



APPENDIX AVAILABLE ON REQUEST

Special Report

Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality

Part I: Replication and Validation

Appendix F. Computer Programs and Output Used in the Replication of the Original Analyses of the American Cancer Society Study

Daniel Krewski, Richard T. Burnett, Mark S. Goldberg, Kristin Hoover, Jack Siemiatycki, Michal Abrahamowicz, Warren H. White, and Others

Correspondence may be addressed to Dr Daniel Krewski, Professor of Epidemiology and Statistics, Department of Epidemiology and Community Medicine, Room 3229C, 451 Smyth Road, University of Ottawa, Ottawa Ontario K1H 8M5, Canada

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

This document was reviewed by the HEI Health Review Committee but did not undergo the HEI scientific editing and production process.

Table of Contents

Program 1 was used to confirm the results summarized in Table 1 of the ACS study (Pope et al. 1995). Table 1 features a broad range of characteristics used to profile the study cohort. These included age, sex, race, smoking status, occupational exposure, education, body mass index and alcohol consumption. Two separate indices of exposure to combustion source particulate air pollution were used with sulfate particles and another with fine particles. The analysis with sulfate particles involved subjects from 151 different metropolitan areas, while the analysis with fine particles involved subjects from 50 metropolitan areas.

Output 1a (Original Data)

Output 1b Modified Data

Program 2 was used to confirm the results summarized in Table 2 of the ACS study (Pope et al. 1995). It reported the adjusted mortality risk ratios and 95% CI by cause of death for current smokers and for two separate indices of exposure to combustion source particulate air pollution.

Output 2a (Original Data)

Output 2b Modified Data

Program 3 was used in the reanalysis to validate the adjusted mortality risk ratios (and 95% CI) which Pope et al. (1995) used to accommodate an inconsistency which they found in the data concerning lung cancer mortality. To identify the source of the inconsistency, they restricted their analysis to 47 metropolitan areas that had both sulfate and fine particle emission data.

Output 3a (Original Data)

Output 3b Modified Data

Program 4 was used to reanalyze the data in Table 2 for the current smoker category. In this analysis, two new variables were entered into the equation. These were "25 years of smoking" and "20 cigarettes per day".

Output 4a (Original Data)

Output 4b Modified Data

Program 5 was used to confirm the results summarized in Table 3 of the ACS study for fine particles (Pope et al. 1995). It calculates the adjusted mortality rates and 95% CI for the most polluted areas compared to the least polluted areas in terms of all-cause and cardiopulmonary deaths. Mortality risk ratios were separated by gender and smoking status and were adjusted for age, sex, race, cigarette smoking, exposure to passive cigarette smoke, body-mass index, alcohol consumption, education and occupational exposure.

Output 5a (Original Data)

Output 5b Modified Data

Program 6 was used to incorporate weather indicator variables (dcold and dhot) into the models in order to validate the findings of the Original Investigators.

Output 6a (Original Data)

Output 6b Modified Data

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by All Cause of Death for the Sulfate Particles Never-smokers

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CENALL
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	219244.365	219077.445	166.920 with 6 DF (p=0.0001)
Wald	.	.	167.776 with 6 DF (p=0.0001)
			167.884 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.001421	0.00376	0.14275	0.7056
EDULOW	1	0.223972	0.02099	113.88590	0.0001
INDUSEXP	1	0.007887	0.02706	0.08492	0.7707
BMI	1	0.006155	0.00223	7.59857	0.0058
ALC	1	-0.033732	0.00777	18.85208	0.0001
SULFATES	1	0.163417	0.05101	10.26195	0.0014

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.001	0.994	1.009	Passive Smoking
EDULOW	1.251	1.201	1.304	Less than high school education
INDUSEXP	1.008	0.956	1.063	Occupational exposure
BMI	1.006	1.002	1.011	Body Mass Index
ALC	0.967	0.952	0.982	Alcohol Drinking
SULFATES	1.178	1.065	1.301	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by All Cause of Death for the Sulfate Particles Never-smokers in Women

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CENALL
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	141276.522	141183.527	92.995 with 6 DF (p=0.0001)
Wald	.	.	90.834 with 6 DF (p=0.0001)
			91.254 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.003857	0.00465	0.68698	0.4072
EDULOW	1	0.186555	0.02595	51.68901	0.0001
INDUSEXP	1	0.058579	0.04411	1.76388	0.1841
BMI	1	0.002916	0.00260	1.25533	0.2625
ALC	1	-0.056983	0.01354	17.71837	0.0001
SULFATES	1	0.184809	0.06543	7.97702	0.0047

