Environmental Research 147 (2016) 108–114

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Spatio-temporal ozone variation in a case-crossover analysis of childhood asthma hospital visits in New York City

Jessie Loving Carr Shmool [a,*], Ellen Kinnee [a], Perry Elizabeth Sheffield [b], Jane Ellen Clougherty [a]

[a] University of Pittsburgh Graduate School of Public Health, Department of Environmental and Occupational Health, 100 Technology Drive, Ste. 350, Pittsburgh, PA 15219, USA
[b] Icahn School of Medicine at Mount Sinai, DPM, 1 Gustave L. Levy Pl., Box 1057, New York, NY 10029, USA

ARTICLE INFO

Article history:
Received 3 November 2015
Received in revised form
8 January 2016
Accepted 15 January 2016

Keywords:
Case-crossover
Childhood asthma
Intraurban variation
Ozone
Spatio-temporal

ABSTRACT

Background: Childhood asthma morbidity has been associated with short-term air pollution exposure. To date, most investigations have used time-series models, and it is not well understood how exposure misclassification arising from unmeasured spatial variation may impact epidemiological effect estimates. Here, we develop case-crossover models integrating temporal and spatial individual-level exposure information, toward reducing exposure misclassification in estimating associations between air pollution and child asthma exacerbations in New York City (NYC).

Methods: Air pollution data included: (a) highly spatially-resolved intra-urban concentration surfaces for ozone and co-pollutants (nitrogen dioxide and fine particulate matter) from the New York City Community Air Survey (NYCCAS), and (b) daily regulatory monitoring data. Case data included citywide hospital records for years 2005–2011 warm-season (June–August) asthma hospitalizations ($n=2353$) and Emergency Department (ED) visits ($n=11{,}719$) among children aged 5–17 years. Case residential locations were geocoded using a multi-step process to maximize positional accuracy and precision in near-residence exposure estimates. We used conditional logistic regression to model associations between ozone and child asthma exacerbations for lag days 0–6, adjusting for co-pollutant and temperature exposures. To evaluate the effect of increased exposure specificity through spatial air pollution information, we sequentially incorporated spatial variation into daily exposure estimates for ozone, temperature, and co-pollutants.

Results: Percent excess risk per 10 ppb ozone exposure in spatio-temporal models were significant on lag days 1 through 5, ranging from 6.5 (95% CI: 0.2–13.1) to 13.0 (6.0–20.6) for inpatient hospitalizations, and from 2.9 (95% CI: 0.1–5.7) to 9.4 (6.3–12.7) for ED visits, with strongest associations consistently observed on lag day 2. Spatio-temporal excess risk estimates were consistently but not statistically significantly higher than temporal-only estimates on lag days 0–3.

Conclusion: Incorporating case-level spatial exposure variation produced small, non-significant increases in excess risk estimates. Our modeling approach enables a refined understanding of potential measurement error in temporal-only versus spatio-temporal air pollution exposure assessments. As ozone generally varies over much larger spatial scales than that observed within NYC, further work is necessary to evaluate potential reductions in exposure misclassification for populations spanning wider geographic areas, and for other pollutants.

© 2016 Published by Elsevier Inc.

1. Background

There is substantial evidence linking short-term variation in ambient air pollution with acute asthmatic response in children (Akinbami et al., 2010; Babin et al., 2007; Dong et al., 2011; Halonen et al., 2010; Samoli et al., 2011; Sheffield et al., 2015; Silverman and Ito, 2010; Strickland et al., 2010). Air pollutant concentrations can vary substantially within urban areas (Jerrett et al.,

Abbreviations: EPAAQS, US Environmental Protection Agency Air Quality System; ED, Emergency Department; LUR, Land use regression; NO_2, Nitrogen dioxide; NOAA, National Oceanic and Atmospheric Administration; NYC, New York City; NYCCAS, New York City Community Air Survey; O_3, Ozone; $PM_{2.5}$, Fine particulate matter; RH, Relative humidity; SPARCS, New York State Department of Health Statewide Planning and Research Cooperative System

* Corresponding author.
E-mail addresses: jlcshmool@pitt.edu (J.L.C. Shmool), ejk40@pitt.edu (E. Kinnee), perry.sheffield@mssm.edu (P.E. Sheffield), jcloughe@pitt.edu (J.E. Clougherty).

http://dx.doi.org/10.1016/j.envres.2016.01.020
0013-9351/© 2016 Published by Elsevier Inc.

2005; Clougherty et al., 2013), and this exposure variation has been associated with asthma exacerbation risk (Lemke et al., 2014). Further, because intraurban variation in ozone is driven by availability of nitrogen oxides and photochemical transformation, patterns of ozone exposures can differ substantially from those of other pollutants [e.g., nitrogen dioxide (NO_2), fine particulate matter ($PM_{2.5}$) (NYC DOHMH, 2009a, 2009b)], complicating the interpretability of associations between ozone concentrations and asthma outcomes. Despite these important challenges – and the tremendous public health burden of urban childhood asthma (CDC, 2014; Akinbami et al., 2009) – few studies to date have been able to account for fine-scale spatio-temporal variation in ozone, temperature and co-pollutant exposures in analyses of childhood asthma exacerbation.

Case-crossover designs have been widely applied in epidemiology studies of acute air pollution exposures (Carracedo-Martinez et al., 2010), with certain advantages compared to time-series analyses. Specifically, because each case serves as their own control, time-invariant case characteristics (e.g., race, sex) are controlled for by design (Maclure, 1991). Further, selecting referent (non-case) days immediately before and/or after the case event day effectively accounts for seasonal patterns (Bateson and Schwartz, 1999, 2001). To date, however, a key advantage of case-crossover design for air pollution epidemiology – the ability to incorporate case-level spatial exposure variation – has been under-explored (Carracedo-Martinez et al., 2010). This analytic advantage may be particularly important for reducing exposure misclassification arising from fine-scale spatial variation for short-term air pollution epidemiology.

While there is growing interest in methods for integrating temporal and spatial variation in air pollution exposure assessment (Özkaynak et al., 2013), there are relatively few examples using the case-crossover design. In one example, Delfino et al. (2014) observed significant effect modification of the association between daily air pollution levels (i.e., temporal variation) and asthma morbidity across increasing levels of modeled near-residence traffic-related pollution concentrations (i.e., spatial variation) across Orange County, California. Alternatively, approaches utilizing spatio-temporal exposure metrics (e.g., Ross et al., 2013; Johnson et al., 2013; Maynard et al., 2007) are well-suited to case-crossover design, but, to date, have also been underutilized.

Here we leverage highly spatially-refined air pollution data from the New York City Community Air Survey (NYCCAS) (Clougherty et al., 2013; Matte et al., 2013) to incorporate individual-level spatial exposure variation in a case-crossover analysis of daily ozone exposure and childhood asthma exacerbation events in New York City (NYC). We compare percent excess risk estimates derived from temporal-only vs. spatio-temporal exposure covariates to evaluate differences in magnitude and precision.

