Estimating the Causal Effect of Fine Particulate Matter Levels on Death and Hospitalization: Are Levels Below the Standards Harmful?

Supplementary material

Summary statistics 1

Propensity score overlap…………………………………………………………………………….5

Sensitivity analyses... 5
SA1: Continuous exposure and Cox proportional hazard model with linear exposure-response (ER) function 6
SA2: Continuous exposure and Cox proportional hazard model with non-linear exposure-response (ER) function 7
SA3: Low pollution cohort results using WHO standard of 10 $\mu g/m^3$ ………………...9

Summary statistics

Table S1: Mean and standard deviation of all variables (potential confounders, p=122) and of the four health outcomes available in the study. Calculations shown are for the full cohort and computed over all observations (total = 68,789). Study period 2002 to 2010. Variables from US census and USDA are available at the area level (zip codes and county) and mean and standard deviations are calculated across these area-level summaries.

	High Pollution (> 12 $\mu g/m^3$)		Low Pollution (< 12) $\mu g/m^3$	
Demographic variables	Mean	SD	Mean	SD
Male (Med)	0.440	0.496	0.464	0.499
Age 65- (Med)	0.197	0.398	0.179	0.383
Age 66-75 (Med)	0.324	0.468	0.327	0.469
Age 76-85 (Med)	0.323	0.468	0.332	0.471
Age 86+ (Med)	0.157	0.364	0.162	0.368
White (Med)	0.830	0.376	0.897	0.303
Black (Med)	0.135	0.342	0.063	0.243
Other race (Med)	0.037	0.188	0.041	0.198
Dual eligible (Med)	0.229	0.420	0.196	0.397
% urban (Us Cen)	73.850	33.844	70.080	36.917
% 65+ (US Cen)	14.963	6.093	15.842	7.666
% White (US Cen)	74.277	24.226	81.271	17.498

% Black (US Cen)	16.354	22.039	8.434	13.646
% Asian (US Cen)	3.374	6.241	2.964	5.077
% Poor (US Cen)	16.006	9.038	14.852	8.688
% in college (US Cen)	25.431	11.290	25.300	12.784
% Bachelors + (US Cen)	25.027	15.874	25.784	14.797
% Foreigen (US Cen)	9.239	11.277	9.647	9.862
House value (US Cen)	190594.700	152377.100	190985.900	128841.000
Metropolitan, score 1-3 (USDA)	0.734	0.442	0.683	0.465
Urban score 4-6 (USDA)	0.203	0.402	0.208	0.406
Rural score 7-9 (USDA)	0.063	0.243	0.109	0.311
MCBS Variables				
Bad hearing/ vision	0.834	0.372	0.838	0.369
Less than high school	0.327	0.469	0.251	0.434
High school or more	0.660	0.474	0.736	0.441
Good health status	0.566	0.496	0.589	0.492
3+ children	0.460	0.498	0.499	0.500
<= 2 children	0.540	0.498	0.501	0.500
Divorce	0.097	0.297	0.104	0.305
Married	0.447	0.497	0.484	0.500
Unkown marital status	0.001	0.031	0.000	0.022
Never married	0.110	0.313	0.101	0.302
Seperated	0.017	0.127	0.012	0.108
Widowed	0.328	0.470	0.299	0.458
Income dec 1	0.108	0.310	0.091	0.287
Income dec 2	0.109	0.312	0.091	0.287
Income dec 3	0.103	0.304	0.092	0.290
Income dec 4	0.098	0.298	0.090	0.286
Income dec 5	0.098	0.298	0.089	0.284
Income dec 6	0.099	0.299	0.099	0.299
Income dec 7	0.101	0.302	0.105	0.307
Income dec 8	0.093	0.291	0.110	0.313
Income dec 9	0.097	0.296	0.116	0.321
Income dec 10	0.093	0.291	0.117	0.321
Employed	0.106	0.308	0.109	0.312

