children (specified risk ratio = 1.2). Complete exposure estimates are calculated as a weighted average of time spent at each residence during the prenatal period, and naive exposure estimates are calculated assuming residence at birth applied to entire prenatal period.

expected bias in epidemiologic associations because of not accounting for this residential mobility. In this cohort, 18.6% of women moved between conception and delivery which was within the range of mobility estimates from previous US studies. Unlike previous studies, which have all found mobility is more likely during the second trimester,[3,4,7] we found that moving was equally common throughout pregnancy. Examining mobility by demographic characteristics, our finding of higher mobility among women who are younger, not married, and have indicators of lower SES replicates findings of several previous studies.[8] One of the strongest predictors of mobility in this cohort was race; 22.5% of women of black race moved during pregnancy compared with only 14.8% of women of white race. Unlike SES, age, and marital status, results from previous studies have found inconsistent mobility patterns by race.[8]

The prenatal move distances (with a median of 10 km) are likely a lower bound of the move distances of all mothers enrolled in Kaiser Permanente Georgia HMO. Our estimates excluded moves by mothers who left KPGA during pregnancy, resided outside of the air pollution modeling region at any time during pregnancy, or whose children lacked first year of life exposure estimates. If we reexamine move distances including all women for whom we have residence data (i.e., including those outside the metropolitan Atlanta area), the calculated median move distance does not change, but the mean move distance is 2 km greater (15 *vs* 13 km). Consequently, we would not expect large moves among women

excluded from our estimate to change the distribution of move distances substantially. The median move distance in this cohort, 10 km, is larger than those calculated in three previous US studies whose median estimates ranged from 4.2 to 6.9 km.[2,3,7] Our study takes place in metropolitan Atlanta, a large urban area with considerable sprawl that includes approximately 22000 square km. Compared with many other metropolitan areas in the United States, a woman in Atlanta can move longer distances and still reside in same metropolitan area. This may be one explanation for the longer move distances in this cohort. The between-study variability in distances moved during pregnancy suggests that move distances depend on both the population studied and the patterns of sprawl in the geographical region of residence.

In this HMO cohort there was high correlation between estimates of prenatal $PM_{2.5}$ exposure calculated accounting and not accounting for residential mobility. Although this is expected as the vast majority of women who do not move during pregnancy have perfect correlation of the two measures, correlations were high even when restricting the sample to only women who moved during pregnancy. In the simulation, we found that not accounting for residential mobility resulted in modest bias of epidemiologic associations, even in groups with a mobility rate as high as 30.7%. Bias was largest when examining the impact of first trimester exposure as one would expect due to the greater amount of time between the first trimester and birth and thus opportunity for a different residence to contribute more time to the weighted average exposure estimate. Although impact is expected to vary by population, overall these results are promising for studies that lack information on residential mobility during pregnancy. However, the result that the magnitude of bias in exposure estimates varied across cohort subgroups because of variation in mobility rates is noteworthy. In the simulation completed among some of the highest and lowest mobility groups in the cohort (30.7% *vs* 12.0%) where using the complete exposure resulted in a median risk difference of ~0.10 in both groups, the resulting median risk difference from the naive exposure was 0.0890 in the first group and 0.0974 in the second group. The effect estimates in these two groups differed solely because of exposure measurement error. In a study where prenatal exposure is calculated without accounting for residential mobility, such discrepant results could be misinterpreted as evidence of effect measure modification if researchers were unaware of the differential measurement error in these two groups. Although in this study the differences in bias between subgroups are modest, we note that our study population is a fully insured cohort with a narrower range of SES than populations outside of an HMO setting. For example, >65% of children in our cohort were born to mothers who attended at least some college. In populations with more socioeconomic diversity, the differences in residential mobility and resulting impact on bias could be larger.

The results of our simulations are dependent on many factors such as the baseline risk of the outcome (10%), the mobility rates in the cohort, the spatial distribution of $PM_{2.5}$, and the specified effect investigated. We assumed non-differential mobility rates by outcome; a study of a specific disease should consider whether mobility could be differential with respect to the outcome. We examined how the magnitude of bias varied based on mobility rates, by completing stratified simulations and with different specified measures of both additive and relative effects. Almost identical magnitudes of bias in the simulation were observed when lowering the baseline risk of disease to 0.05% (results not shown) suggesting that these estimates of bias would be relevant to diseases with different prevalences. The increasing bias with increasing mobility rates, as well as other factors dictating magnitude of bias, have previously been discussed by two related simulation studies.[21,22] Our results would change dramatically if exposure was assigned at a different spatial resolution. In this cohort with a median move distance of 10 km, if exposure was assigned at a 20 km spatial resolution, instead of a 250 m spatial resolution, there would be minimal differences between exposure estimates accounting and not accounting for mobility and subsequently even less bias observed in the simulation.