2. Methods

2.1. Asthma case data

Emergency Department (ED) visit (outpatient) and hospitalization (inpatient) event data for asthma (ICD-9 code: 493) among children aged 5–17 years old in NYC from 2005 to 2011 (June 1–August 31) were obtained from the New York State Department of Health Statewide Planning and Research Cooperative System (SPARCS). Cases were limited to warm season months (June–August) to match the period over which the NYCCAS spatial ozone assessment occurred, and to capture annual-peak ozone levels resulting from higher rates of photochemical transformation (DOHMH, 2009a). We excluded case events for children younger than 5 years, due to questionable reliability of asthma diagnosis among younger children (Potter, 2010).

2.2. Air pollution data and exposure assignment

We estimated short-term near-residence exposures to ozone and co-pollutants by integrating temporally- and spatially-refined data sources. Spatial data consisted of fine-scale summertime O_3 and co-pollutant (annual average NO_2 and $PM_{2.5}$) concentrations, derived from two years of NYCCAS monitoring data (2009–2010). NYCCAS is an on-going air pollution surveillance initiative of the New York City Department of Health and Mental Hygiene (NYC DOHMH), methods for which are detailed elsewhere (Matte et al., 2013; Clougherty et al., 2013). Briefly, two-week integrated street-level (10–12 feet above the ground) samples were collected at 150 monitoring locations across NYC. Land Use Regression (LUR) was used to predict fine-scale seasonal-average concentration estimates corresponding to 300 m grid centroids, enabling fine-scale exposure estimates.

Temporal data included hourly monitoring data from five US Environmental Protection Agency Air Quality System (EPA AQS) regulatory monitoring stations in NYC, for the years 2005–2011 (May–September). Methods for constructing city-wide daily time-series are detailed elsewhere (Sheffield et al., 2015). Briefly, missing values were interpolated and daily 24-h average O_3 concentration values were computed for each regulatory monitor. We considered the spatial distribution of, and data density within, regulatory monitors to prevent biasing the city-wide time trend. To enable adjustment for potential confounding by co-pollutant exposures, fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) daily time series were constructed using the average of scaled daily values to account for between-site differences in variance and mean of regulatory monitoring data, as in Schwartz (2000).

We combined these spatial and temporal air pollution data to estimate spatio-temporal near-residence exposures. We first applied a multi-step validation and geocoding process, methods for which are detailed in Supplemental Materials. Briefly, our approach used three address locators – the geographic reference data that translate addresses to latitude and longitude point locations – to sequentially geocode addresses, allowing for maximal accuracy and match rates. This approach also enabled sensitivity analysis for geocoding-induced exposure misclassification, wherein we re-fit all models including only the subset of subjects whose addresses were geocoded with the highest positional accuracy.

We estimated near-residence seasonal-average pollution exposures from NYCCAS spatial data as the mean concentration within a 300 m radial buffer around the geocoded point. To generate daily spatio-temporal exposure estimates, we multiplied daily EPA AQS concentrations by the ratio of near-residence (i.e., 300 m buffer mean) concentration to the citywide average of NYCCAS concentrations across populated areas (i.e., census tracts with residential population > 20).

2.3. Meteorological covariates

Spatial temperature data consisted of a NYCCAS-derived fine-scale spatial temperature surface. NYCCAS temperature monitoring data included 15-min measurements from continuous HOBO sensors (Pocasset, MA) deployed at each site throughout all sampling periods, averaged by hour and site. For calculating the spatial predictive surface, overnight (3–5 AM) averages (adjusted for trends at the five NYCCAS reference sites) were used as the dependent variable because they presented the most consistent spatial patterns across seasons. As such, overnight temperatures may be more indicative of consistent spatial differences in ambient

temperature, compared to average or maximum temperature (due to localized intermittent shading in street canyons, etc.).

Temporal temperature data from the four meteorological stations in the NYC area (JFK International Airport, LaGuardia Airport, Central Park, Newark International Airport) was retrieved from the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NOAA, 2014a). Daily minimum temperature, average temperature (T_{avg}), daily maximum temperature (T_{max}), and dew point temperature (T_{DP}) were averaged across the four stations, and these values were highly correlated across stations: T_{min} vs. T_{avg} *rho*=0.93; T_{min} vs. T_{max} *rho*=0.84. Relative humidity (RH) was calculated from T_{avg} and T_{DP} using the standard NOAA equation (NOAA, 2014b). To best match the spatial temperature surface, we created a daily time-series including 3–5 AM hourly average temperature from the four NOAA sites (hereafter T_{min}). As with spatio-temporal air pollution exposure estimates, we adjusted the daily temperature time-series using the spatial ratio of near-residence (i.e., 300 m buffer mean) T_{min} to the city-wide average T_{min} for each case.

2.4. Statistical analysis

We used conditional logistic regression with symmetric bi-directional referent sampling in the case-crossover design. Specifically, referent days were defined as the same day of week and year, +/− two weeks from the case event day (as in Bateson and Schwartz (1999)), such that each case had four referent days. The functional form of model covariates was determined using Likelihood Ratio Test to compare fit across models of increasing complexity [i.e., linear form, natural spline (ns) with defined degrees of freedom (df), or penalized spline (ps) with unlimited df], one variable at a time. Temperature covariates were modeled using natural splines with 3 degrees of freedom, and co-pollutants were modeled using penalized splines.

Building upon previous analysis of this hospitalization and emergency visit data that did not include spatial exposure information (Sheffield et al., 2015), we estimated percent excess risk per 10 ppb change in estimated ozone exposure for lag days 0–6 to capture potential delayed effects of ozone beyond the lag period commonly investigated (e.g., Halonen et al., 2010). Adjustment covariates included: same-day T_{min} (ns, df=3), lagged 6-day average T_{min} (ns, df=3), and 4-day average (lag days 0–3) of estimated $PM_{2.5}$ exposures (ps, df=3). NO_2 was included in the sensitivity analysis detailed below.

To evaluate how incorporating individual-level (spatial) exposure information changed exposure-response relationships, we sequentially fit four models (A–D) moving from temporal-only (A) to spatio-temporal exposure covariates (D), as follows:

1. Model A: Temporal O_3+Temporal T_{min}+Temporal $PM_{2.5}$
2. Model B: Spatio-temporal O_3+Temporal T_{min}+Temporal $PM_{2.5}$
3. Model C: Spatio-temporal O_3+Spatio-temporal T_{min}+Aspatial $PM_{2.5}$
4. Model D: Spatio-temporal O_3+Spatio-temporal T_{min}+Spatio-temporal $PM_{2.5}$

Temporal therefore refers to the model or specific exposure estimates that include only the temporal variation component. *Spatio-temporal* includes both the temporal component (i.e. the estimate for the case day) adjusted spatially per the case residential address.