Excellent health status	0.121	0.327	0.148	0.355
Fair health status	0.217	0.412	0.186	0.389
Poor health status	0.095	0.294	0.077	0.267
Health better	0.148	0.355	0.148	0.355
Health same	0.605	0.489	0.620	0.485
Health worse	0.247	0.431	0.231	0.422
Mammogram	0.249	0.432	0.253	0.435
Pap smear	0.171	0.377	0.153	0.360
Hsterectomy	0.180	0.385	0.161	0.368
Rectal prostate exam	0.175	0.380	0.183	0.387
Blood test	0.262	0.440	0.280	0.449
Flu shot	0.670	0.470	0.707	0.455
Pneumonia shot	0.674	0.469	0.718	0.450
Hx smoking	0.556	0.497	0.578	0.494
Current smoker	0.145	0.352	0.131	0.338
Cataract op	0.265	0.441	0.297	0.457
No difficulty stooping	0.456	0.498	0.482	0.500
Difficulty stooping	0.365	0.482	0.334	0.472
Some difficulty stooping	0.179	0.383	0.184	0.388
No difficulty lifiting	0.672	0.469	0.707	0.455
Difficulty lifting	0.230	0.421	0.198	0.398
Some difficulty lifting	0.098	0.297	0.095	0.293
No difficulty reaching	0.786	0.410	0.806	0.395
Difficulty reaching	0.115	0.320	0.103	0.303
Some difficulty reaching	0.098	0.298	0.091	0.288
No difficulty walking	0.564	0.496	0.604	0.489
Difficulty walking	0.336	0.472	0.298	0.457
Some difficulty walking	0.100	0.300	0.098	0.298
Problems with phone	0.084	0.277	0.081	0.272
Problems shopping	0.143	0.350	0.120	0.325
Problems bills/money	0.089	0.285	0.079	0.269
Hx cancer: skin	0.176	0.381	0.212	0.409
Hx cancer	0.166	0.372	0.182	0.386
Hx cancer: colon	0.023	0.151	0.024	0.152
Hx cancer: breast	0.040	0.196	0.047	0.211
Hx cancer: uterus	0.016	0.125	0.016	0.124

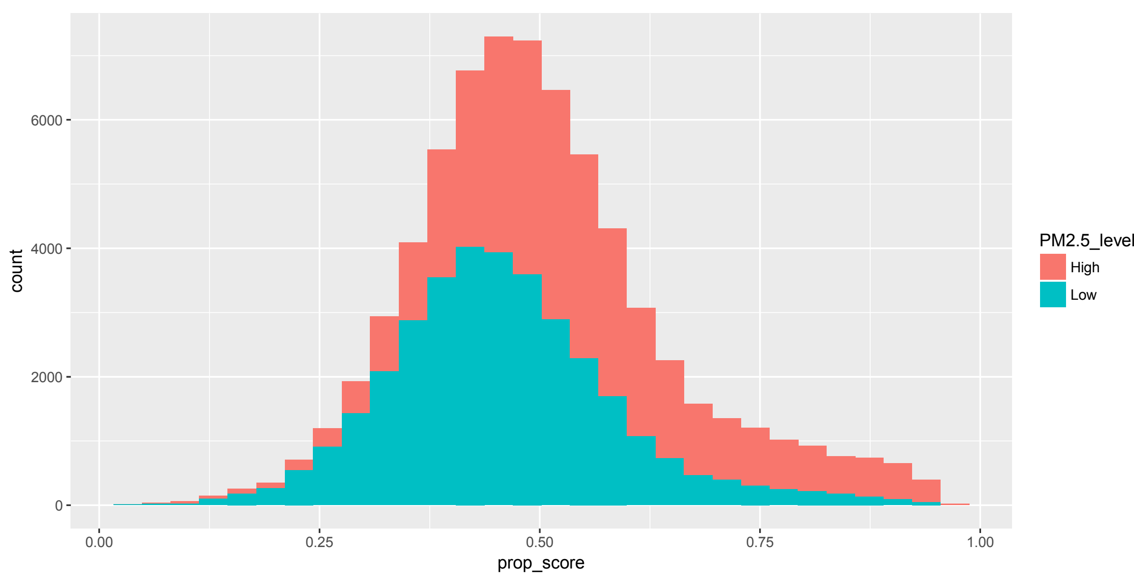
Hx cancer: prostate	0.036	0.186	0.041	0.198
Hx cancer: bladder	0.008	0.086	0.010	0.100
Hx cancer: ovary	0.006	0.079	0.007	0.080
Hx cancer: stomach	0.004	0.065	0.004	0.062
Hx cancer: cervix	0.006	0.078	0.006	0.076
Hx diabetes	0.213	0.410	0.203	0.402
Hx arthritis	0.600	0.490	0.580	0.494
Hx alzheimer	0.043	0.203	0.036	0.187
Hx psych	0.155	0.361	0.137	0.344
Hx osteoperosis	0.206	0.404	0.208	0.406
Hx broken hip	0.039	0.193	0.037	0.190
Hx Parkinsons	0.017	0.128	0.017	0.129
BMI	27.223	6.064	27.227	6.016
Obese	0.253	0.435	0.252	0.434
Medicare Variables				
Proc digestive sys	0.0352	0.1843	0.0294	0.1689
Proc urinary sys	0.0075	0.0863	0.0079	0.0887
Proc male genital org	0.0030	0.0547	0.0031	0.0554
Proc female genital org	0.0031	0.0552	0.0037	0.0604
Obstetrical Procs	0.0006	0.0235	0.0007	0.0258
Musculoskeletal procs	0.0277	0.1640	0.0294	0.1689
Integumentary sys procs	0.0091	0.0947	0.0078	0.0879
Nervous sys procs	0.0065	0.0803	0.0072	0.0846
Endocrine procs	0.0009	0.0295	0.0008	0.0290
Eye procs	0.0003	0.0162	0.0003	0.0162
Nose/mouth/Pharynx procs	0.0013	0.0358	0.0011	0.0328
Infectious/Parasitic Dx	0.0318	0.1756	0.0296	0.1696
Genitourinary Dx	0.0726	0.2594	0.0725	0.2593
Skin Dx	0.0201	0.1402	0.0176	0.1314
Musculoskeletal Dx	0.0804	0.2719	0.0776	0.2676
Congential Anomalies	0.0034	0.0583	0.0039	0.0623
Perinatal conditions	0.0001	0.0076	0.0001	0.0093
Ill-Defined signs/Symptoms	0.0950	0.2932	0.0887	0.2843
Injury/poison	0.0468	0.2113	0.0479	0.2135
Supp: health status	0.1097	0.3125	0.1074	0.3096
External Injury	0.0114	0.1061	0.0136	0.1157