The residence data used for this analysis come from KPGA administrative records that are prospectively collected and include addresses and dates of residence. Administrative data have limitations. Residence information is updated in the KPGA system when the HMO is notified by a member of a new address. There were likely some changes of address that were not reported to KPGA, or for which there was uncertainty about when addresses changed, as evidenced by gaps in residence data for some women. Our residences were geocoded at a 250 m grid resolution. If a mother moved to a new residence within the same 250 m grid as her current residence, then we would be unaware that she changed residences. Although such short distance moves are likely to be rare, our inability to track within-grid movement may have contributed to a slight underestimation of the proportion of women who moved in this cohort. In addition, residences are stored in the KPGA system at monthly, not daily, resolution that masks the exact start date of each residence. This challenge, which has been encountered by previous studies,[3,5] is of most concern for calculating mobility by trimester for which exact timing of changes in residence is important. Because of this imprecision, we did not conduct analyses related to the specific timing of maternal changes in residence (e.g., modeling timing of moves in a time-to-event analysis).

Primary air pollution from mobile sources is one component of total ambient air pollution that encompasses primary and secondary pollution from traffic and other sources. Our RLINE-based exposure model incorporates emissions and meteorology data and is calibrated using observation-based mobile source impacts. Although the incorporation of these factors is anticipated to increase model validity, we do not have estimates of exposure measurement error due to model error for each child in the cohort; it is possible that this source of error is larger than error due to residential mobility. Because of variation in spatial distributions of pollution, our results may not be applicable to estimates of total ambient exposure. Similarly, our study did not examine personal air pollution exposure that is affected by factors such as ambient pollution concentrations, indoor air pollution exposure, housing air exchange rates, and time–activity patterns. There is some evidence from the literature indicating high correlations between estimates of pollution exposures based on maternal residence alone and those incorporating information on maternal time–activity patterns.[23] Regardless, we do not expect the results of the study to reflect the impact of residential mobility on estimates of personal exposure to air pollution. Considering the population of this study, our results are most generalizable to studies of prenatal exposure completed in other insured HMO populations. Mobility rates are expected to differ by demographic characteristics, and based on the patterns of mobility observed in our study these may be higher in uninsured or lower SES populations where factors such as housing instability are more likely to influence residential mobility.

Understanding residential mobility during pregnancy is critical for research on the impact of environmental exposures during pregnancy. This study contributes to our knowledge by describing patterns of residential movement among a cohort of pregnant women and by estimating its impact on fine-scale estimates of one environmental exposure, primary mobile source $PM_{2.5}$. Overall, we observed a modest amount of bias in prenatal exposure estimates and expected epidemiologic associations due to not accounting for residential mobility during pregnancy. The estimated bias would have been smaller if we were interested in more spatially homogeneous exposures, for example, those that can be estimated accurately at the county level. The most bias was

seen in estimates of associations with first trimester exposure and estimates among subgroups of women with the highest levels of residential mobility. Our results show that in extreme situations when comparing results among groups with very different mobility rates, not accounting for residential mobility when estimating exposure can lead to results that look like effect measure modification. The results of this study provide some insight into the potential implications of not accounting for residential mobility during pregnancy and suggest that in the absence of these data future studies still have the potential to produce fairly reliable estimates of association.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Support for this work was provided by US EPA Grant R834799, NIH/NICHD Grant R03HD084884-01, and NIH Reproductive, Perinatal, & Pediatric Training Grant T32HD052460. This publication's contents are solely the responsibility of the grantee and do not necessarily represent the official view of the US EPA. Furthermore, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication.