Analyses were conducted using R statistical package (version 2.14.0; R Development Core Team, 2013).

2.5. Sensitivity analyses

To test the sensitivity of model results to covariate formulation, we re-fit all models with 4 and 5 df and with 7-day average co-pollutant exposure estimates. In previous analyses (Sheffield et al., 2015), we observed no difference in case-crossover model results using the same temporal exposure data for temperature covariates at 3–5 df, or by replacing RH with dew point. Additionally, we re-fit all models with the subset of cases geocoded using the highest positional accuracy address locator (inpatient n=2251; outpatient n=11,121). Because NOx is associated with ground-level O_3 formation (as nitrogen oxides react with ambient ozone, producing localized ozone "scavenging" in very dense urban areas), treating NO_2 as a confounder may not be correct in this city-level domain, and as such we chose not to include NO_2 exposures in our main models, but do explore its effect in sensitivity analyses.

All study protocols were reviewed and approved by the Institutional Review Boards of the University of Pittsburgh and Icahn School of Medicine at Mount Sinai.

3. Results

Our analyses included 2353 asthma inpatient hospitalizations and 11,719 outpatient ED visits (Table 1). Overall, characteristics of in- and outpatient cases were similar. Case events were most common in the youngest age range (5–9 years), and among male children.

Table 2 summarizes distributions for temporal and spatio-temporal exposure distributions, across in- and outpatient populations. Table 3 reports correlations among air pollution and temperature exposure estimates among in- and outpatient cases, respectively. In both populations, temporal and spatio-temporal exposure estimates were highly-correlated (rho=0.99). Temporal and spatio-temporal ozone and T_{min} exposures were correlated at 0.38 and 0.36, respectively, in both in- and outpatient populations. Temporal and spatio-temporal ozone exposure estimates were more highly correlated with $PM_{2.5}$ than with T_{min}, in both case populations. Temporal and spatio-temporal ozone estimates were less correlated with NO_2 (rho=0.29 and 0.22 among inpatient cases; 0.27 and 0.19 among outpatient cases, respectively).

Spatial patterns of ozone ratios for in- (Fig. 1a) and outpatient (Fig. 1b) cases reflect the lower surface concentrations in the city core (Manhattan), and higher concentrations in the outer boroughs. As such, applying spatial ratios to daily regulatory concentrations resulted in lower spatio-temporal than temporal-only O_3 exposure estimates, for approximately 90% of cases. Compared to O_3, $PM_{2.5}$ and NO_2 surface concentrations are higher in the city core, and more spatially-varying across boroughs.

Fig. 2A and B report estimated percent excess risk of an event per 10 ppb increase in ozone exposure for models A–D, ordered from the temporal-only model (A), to the fully spatially-resolved model (D). Overall, trends in association across lag days were similar between in- and outpatient cases, with the strongest excess risk observed from one to three days following exposure (i.e., lags 1–3). For inpatient cases, highest percent excess risks were consistently observed on lag day 2, for which the fully spatio-temporal (model D) excess risk estimates were higher (+2.0%) than temporal (model A) estimates. Excess risk estimates did not, however, differ significantly between these models [13.0% (95% CI: 6.0–20.6) vs. 11.0% (95% CI: 3.6–18.6), respectively]. For outpatient cases, strongest effects were also observed on lag day 2, at slightly lower magnitude and more narrow 95% CIs (due to higher statistical power) compared to inpatient cases. The difference in magnitude of excess risk estimates between models D and A was small (+0.3%) and not statistically significant [9.4% (95% CI 6.3–12.7) vs.

Table 1
Population characteristics. Note that age groupings represent biologically-relevant categories for asthma etiology (Sheffield et al., 2015).

	Inpatient ($n=2353$)		Outpatient ($n=11,719$)	
	N	%	*N*	%
Age categories				
5–9 yr	1246	52.3	5809	49.6
10–13 yr	645	27.4	3253	27.7
14–17 yr	462	19.6	2657	22.7
Sex				
Female	1037	44.1	5108	43.6
Male	1316	55.9	6611	56.4
Borough of residence				
Bronx	845	35.9	4241	36.2
Brooklyn	764	32.5	3189	27.2
Manhattan	281	11.9	2152	18.4
Queens	405	17.4	1911	16.3
Staten Island	54	2.3	226	1.9
Month of asthma event				
June	989	42.0	5004	42.7
July	568	24.1	3178	27.1
August	796	33.8	3537	30.2
Year of asthma event				
2005	423	18.0	2132	18.2
2006	395	16.8	1863	15.9
2007	328	13.9	1634	13.9
2008	264	11.2	1516	12.9
2009	432	18.4	1756	15.0
2010	260	11.1	1428	12.2
2011	251	10.1	1390	11.9

Table 2
Temporal and spatio-temporal air pollution and temperature exposure estimates.

	Mean	Std	Min	Max	IQR
Inpatient ($n=2353$)					
Ozone (case-day) (ppb)					
Temporal	30.4	9.4	5.0	60.0	13.0
Spatio-temporal	29.0	9.1	4.6	60.3	12.5
T_{min} (case-day) (°F)					
Temporal	66.5	5.7	50.5	81.4	7.8
Spatio-temporal	67.4	5.8	50.2	82.9	8.1
$PM_{2.5}$ (4-day average) ($\mu g/m^3$)					
Temporal	13.8	9.1	1.0	52.5	11.6
Spatio-temporal	14.9	10.1	1.0	58.5	12.7
Outpatient ($n=11,719$)					
Ozone (case-day) (ppb)					
Temporal	30.4	9.5	5.0	60.0	13.0
Spatio-temporal	28.7	9.1	3.2	64.4	12.5
T_{min} (case-day) (°F)					
Temporal	66.4	5.8	50.5	81.4	7.8
Spatio-temporal	67.4	5.9	49.7	84.0	8.3
$PM_{2.5}$ (4-day average) ($\mu g/m^3$)					
Temporal	13.6	8.8	1.0	52.5	11.2
Spatio-temporal	14.9	9.8	0.9	73.2	12.4

9.1% (95% CI 5.7–12.5), respectively].

When we considered the sequential incorporation of spatial information into each covariate, between-model differences were consistent, but not statistically significant. The largest increase in effect estimates occurred after substituting spatio-temporal $PM_{2.5}$ (model C to D). These differences were most pronounced on lag days 0–2, with the greatest increase on lag 1 of 3.2 and 1.6% excess risks among inpatient and outpatient populations, respectively. In contrast, substituting spatio-temporal T_{min} (model B to C) produced differences in percent excess risk < 0.1. In addition, we observed small ($\leq 1\%$ excess risk) but consistent gains in precision between model A and D on lag days 2 and 3.