Neoplasms	0.0224	0.1480	0.0203	0.1411
Endocrine Dx	0.1419	0.3489	0.1331	0.3397
Mental disorders	0.0707	0.2563	0.0663	0.2487
Nervous sys dx	0.0519	0.2219	0.0486	0.2151
Digestive sys Dx	0.0835	0.2767	0.0760	0.2650
Ear procs	0.0000	0.0054	0.0001	0.0076
Outcomes				
All cause death	0.046	0.210	0.045	0.207
All cause admissions	0.233	0.423	0.211	0.408
Circulatory admissions	0.199	0.399	0.181	0.385
Respiratory admissions	0.101	0.301	0.093	0.291

Propensity score overlap

To ensure that the distribution of the covariates are overlapping in the region of high ad low pollution, we have plotted a histogram of the estimated propensity scores in both the high (> 12) and low pollution (< 12) groups, respectively (see Figure **S1**). We found that there is adequate overlap between these two distributions suggesting that positivity assumption is not violated in our study.

Figure S1: Distribution of the estimated propensity scores for the high and low exposure groups



Sensitivity Analyses

Table S2: Summary of the main analysis and of the sensitivity analyses

<table>
<tr><td>Model name</td><td>Exposure variable</td><td>Adjusting for confounders</td><td>Results reported in</td></tr>
<tr><td>Main Analysis:
Inverse probability weighting with Binary Exposure</td><td>Binary (<12 for the full cohort, <8 for the low pollution cohort)</td><td>Inverse probability weighting</td><td>Main manuscript
Table 2, and Figure 4</td></tr>
<tr><td>SA1: Continuous exposure, linear exposure-response function and confounding adjustment using a regression model</td><td>Continuous</td><td>Cox proportional hazard model, linear E-R function, potential confounders are included into the model as linear terms</td><td>Supplementary material:
Table S3, and Figure S2</td></tr>
<tr><td>SA2:
Continuous exposure, nonlinear (spline) exposure response function and adjustment using a regression model</td><td>Continuous</td><td>Cox proportional hazard model with a nonlinear E-R function, potential confounders are included into the model as linear terms</td><td>Supplementary material:
Table S4, and Figure S3</td></tr>
<tr><td>SA3:
Low pollution cohort analysis where binary exposure is defined using the WHO cutoff of 10 $\mu g/m^3$.</td><td>Binary</td><td>Inverse probability weighting</td><td>Supplementary material:
Table S5, and Figure S4</td></tr>
<tr><td colspan="4">All analyses were tested for robustness against exclusion of all the behavioral risk factors measured from MCBS (e.g. smoking, BMI, etc.) from the confounding adjustment (exclusion of the 73 covariates available from MCBS only).</td></tr>
<tr><td colspan="4">All analyses were done using the full cohort (N=32,119) and the low pollution cohort (N=18,144)</td></tr>
</table>

Main Analysis: Inverse probability weighting with Binary Exposure: In the main analyses we presented results by using inverse probability weighting (IPW). IPW estimates the effect of the exposure by re-weighting our sample using the stabilized inverse probability weights (Table 2, and Figure 4).