REFERENCES

1 Pereira G, Bracken MB, Bell ML. Particulate air pollution, fetal growth and gestational length: the influence of residential mobility in pregnancy. *Environ Res* 2016; **147**: 269–274.
2 Saadeh FB, Clark MA, Rogers ML, Linkletter CD, Phipps MG, Padbury JF *et al*. Pregnant and moving: understanding residential mobility during pregnancy and in the first year of life using a prospective birth cohort. *Matern Child Health J* 2013; **17**: 330–343.
3 Chen L, Bell EM, Caton AR, Druschel CM, Lin S. Residential mobility during pregnancy and the potential for ambient air pollution exposure misclassification. *Environ Res* 2010; **110**: 162–168.
4 Miller A, Siffel C, Correa A. Residential mobility during pregnancy: patterns and correlates. *Matern Child Health J* 2010; **14**: 625–634.
5 Canfield MA, Ramadhani TA, Langlois PH, Waller DK. Residential mobility patterns and exposure misclassification in epidemiologic studies of birth defects. *J Expo Sci Environ Epidemiol* 2006; **16**: 538–543.
6 Zender R, Bachand AM, Reif JS. Exposure to tap water during pregnancy. *J Expo Anal Environ Epidemiol* 2001; **11**: 224–230.
7 Lupo PJ, Symanski E, Chan W, Mitchell LE, Waller DK, Canfield MA *et al*. Differences in exposure assignment between conception and delivery: the impact of maternal mobility. *Paediatr Perinat Epidemiol* 2010; **24**: 200–208.
8 Bell ML, Belanger K. Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. *J Expo Sci Environ Epidemiol* 2012; **22**: 429–438.
9 DeFranco E, Hall E, Hossain M, Chen A, Haynes EN, Jones D *et al*. Air pollution and stillbirth risk: exposure to airborne particulate matter during pregnancy is associated with fetal death. *PLoS One* 2015; **10**: e0120594.
10 Capobussi M, Tettamanti R, Marcolin L, Piovesan L, Bronzin S, Gattoni ME *et al*. Air pollution impact on pregnancy outcomes in Como, Italy. *J Occup Environ Med* 2016; **58**: 47–52.
11 Hannam K, McNamee R, Baker P, Sibley C, Agius R. Air pollution exposure and adverse pregnancy outcomes in a large UK birth cohort: use of a novel spatio-temporal modelling technique. *Scand J Work Environ Health* 2014; **40**: 518–530.
12 Lee PC, Roberts JM, Catov JM, Talbott EO, Ritz B. First trimester exposure to ambient air pollution, pregnancy complications and adverse birth outcomes in Allegheny County, PA. *Matern Child Health J* 2013; **17**: 545–555.
13 Hodgson S, Lurz PW, Shirley MD, Bythell M, Rankin J. Exposure misclassification due to residential mobility during pregnancy. *Int J Hyg Environ Health* 2015; **218**: 414–421.
14 Brokamp C, LeMasters GK, Ryan PH. Residential mobility impacts exposure assessment and community socioeconomic characteristics in longitudinal epidemiology studies. *J Expo Sci Environ Epidemiol* 2016; **26**: 428–434.
15 Georgia Department of Public Health, Office of Health Indicators for Planning (OHIP). Online Analytical Statistical Information System: Demographic Clusters of Georgia: Accessing the Georgia Department of Public Health's Data Warehouse. Available from https://oasis.state.ga.us/gis/demographiccluster/DemoClusters2011.htm. Accessed 11 April 2016.
16 Community Modeling and Analysis System. R-LINE: A Research LINE-source dispersion model for near-surface releases. Available from https://www.cmascenter.org/r-line/. Accessed 28 September 2015.
17 Atlanta Regional Commission. AREES - Near Road Emissions. Available from http://atlantaregional.com/environment/air/arees-near-road-emissions. Accessed 6 May 2016.
18 Cimorelli AJ, Perry SG, Venkatram A, Weil JC, Paine RJ, Wilson RB *et al*. AERMOD: a dispersion model for industrial source applications. Part I: general model formulation and boundary layer characterization. *J Appl Meteorol* 2005; **44**: 682–693.
19 U.S. Environmental Protection Agency. EPA-CMB8.2 Users Manual. Report No. EPA-452/R-04-011. December 2004. Research Triangle Park, NC.
20 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria. http://www.R-project.org. 2014.
21 Khoury MJ, Stewart W, Weinstein A, Panny S, Lindsay P, Eisenberg M. Residential mobility during pregnancy: implications for environmental teratogenesis. *J Clin Epidemiol* 1988; **41**: 15–20.
22 Schulman J, Selvin S, Shaw GM, Malcoe LH. Exposure misclassification due to residential mobility during pregnancy in epidemiologic investigations of congenital malformations. *Arch Environ Health* 1993; **48**: 114–119.
23 Iniguez C, Ballester F, Estarlich M, Llop S, Fernandez-Patier R, Aguirre-Alfaro A *et al*. Estimation of personal NO2 exposure in a cohort of pregnant women. *Sci Total Environ* 2009; **407**: 6093–6099.

Supplementary Information accompanies the paper on the Journal of Exposure Science and Environmental Epidemiology website (http://www.nature.com/jes)