In sensitivity analyses, all model results were robust to adjustment for NO_2 as a co-pollutant confounder, to varying the functional form of co-pollutant covariates (i.e., three to five d.f.), and were consistent among the subset of cases geocoded using the most accurate geocoding method. Models using 7-day co-pollutant averages showed significant effects over shorter periods: lag 1–3 for inpatient, and lag 1–4 for outpatient, compared to 4-day averages in main models.

4. Discussion

We incorporated residential-level spatial exposure contrasts into a case-crossover design, to reduce exposure misclassification in estimating short-term effects of ambient ozone on asthma exacerbation. We found significant positive associations on lag days 1–5, with the strongest effects on lag days 1–3, on which percent excess risk ranged from 10.2% to 13.0% among inpatient populations, and 7.3–9.4% among outpatient populations, in fully-adjusted spatio-temporal models. Our modeling approach enables a refined understanding of potential measurement error in temporal-only versus spatio-temporal air pollution exposure assessments.

Our comparison of temporal vs. spatio-temporal models found generally similar percent excess risks for short-term ozone exposures, with slightly higher estimates on lag days 1 and 2 (+3.2 and 2.0 for inpatients, and+1.6 and 0.3 for outpatients). The greatest (albeit small) change in effect estimates followed the substitution of spatio-temporal adjustment covariate $PM_{2.5}$ (model D), into the model including spatio-temporal forms of ozone and other covariates. One potential explanation is that intra-urban concentration gradients within NYC are more spatially-varying for $PM_{2.5}$ and NO_2, compared to O_3. Because the ozone gradients in dense urban areas is driven by scavenging, NOx and VOC availability, and transformation rates, the highest ozone concentrations occur in the outer-most, sparsely-populated areas of NYC, and in areas downwind of the metropolitan core. As such, adding spatial information on O_3 into case-crossover studies may yield greater exposure contrasts across much larger and varied, regional-scale domains. Likewise, our results indicate that this approach for integrating temporal and spatial exposures in city-level studies may be particularly useful for reducing exposure misclassification for pollutants with more spatial variability.

Interestingly, we found that addition of spatial covariates resulted in a larger change in effect estimates for inpatient child asthma cases than for ED visits. This difference could be driven by underlying population differences, if younger children have more severe asthma and spend more time at home, improving precision in the spatio-temporal exposure estimates. Alternatively, this population may have greater susceptibility to ozone, given more severe or more poorly-controlled disease.

Comparing our findings with the few studies on spatio-temporal ozone variation in relation to asthma exacerbation is challenging, due to differing exposure metrics (e.g., per 10 ppb vs. IQR-increase; single-day lags vs. multi-day moving averages), case definitions (e.g., age categories, ICD-9 codes), and study area characteristics (e.g., spatial variation in ozone; scavenging). In a similar case-crossover study examining summertime all-ages asthma hospitalization cases in NYC, Jones et al. (2013) compared daily O_3 and residential census tract-level O_3 variation [predicted by EPA's Stochastic Human Exposure and Dose Simulation (SHEDS) model], and observed a high correlation between temporal and spatio-temporal exposure estimates, and no difference between hazard ratios (HR) using each exposure metric (HR = 1.029 per IQR,

Table 3
Pearson correlation coefficients across exposure metrics.

	Ozone (temporal)	$PM_{2.5}$ (temporal)	T_{min} (temporal)	Ozone (spatio-temporal)	$PM_{2.5}$ (spatio-temporal)	T_{min} (spatio-temporal)
Inpatient ($n=2353$)						
Ozone (temporal)	1.00					
$PM_{2.5}$ (temporal)	0.60	1.00				
T_{min} (temporal)	0.38	0.46	1.00			
Ozone (spatio-temporal)	0.99	0.60	0.37	1.00		
$PM_{2.5}$ (spatio-temporal)	0.60	0.99	0.46	0.57	1.00	
T_{min} (spatio-temporal)	0.37	0.46	0.99	0.36	0.46	1.00
Outpatient ($n=11,719$)						
Ozone (temporal)	1.00					
$PM_{2.5}$ (temporal)	0.58	1.00				
T_{min} (temporal)	0.38	0.45	1.00			
Ozone (spatio-temporal)	0.99	0.57	0.37	1.00		
$PM_{2.5}$ (spatio-temporal)	0.57	0.99	0.45	0.54	1.00	
T_{min} (spatio-temporal)	0.38	0.45	0.99	0.36	0.46	1.00

95% CI: 1.01–1.05). We sought further spatial refinement by using near-residence (300 m buffer) spatio-temporal estimates, and also found high correlations with temporal exposure estimates, but did observe small though non-significant between-model differences. In a separate study utilizing temporal (Poisson) case-crossover design to evaluate the association between 8-h max ozone and asthma ED visits among children (aged 5–17), Strickland et al. (2010) observed elevated relative risk [RR=1.08 (1.04–1.12)] during warm season months. Because an IQR for the 3-day moving average in ozone, in their Atlanta study area, was reported as 29 ppb, this increased risk of 8% appears substantially lower than our finding of 7.3% and 9.4% excess risk per 10 ppb on lag days 1 and 2, respectively, potentially due to unmeasured spatial exposure variation.