Sensitivity Analyses

SA1: Continuous exposure model

We also fitted to the data a Cox proportional hazard model with a linear E-R function, where the p=122 potential confounders are included into the model as linear terms. We define the continuous exposure as average of $PM_{2.5}$ two years prior the interview date.

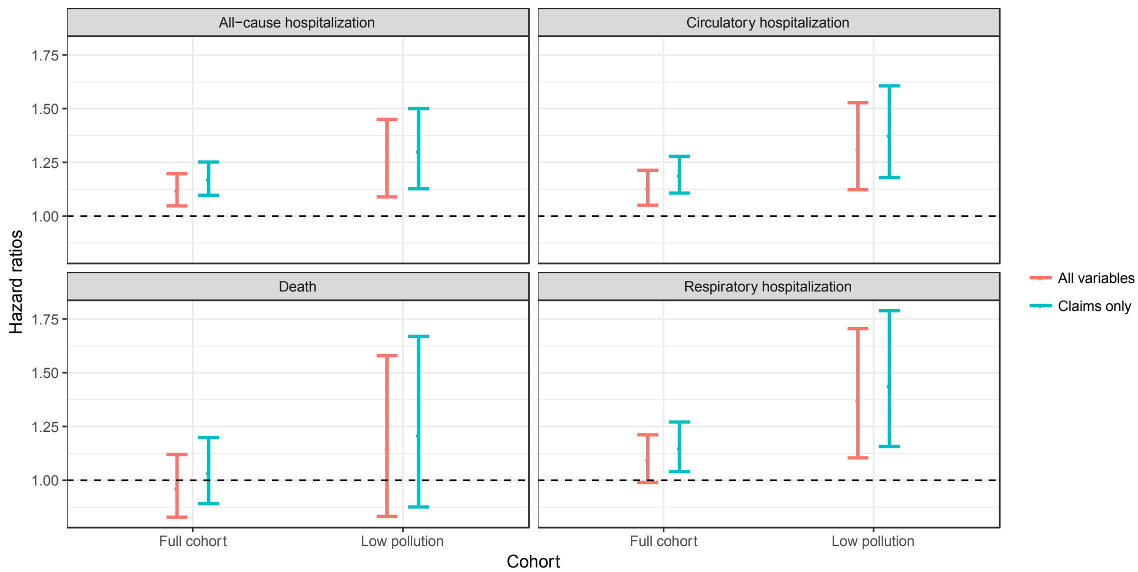
Table S3 shows the estimated hazard ratios and confidence intervals. Hazards ratio estimates are calculated as exponents of the coefficients and confidence intervals are calculated based on robust variances accounting for multiple observations per patient. Our estimate for the effect of $PM_{2.5}$ on all cause mortality is comparable to the same estimate in Table 2 of the ACS although it is not statistically significant. The ACS study reports an estimated HR of 1.06 (1.02, 1.11) associated with a 10 μg/m^3 increase in $PM_{2.5.}$ This is largely consistent with our findings in light of differences between the two studies as outlined in Table 1 in the main text. For example, since the ACS study was conducted at an earlier time period with higher average $PM_{2.5}$ and a larger study population, it does not come as a surprise that their estimate is significant while ours is not.

Table S3: Hazard ratios showing the effect of an increase of 10 μg/m^3 increase in $PM_{2.5.}$ obtained from fitting a CPH with the average exposure to $PM_{2.5}$ as the main exposure variable. Table reports 95% confidence intervals based on robust, sandwich variance estimators (Results of SA1).

	Full cohort N = 32,119 person years = 68,789	Low pollution cohort, N = 18,144 person years = 34,429
All cause mortality	0.96 (0.83, 1.12)	1.15 (0.83, 1.58)
All cause hospitalization	1.12 (1.05, 1.20)	1.26 (1.09, 1.45)
Circulatory hospitalization	1.13 (1.05, 1.21)	1.31 (1.12, 1.53)
Respiratory hospitalization	1.10 (0.99, 1.21)	1.37 (1.10, 1.71)

Figure S2 shows the CPH estimates using the linear exposure term and after direct adjustment for confounders including and excluding MCBS variables and is similar to Figure 2 in the main analysis. One again we find that the results are consistent with the main analysis, suggesting that our estimates are robust to the exclusion of the MCBS variables into the model to adjust for confounding.