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.004	0.995	1.013	Passive Smoking
EDULOW	1.205	1.145	1.268	Less than high school education
INDUSEXP	1.060	0.973	1.156	Occupational exposure
BMI	1.003	0.998	1.008	Body Mass Index
ALC	0.945	0.920	0.970	Alcohol Drinking
SULFATES	1.203	1.058	1.368	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by All Cause of Death for the Sulfate Particles Never-smokers in Men

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CENALL
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	77967.843	77873.348	94.495 with 6 DF (p=0.0001)
Wald	.	.	97.913 with 6 DF (p=0.0001)
			97.544 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	-0.003587	0.00639	0.31523	0.5745
EDULOW	1	0.295769	0.03556	69.16637	0.0001
INDUSEXP	1	-0.028119	0.03419	0.67642	0.4108
BMI	1	0.015555	0.00440	12.48400	0.0004
ALC	1	-0.021560	0.00924	5.44518	0.0196
SULFATES	1	0.133129	0.08157	2.66374	0.1027

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	0.996	0.984	1.009	Passive Smoking
EDULOW	1.344	1.254	1.441	Less than high school education
INDUSEXP	0.972	0.909	1.040	Occupational exposure
BMI	1.016	1.007	1.024	Body Mass Index
ALC	0.979	0.961	0.997	Alcohol Drinking
SULFATES	1.142	0.974	1.340	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Lung Cancer Related Death for the Sulfate Particles Never-smokers
The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CEN62
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	4462.097	4439.410	22.688 with 6 DF (p=0.0009)
Wald	.	.	22.462 with 6 DF (p=0.0010)
			22.636 with 6 DF (p=0.0009)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.004722	0.02475	0.03639	0.8487
EDULOW	1	0.355026	0.15496	5.24916	0.0220
INDUSEXP	1	0.360148	0.17482	4.24408	0.0394
BMI	1	-0.061575	0.01722	12.78509	0.0003
ALC	1	-0.031798	0.05490	0.33550	0.5624
SULFATES	1	0.412453	0.36806	1.25575	0.2625

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.005	0.957	1.055	Passive Smoking
EDULOW	1.426	1.053	1.932	Less than high school education
INDUSEXP	1.434	1.018	2.019	Occupational exposure
BMI	0.940	0.909	0.973	Body Mass Index
ALC	0.969	0.870	1.079	Alcohol Drinking
SULFATES	1.511	0.734	3.108	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Lung Cancer Related Death for the Sulfate Particles Never-smokers in Women

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CEN62
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	3128.366	3114.715	13.652 with 6 DF (p=0.0338)
Wald	.	.	12.901 with 6 DF (p=0.0446)
			13.097 with 6 DF (p=0.0415)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	-0.001676	0.03036	0.00305	0.9560
EDULOW	1	0.234628	0.19173	1.49751	0.2211
INDUSEXP	1	0.231220	0.27282	0.71830	0.3967
BMI	1	-0.065333	0.02004	10.63228	0.0011
ALC	1	-0.011611	0.07314	0.02520	0.8739
SULFATES	1	0.474699	0.45429	1.09186	0.2961

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	0.998	0.941	1.060	Passive Smoking
EDULOW	1.264	0.868	1.841	Less than high school education
INDUSEXP	1.260	0.738	2.151	Occupational exposure
BMI	0.937	0.901	0.974	Body Mass Index
ALC	0.988	0.856	1.141	Alcohol Drinking
SULFATES	1.608	0.660	3.916	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Lung Cancer Related Death for the Sulfate Particles Never-smokers in Men

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CEN62
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	1333.731	1322.377	11.354 with 6 DF (p=0.0780)
Wald	.	.	12.068 with 6 DF (p=0.0605)
			11.901 with 6 DF (p=0.0642)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.019937	0.04286	0.21635	0.6418
EDULOW	1	0.592010	0.26458	5.00666	0.0252
INDUSEXP	1	0.437232	0.23413	3.48737	0.0618
BMI	1	-0.047427	0.03429	1.91356	0.1666
ALC	1	-0.056325	0.08302	0.46034	0.4975
SULFATES	1	0.307581	0.62806	0.23983	0.6243