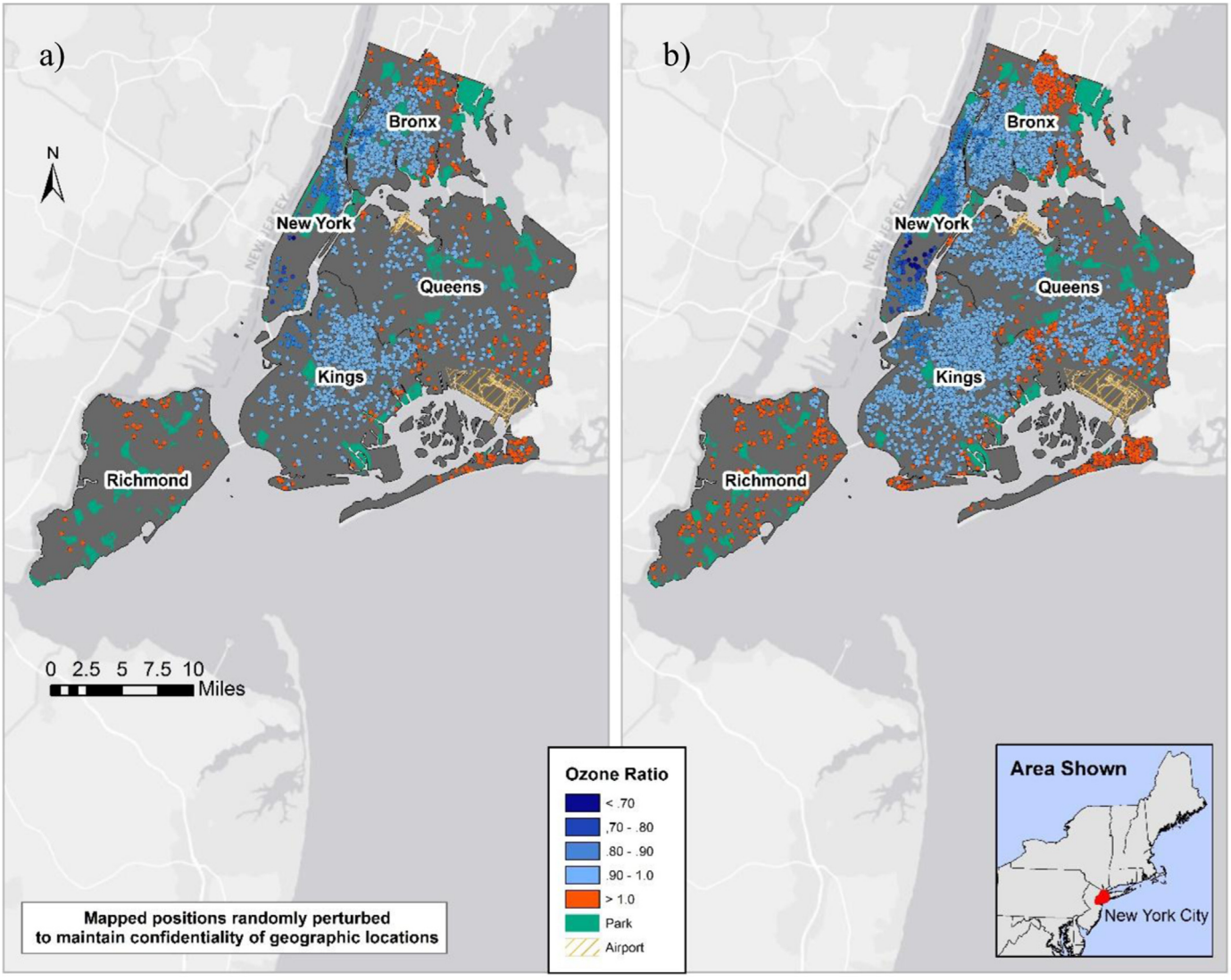


Fig. 1. Spatial ozone ratios for (a) Inpatient ($n=2353$; range=0.69–1.13) and (b) Outpatient ($n=11,719$; range=0.57–1.16).

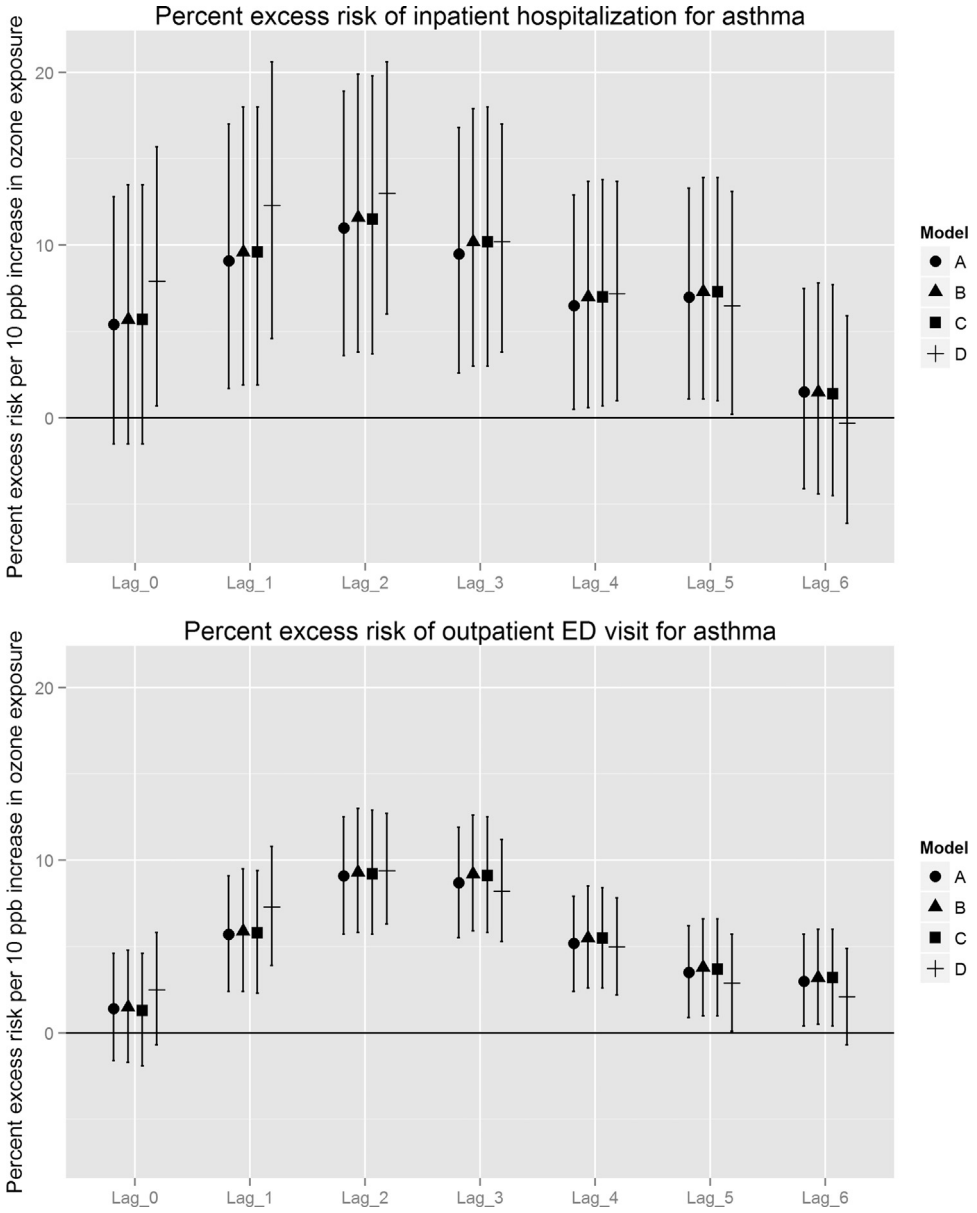


Fig. 2. A. Percent excess risk of inpatient hospitalization for asthma, per 10 ppb increase in ozone exposure: comparison of models A–D. B. Percent excess risk of outpatient ED visit for asthma, per 10 ppb increase in ozone exposure: comparison of models A–D.