Figure S2: shows the estimated hazard ratios and 95% CI obtained by fitting a Cox Proportional Hazard Model (CPH) with a continuous exposure and confounding adjustment obtained by including the covariates linearly into the model. Estimates in blue are obtained by excluding from the model the
MCBS variables. (Results of SA1)



SA2: Continuous Exposure and Cox Proportional Hazard Model with Non Linear Exposure-Response (ER) function

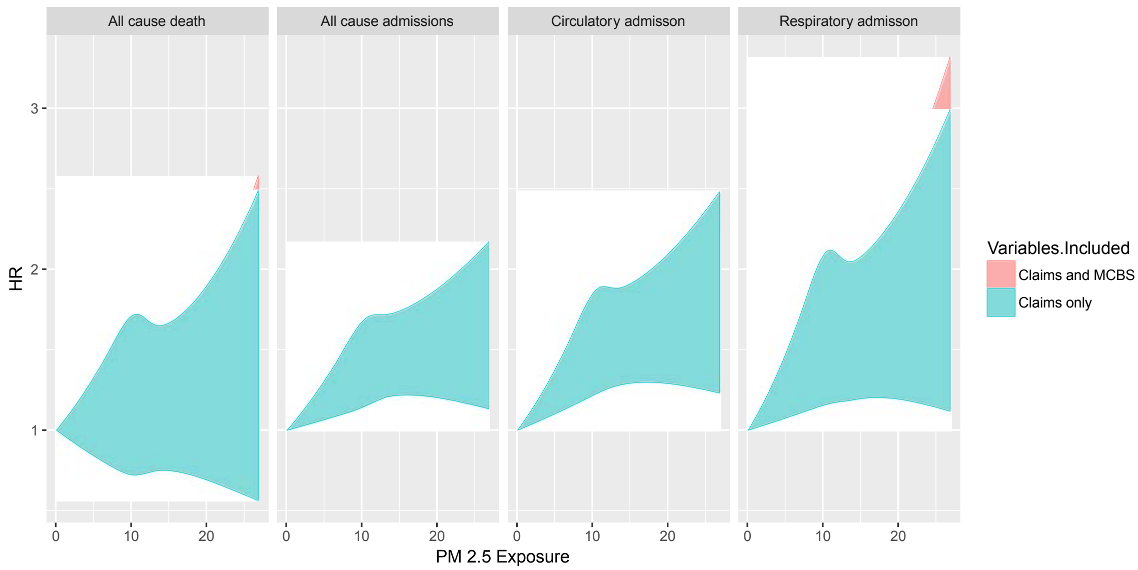
Table S4 shows the p-value of the Wald test for nested models testing as whether globally there is evidence of non-linearity in the ER function. P-values suggest that there might be a nonlinear relationship between exposure to $PM_{2.5}$ and circulatory hospitalizations in the low pollution cohort. Further investigations of this relationship are left for future work (Results for SA2).

Table S4: Wald test for the significance of the spline parameters

	Low pollution cohort	Full cohort
All cause mortality	0.3636	0.6730
All cause hospitalization	0.0128	0.2712

Circulatory hospitalization	0.0019	0.0358
Respiratory hospitalization	0.3481	0.1034

Figure S3: Shows the exposure-response curve for $PM_{2.5}$ and the four outcomes looked at. The nonlinear curve was fit using splines with 3 degrees of freedom. (Results of SA2)



SA3: Low pollution cohort using cutoff of 10 $\mu g/m^3$. This analysis again restricts subjects to those living in areas lower than 12 $\mu g/m^3$, but now defines the binary exposure to be an indicator whether a subject lives in an area with average pollution levels below 10 $\mu g/m^3$ instead than 8 $\mu g/m^3$ as in the main analysis.

Table S5: Effect estimates for low pollution cohort using the exposure to be an indicator that $PM_{2.5}$ is below 10 $\mu g/m^3$.

	Low pollution cohort using WHO cutoff, N = 18,144 person years = 34,429
All cause mortality	1.09 (0.98, 1.22)
All cause hospitalization	1.03 (0.98, 1.08)
Circulatory hospitalization	1.03 (0.98, 1.09)
Respiratory hospitalization	1.04 (0.97, 1.12)

Figure S4: shows the estimated hazard ratios and 95% CI obtained by fitting a Cox Proportional Hazard Model (CPH) with a binary exposure and confounding adjustment done via inverse probability weighting. Estimates in blue are obtained by excluding from the PS model the MCBS variables. (Results of SA3)