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.020	0.938	1.110	Passive Smoking
EDULOW	1.808	1.076	3.036	Less than high school education
INDUSEXP	1.548	0.979	2.450	Occupational exposure
BMI	0.954	0.892	1.020	Body Mass Index
ALC	0.945	0.803	1.112	Alcohol Drinking
SULFATES	1.360	0.397	4.658	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Cardiopulmonary Death for the Sulfate Particles Never-smokers

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CENCOMB
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	106712.082	106535.194	176.888 with 6 DF (p=0.0001)
Wald	.	.	176.766 with 6 DF (p=0.0001)
			177.133 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.010235	0.00560	3.33441	0.0678
EDULOW	1	0.289471	0.02810	106.14105	0.0001
INDUSEXP	1	-0.037928	0.03891	0.95015	0.3297
BMI	1	0.010172	0.00318	10.23480	0.0014
ALC	1	-0.051334	0.01184	18.80280	0.0001
SULFATES	1	0.318727	0.07150	19.87153	0.0001

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.010	0.999	1.021	Passive Smoking
EDULOW	1.336	1.264	1.411	Less than high school education
INDUSEXP	0.963	0.892	1.039	Occupational exposure
BMI	1.010	1.004	1.017	Body Mass Index
ALC	0.950	0.928	0.972	Alcohol Drinking
SULFATES	1.375	1.196	1.582	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Cardiopulmonary Death for the Sulfate Particles Never-smokers in Women

The PHREG Procedure

Data Set: WORK.SULF_N
 Dependent Variable: FAIL
 Censoring Variable: CENCOMB
 Censoring Value(s): 1
 Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	64190.517	64076.973	113.544 with 6 DF (p=0.0001)
Wald	.	.	107.832 with 6 DF (p=0.0001)
			109.147 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.014617	0.00731	3.99725	0.0456
EDULOW	1	0.261147	0.03523	54.93577	0.0001
INDUSEXP	1	-0.002777	0.07018	0.00157	0.9684
BMI	1	0.002921	0.00381	0.58786	0.4432
ALC	1	-0.112273	0.02378	22.29917	0.0001
SULFATES	1	0.373289	0.09464	15.55726	0.0001

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.015	1.000	1.029	Passive Smoking
EDULOW	1.298	1.212	1.391	Less than high school education
INDUSEXP	0.997	0.869	1.144	Occupational exposure
BMI	1.003	0.995	1.010	Body Mass Index
ALC	0.894	0.853	0.936	Alcohol Drinking
SULFATES	1.453	1.207	1.749	Sulfate Particles

Table3: Adjusted Mortality Risk Ratios (and 95% Confidence Intervals) by Cardiopulmonary Death for the Sulfate Particles Never-smokers in Men
The PHREG Procedure

Data Set: WORK.SULF_N
Dependent Variable: FAIL
Censoring Variable: CENCOMB
Censoring Value(s): 1
Ties Handling: BRESLOW

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L Score	42521.565	42431.999	89.566 with 6 DF (p=0.0001)
Wald	.	.	92.950 with 6 DF (p=0.0001)
			92.675 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
CURCIG	0	0	.	.	.
EVPCONLY	0	0	.	.	.
SMKCPD	0	0	.	.	.
XSMKCPD	0	0	.	.	.
SMKCYR	0	0	.	.	.
XSMKCYR	0	0	.	.	.
PASSIVE	1	0.003690	0.00873	0.17848	0.6727
EDULOW	1	0.335106	0.04635	52.26692	0.0001
INDUSEXP	1	-0.058560	0.04675	1.56875	0.2104
BMI	1	0.027536	0.00587	22.01570	0.0001
ALC	1	-0.027828	0.01297	4.60070	0.0320
SULFATES	1	0.242659	0.10931	4.92785	0.0264

Analysis of Maximum Likelihood Estimates

Conditional Risk Ratio and 95% Confidence Limits

Variable	Risk Ratio	Lower	Upper	Label
CURCIG	.	.	.	Current Smoker
EVPCONLY	.	.	.	Pipe/cigar smoker
SMKCPD	.	.	.	Current cigarettes per day
XSMKCPD	.	.	.	Former cigarettes per day
SMKCYR	.	.	.	Current years smoke
XSMKCYR	.	.	.	Former years smoked
PASSIVE	1.004	0.987	1.021	Passive Smoking
EDULOW	1.398	1.277	1.531	Less than high school education
INDUSEXP	0.943	0.861	1.034	Occupational exposure
BMI	1.028	1.016	1.040	Body Mass Index
ALC	0.973	0.948	0.998	Alcohol Drinking
SULFATES	1.275	1.029	1.579	Sulfate Particles