5. Limitations

Our exposure assignment has multiple limitations. First, near-residence exposure estimates do not capture the day-to-day activity patterns of individuals (e.g., travel to school), and we cannot know the degree of exposure misclassification induced, or whether it is non-differential. In our integration of spatial and temporal exposure data, we assumed similar spatial variation in air pollution and temperature across study years, while the spatial data reflect seasonal (O_3) and annual (temperature, NO_2 and $PM_{2.5}$) averages for 2008–2010. While spatial temperature data used overnight temperatures to reflect the most robust spatial pattern over time, season-specific patterns in maximum daytime temperatures vary and may represent different physiologic risk for acute health events, compared to minimum temperature. The choice of buffer size for exposure assignment is another potential source of error measurement, common in air pollution epidemiology. However, prior work with NO_2 and $PM_{2.5}$ exposure assessment using NYCCAS data evaluated the effect of buffer size (address point, 300 m, and 800 m), and observed that health effects analyses would not be sensitive to buffer size within this near-residence range (Ross et al., 2013). Though between-season differences in NO_2 and $PM_{2.5}$ are substantially less than for ozone (NYC DOHMH, 2009a, 2009b), we were not able to use summer seasonal average co-pollutant data. We were unable to account for the potential contribution of pollen to respiratory outcomes due to lack of data; limiting our analyses to June–August, however, effectively excluded peak spring pollen periods, which occur March–May in NYC (Sheffield et al., 2011). Finally, we were unable to account for potential "avoidance behaviors," wherein people may change their behaviors (and resultant exposures) in response to air quality information (e.g., public alerts); controlling for these behavioral changes has been shown to increase observed strengths of associations (Neidell and Kinney, 2010), presumably by reducing exposure misclassification.

6. Strengths

The case-crossover design inherently controlled for individual-

level time-invariant confounders (e.g., age, sex), and enabled incorporation of both spatial and temporal exposure information. A unique contribution of our study is the inclusion of spatial variation in temperature as an adjustment covariate. Our results were robust to a range of sensitivity analyses, and demonstrated consistent patterns between in- and outpatient case populations. Although our approach for reducing exposure misclassification did not reveal significant differences in effect estimates when incorporating spatial information on exposures, we developed a useful approach for case-crossover studies to distinguish relative contributions of information on spatial and temporal exposure variation.

7. Conclusions

Incorporating spatial information on variation in ambient ozone exposures slightly increased excess risk estimates in our study. As ozone normally varies over much larger spatial scales than that observed in NYC, further work is necessary to evaluate potential reductions in exposure misclassification for populations spanning wider geographic areas, and for other pollutants.

Acknowledgments

The authors are grateful to Zev Ross (Zev Ross Spatial Analysis) for methods consultation, and to Jiang Zhou (formerly of University of Pittsburgh) for work on early statistical models. This research was funded by US NIH (NIEHS) Grants R21ES021429, R01ES019955 and K23ES024127 (PS); US EPA Grant RD-83457601-0.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2016.01.020.

References

Akinbami, L.J., Moorman, J.E., Garbe, P.L., Sondik, E.J., 2009. Status of childhood asthma in the United States, 1980–2007. Pediatrics 123 (Suppl. 3), S131–S145. http://dx.doi.org/10.1542/peds.2008-2233C.

Akinbami, L.J., Lynch, C.D., Parker, J.D., Woodruff, T.J., 2010. The association between childhood asthma prevalence and monitored air pollutants in metropolitan areas, United States, 2001–2004. Environ. Res. 110 (3), 294–301. http://dx.doi.org/10.1016/j.envres.2010.01.001.

Babin, S.M., Burkom, H.S., Holtry, R.S., Tabernero, N.R., Stokes, L.D., Davies-Cole, J.O., et al., 2007. Pediatric patient asthma-related emergency department visits and admissions in washington, DC, from 2001–2004, and associations with air quality, socio-economic status and age group. Environ. Health 6, 9. http://dx.doi.org/10.1186/1476-069X-6-9.

Bateson, T.F., Schwartz, J., 2001. Selection bias and confounding in case-crossover analyses of environmental time-series data. Epidemiology 12 (6), 654–661.

Bateson, T.F., Schwartz, J., 1999. Control for seasonal variation and time trend in case-crossover studies of acute effects of environmental exposures. Epidemiology 10 (5), 539–544.

Carracedo-Martinez, E., Taracido, M., Tobias, A., Saez, M., Figueiras, A., 2010. Case-crossover analysis of air pollution health effects: a systematic review of methodology and application. Environ. Health Perspect. 118 (8), 1173–1182. http://dx.doi.org/10.1289/ehp.0901485.

CDC, 2014. Asthma Facts: CDC's National Asthma Control Program Grantees. ⟨http://www.cdc.gov/asthma/pdfs/asthma_facts_program_grantees.pdf⟩ ed.

Clougherty, J.E., et al., 2013. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: the New York City Community Air Survey (NYCCAS). J Expo. Sci Environ. Epidemiol. 23 (3), 232–240.

Delfino, R.J., Jun, W., Tjoa, T., Gullerserian, S.K., Nickerson, B., Gillen, D.L., 2014. Asthma morbidity and ambient air pollution: effect modification by residential traffic-related air pollution. Epidemiology 25 (1), 48–57. http://dx.doi.org/10.1097/EDE.0000000000000016.

Dong, G.H., Chen, T., Liu, M.M., Wang, D., Ma, Y.N., Ren, W.H., et al., 2011. Gender differences and effect of air pollution on asthma in children with and without allergic predisposition: Northeast Chinese children health study. PLoS One 6 (7), e22470. http://dx.doi.org/10.1371/journal.pone.0022470.

Halonen, J.I., Lanki, T., Tiittanen, P., Niemi, J.V., Loh, M., Pekkanen, J., 2010. Ozone and cause-specific cardiorespiratory morbidity and mortality. J. Epidemiol. Community Health 64 (9), 814–820. http://dx.doi.org/10.1136/jech.2009.087106.

Jerrett, M., Burnett, R.T., Ma, R., Pope 3rd, C.A., Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E., Then, M.J., 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiol 16 (6), 727–736.

Johnson, M., Macneill, M., Grgicak-Mannion, A., Nethery, E., Xu, X., Dales, R., et al., 2013. Development of temporally refined land-use regression models predicting daily household-level air pollution in a panel study of lung function among asthmatic children. J. Expo. Sci. Envirn. Epidemiol. 23 (3), 259–267. http://dx.doi.org/10.1038/jes.2013.1.

Jones, R.R., Özkaynak, H., Nayak, S.G., Garcia, V., Hwang, S., Lin, S., 2013. Associatrions between Summertime ambient pollutants and respiratory morbidity in New York City: comparison of results using ambient concentrations versus predicted exposures. J. Expo. Sci. Environ. Epidemiol. 23, 616–626.

Lemke, L.D., Lamerato, L.E., Xu, X., Booza, J.C., Reiners, J.J., Raymond III, D.M., Krouse, H.J., 2014. Geospatial relationships of air pollution and acute asthma events across the Detroit–Windsor international border: study design and preliminary results. J. Expo. Sci. Environ. Epidemiol. 24 (4), 346–357. http://dx.doi.org/10.1038/jes.2013.78.

Maclure, M., 1991. The case-crossover design: a method for studying transient effects on the risk of acute events. Am. J. Epidemiol. 133 (2), 144–153.

Matte, T.D., Ross, Z., Kheirbek, I., Eisl, H., Johnson, S., Gorczynski, J.E., et al., 2013. Monitoring intraurban spatial patterns of multiple combustion air pollutants in New York City: design and implementation. J. Expo. Sci. Environ. Epidemiol. 23 (3), 223–231. http://dx.doi.org/10.1038/jes.2012.126.

Maynard, D., Coull, B.A., Gryparis, A., Schwartz, J., 2007. Mortality risk associated with short-term exposure to traffic particles and sulfates. Environ. Health Perspect. 115 (5), 751–755.

New York City Department of Health and Mental Hygiene (NYC DOHMH), 2009a. New York City Community Air Survey (NYCCAS): Results from Summer Monitoring 2009. Available: ⟨http://www.nyc.gov/html/doh/html/environmental/community-air-survey.shtml⟩.

New York City Department of Health and Mental Hygiene (NYC DOHMH), 2009b. New York City Community Air Survey (NYCCAS): Results from Winter Monitoring 2008–2009. [Available: ⟨http://www.nyc.gov/html/doh/html/environmental/community-air-survey.shtml⟩.

NOAA, 2014a. Climate Data Online – Select Area. ⟨http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD&countryabbv=&georegionabbv=⟩.

NOAA, 2014b. NOAA's National Weather Service-Binghamton, NY – Relative Humidity. ⟨http://www.erh.noaa.gov/bgm/tables/rh.shtml⟩.

Özkaynak, H., Baxter, L.K., Burke, J., 2013. Evaluation and application of alternative air pollution exposure metrics in air pollution epidemiology studies. J. Expo. Sci. Environ. Epidemiol. 23, 566–572. http://dx.doi.org/10.1038/jes.2013.50.

Potter, P.C., 2010. Current guidelines for the management of asthma in young children. Allergy Asthma Immunol. Res. 2 (1), 1–13. http://dx.doi.org/10.4168/aair.2010.2.1.1.

Ross, Z., Ito, K., Johnson, S., Yee, M., Pezeshki, G., Clougherty, J.E., et al., 2013. Spatial and temporal estimation of air pollutants in New York City: exposure assignment for use in a birth outcomes study. Environ. Health 12, 51. http://dx.doi.org/10.1186/1476-069X-12-51.

Samoli, E., Nastos, P.T., Paliatsos, A.G., Katsouyanni, K., Priftis, K.N., 2011. Acute effects of air pollution on pediatric asthma exacerbation: evidence of association and effect modification. Environ. Res. 111 (3), 418–424. http://dx.doi.org/10.1016/j.envres.2011.01.014.

Schwartz, J., 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11 (3), 320–326.

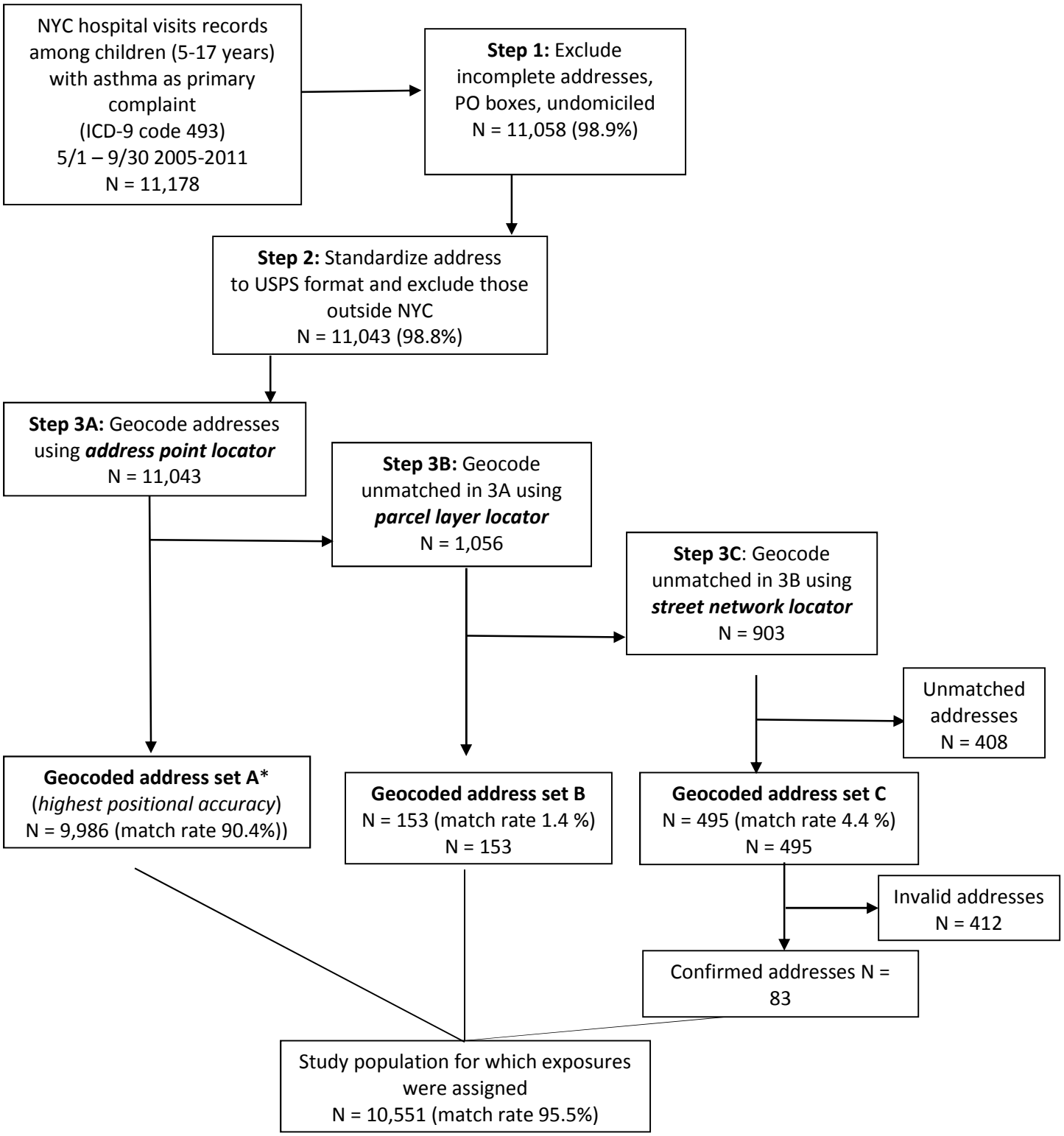
Sheffield, P.E., Zhou, J., Shmool, J.L.C., Clougherty, J.E., 2015. Ambient ozone exposure and children's acute asthma in New York City: a case-crossover analysis. Environ. Health . http://dx.doi.org/10.1186/s12940-015-0010-2.

Sheffield, P.E., Weinberger, K.R., Ito, K., Matte, T.D., Mathes, R.W., Robinson, G.S., Kinney, P.L., 2011. The association of tree pollen concentration peaks and allergy medication sales in New York City: 2003–2008. ISRN Allergy 2011, 537194. http://dx.doi.org/10.5402/2011/537194.

Silverman, R.A., Ito, K., 2010. Age-related association of fine particles and ozone with severe acute asthma in New York City. J. Allergy Clin. Immunol. 125 (2), 367–373. http://dx.doi.org/10.1016/j.jaci.2009.10.061.

Strickland, M.J., Darrow, L.A., Klein, M., Flanders, W.D., Sarnat, J.A., Waller, L.A., et al., 2010. Short-term associations between ambient air pollutants and pediatric asthma emergency department visits. Am. J. Respir. Crit. Care Med. 182 (3), 307–316. http://dx.doi.org/10.1164/rccm.200908-1201OC.

NYC hospitalization case data exclusions and geocoding methodology:



* Because addresses geocoded using the point locator method (Set A) have the highest positional accuracy, we use this set to test the sensitivity of our epidemiological models to potential exposure errors induced by alternative methods (i.e., parcel and street network locators methods).

Geocoding methods:

Geocoding residential address data to obtain location coordinates is a necessary first step for assigning fine-scale exposure measures (e.g., mean air pollution concentration within 300m buffer). The need to geocode study subjects' residential location is increasing with greater use of GIS-based environmental information and spatial statistical techniques. Geocoding process rely on an *address locator,* which specifies reference data and defines the address format and rules for translating street addresses into geospatial data (i.e., assign longitude and latitude coordinates) that can be mapped. Address locators can take many forms depending on the availability of reference data, and the choice of which locator to use has implicit trade-offs between maximizing the match rate (increasing analytic power), versus positional accuracy (minimizing exposure misclassification). As such, we developed a multi-step geocoding process that first filters and standardizes addresses before running them through a composite address locator. The composite locator utilizes three separate sets of reference data to obtain the highest degree of positional accuracy while maximizing match rates using different address sensitivity rules.

Step 1. We first excluded all incomplete address records (e.g., having no street number). To restrict our cohort to NYC residents, we excluded all records with ZIP Codes outside of the NYC. Because PO boxes represent the location of a Post Office, rather than a residence, we also exclude PO Box records. Similarly, address records coded for homeless hospital patients (e.g., 999 Homeless; Undomiciled) were excluded.

Step 2. Patient address data is collected by hospitals, and must undergo substantial format standardization in order to meet U.S. Postal Service specifications and be compatible with GIS-based address locators. We performed a variety of address cleaning processes, such as standardizing common address abbreviations (e.g., BLVD for Boulevard). Data cleaning steps were performed in SAS™ (Cary, NC) and gVim 7.4, an open source PC version of the Unix vi editor. Next, we used U.S. Postal Service reference databases, available through ZP4™ address standardization software (Semaphore Corporation, Monterey, CA) to correct and standardize address format. Standardizing address format helps to ensure that locally-specific differences in address structures (e.g., 12-34 Queens Blvd., vs. 1234 Manhattan Blvd.) do not inadvertently induce spatial patterns in un-matched cohort addresses (e.g., systematically missing non-geocoded Queens addresses, where a distinct address format uses hyphenated street numbers).

Step 3. We constructed a three-level composite address locator to ensure that the greatest number of addresses would be matched using the most accurate reference data. Addresses were geocoded sequentially through each of three separate locators, beginning with the locator with highest locational accuracy. Addresses not successfully geocoded by a given locator (i.e., matched) were then run through the subsequent locator to attempt a match, albeit with lower locational accuracy.

> **Step 3A:** The first level used an *address point locator*, which geolocated address records based on address-level point reference data. These are delivery point locations, precise enough to assign latitude and longitude coordinates as the centroid of individual

building footprints. This first layer provides the highest degree of positional accuracy, but is the most restrictive with respect to exact address formatting requirements.

Step 3B: Unmatched addresses from 3A were run through the second level, which used a *parcel-based locator* to match addresses to the corresponding tax parcels. In NYC, tax parcels are determined by the Department of Finance for the purpose of assessing and collecting property taxes, and as such generally represent building lots which vary widely with respect to both size and the location of buildings within the lot area. Once the locator matches an address to a tax parcel, the address is assigned the latitude and longitude coordinates corresponding to the parcel centroid. Thus, the positional accuracy of geocoded points varies in relation to parcel size and building orientation.

Step 3C: Unmatched addresses from 3B were run through the third level, which used a *street network locator* to match addresses to street segments and assign coordinates along these segments based on the address number. Each street segment contains the range of address numbers for both sides of the street. The locator first matches address records by street name, and then uses the address number range to interpolate an address point location along the appropriate side (i.e., even vs. odd numbers) of the street center line. For example, the address 1250 Manhattan Blvd. would be located at the approximate midpoint of the street segment with numeric range of 1200 – 1298 Manhattan Blvd. Street locators are the most common geocoding method applied in public health research, however, this method has been shown to induce positional errors as great as a few hundred meters (Jacquemin et. al, 2013; Zandbergen and Green, 2007), with the potential for systematic misclassification across heterogeneous study domains (e.g., urban vs. rural differences, intra-urban differences by street network complexity). Because addresses successfully matched in this third layer are not validated against real-world address data, we used the ZP4 software to verify that these addresses were eligible for mail delivery and excluded any points that were non-confirmed.

Geocoded residential locations were then used to assign individual-level air pollution exposure estimates. Exposure assessment for the main epidemiological analysis used all successfully geocoded records from each locator (i.e., address set A + B + C), and we tested the sensitivity of model results to exposure misclassification due to positional errors by re-fitting models with only population set A (n = 9986), which are geocoded with highest positional accuracy.

References:

Jacquemin B, Lepeule J, Boudier A, Arnould C, Benmerad M, Chappaz C, Ferran J, Kauffmann F, Morelli X, Pin I, Pison C, Rios I, Temam S, Kunzli N, Slama R, Siroux V. 2013. Impact of geocoding methods on associations between long-term exposure to urban air pollution and lung function. *Environ Health Perspect* 9:1054-1059.

Zandbergen PW, Green JW. 2007. Error and bias in determining potential of children at school locations using proximity-base GIS techniques. *Environ Health Perspect* 9:1363-1370.