**ONLINE SUPPLEMENT**

**Title:** Particulate Air Pollution and Liver Cancer Survival

**Authors:** Huiyu Deng, MA,1 Sandrah P Eckel, PhD,1 Lihua Liu, PhD,1Frederick W. Lurmann, MS,2 Myles Cockburn, PhD,1 and Frank D Gilliland, MD PhD1

**Contents:**

eTable 1. Summary of demographic, clinical, treatment characteristics, and distances of residential addresses at diagnosis from highways and PM2.5 exposures for liver cancer patients newly diagnosed in California from 2000-2009, by PM2.5 concentration levels.

eTable 2. Adjusted hazard ratios for all-cause mortality and liver cancer specific survival from pooled and stage-specific Cox proportional hazard models.

eTable 3. Sensitivity analysis for PM2.5: adjusted all-cause mortality hazard ratios (and 95% confidence intervals) associated with a standard deviation (SD) increase in PM2.5 exposure, from models stratified by stage at diagnosis (see Table 2) but with additional stratification by factors of interest.

Information about the socioeconomic variable (SES).Page 2

Page 3

Page 3

Page 4

**eTable 1**. Summary of demographic, clinical, treatment characteristics, and distances of residential addresses at diagnosis from highways and PM2.5 exposures for liver cancer patients newly diagnosed in California from 2000-2009, by PM2.5 concentration levelsa.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Characteristics (Mean ± SD or %) | Lown=5,364 (26.5%) | Mediumn=7,431 (36.7%) | Highn=5,668 (28.0%) | Missingbn=1,758(8.7%) | Totaln=20,221 |
| Age at diagnosis (years) | 63.6±12.2 | 63.8±12.3 | 64.0±12.8 | 62.9±12.1 | 63.7±12.4 |
| Male  | 76.0 | 74.2 | 74.0 | 78.8 | 75.0 |
| Race (%) |
|  Non-Hispanic whites | 45.8 | 36.6 | 29.0 | 59.8 | 38.9 |
| Hispanic | 19.0 | 26.4 | 32.9 | 22.0 | 25.9 |
| Non-Hispanic blacks | 7.2 | 7.2 | 9.9 | 5.1 | 7.8 |
| Asian/Pacific islanders | 26.5 | 28.7 | 27.5 | 9.0 | 26.1 |
| Others/Unknown | 1.5 | 1.0 | 0.6 | 4.0 | 1.3 |
| Marital status (%) |
| Single | 19.3 | 18.1 | 18.0 | 22.2 | 18.7 |
| Married | 56.2 | 55.6 | 56.6 | 49.3 | 55.5 |
| Formerly married | 22.4 | 23.5 | 23.3 | 24.6 | 23.3 |
|  Unknown | 2.1 | 2.7 | 2.2 | 3.9 | 2.5 |
| Rural-urban commuting area (%) |
| Metropolitan core | 90.2 | 95.8 | 97.0 | 49.1 | 90.6 |
| Non-metropolitan core | 9.8 | 4.2 | 3.0 | 50.9 | 9.4 |
| Socioeconomic status (SES, %) |
| Lowest | 13.1 | 19.4 | 30.1 | 19.4 | 20.7 |
| Lower-middle | 17.7 | 21.5 | 24.9 | 27.6 | 22.0 |
| Middle | 22.1 | 20.7 | 19.1 | 23.1 | 20.8 |
| Higher-middle | 23.8 | 20.5 | 15.5 | 17.2 | 19.7 |
| Highest | 22.6 | 17.1 | 10.3 | 12.1 | 16.2 |
| Unknown | 0.7 | 0.8 | 0.1 | 0.6 | 0.6 |
| Stage of diagnosis  |
|  Local | 47.5 | 48.3 | 38.6 | 42.4 | 44.8 |
|  Regional  | 28.1 | 27.0 | 28.0 | 27.0 | 27.5 |
|  Distant | 16.2 | 16.0 | 20.9 | 18.3 | 17.6 |
|  Unknown c | 8.2 | 8.8 | 12.5 | 12.3 | 10.0 |
| First-course treatment types (%) |
|  Surgery | 24.6 | 25.9 | 15.3 | 18.7 | 21.9 |
|  Radiation | 3.0 | 3.4 | 3.6 | 2.7 | 3.3 |
|  Chemotherapy | 37.0 | 33.9 | 27.7 | 27.6 | 32.4 |
| Geocode match quality (%) |
| Street address  | 92.0 | 93.7 | 93.5 | 73.1 | 91.4 |
| Area-level  | 8.0 | 6.3 | 6.5 | 26.5 | 8.6 |
|  Other or missing | <0.1 | <0.1 | <0.1 | 0.4 | <0.1 |
| Distance to primary interstate highwayb |
|  < 300 m | 10.8 | 11.3 | 10.2 | 3.8 | 10.2 |
|  300 – 1500 m  | 38.5 | 41.8 | 43.0 | 12.5 | 38.7 |
|  >1500 m  | 42.7 | 40.6 | 40.4 | 46.0 | 41.5 |
|  % missing | 8.0 | 6.3 | 6.5 | 37.7 | 9.5 |
| Distance to primary US and State highwaysd |
|  < 300 m | 4.9 | 3.4 | 4.0 | 5.5 | 4.1 |
|  300 – 1500 m  | 17.9 | 12.9 | 14.1 | 10.6 | 14.4 |
|  >1500 m  | 69.2 | 77.4 | 75.5 | 46.1 | 71.9 |
|  % missing | 8.0 | 6.3 | 6.5 | 37.7 | 9.5 |

a Percentage of patients with exposure assignment available (requires a monitor for that pollutant ≤25 km from residential address and non-missing geocode).

b The categories of PM2.5 exposure were 0-10, 10-15, 15+µg/m3.

c Insufficient evidence available to assign a stage (e.g., patient dies before workup is complete, patient refuses diagnostic procedure, or limited workup is performed due to patient’s age or simultaneous contraindicating condition).

d Distance values are primarily missing for participants with poor geocode matches (worse than street address match).

 **eTable 2**. Adjusted hazard ratios for all-cause mortality and liver cancer specific survival from pooled and stage-specific Cox proportional hazard models.

|  |
| --- |
|  Hazard Ratio (95% confidence interval)c |
|  |  | Local | Regional | Distant | Overall |
| All-cause mortality | Base modela | 1.41 (1.37 – 1.45) | 1.24 (1.20 – 1.28) | 1.02 (0.99 – 1.06) | 1.28 (1.26 – 1.31) |
| Fully adjusted modelb | 1.31 (1.26 – 1.35) | 1.19 (1.14 – 1.23) | 1.05 (1.01 – 1.10) | 1.18 (1.16 – 1.20) |
| Liver cancer specific survival | Base modela | 1.37 (1.32 – 1.42) | 1.24 (1.19 – 1.29) | 1.03 (0.99 – 1.07) | 1.27 (1.24 – 1.30) |
| Fully adjusted modelb | 1.24 (1.19 – 1.30) | 1.19 (1.13 – 1.24) | 1.05 (1.00 – 1.10) | 1.15 (1.12 – 1.18) |

a Adjusted for age, sex

b Adjusted for age, sex, race/ethnicity, marital status, socioeconomic status, rural–urban commuting area, distance to primary interstate highway, distance to primary US and state highways, month of diagnosis, year of diagnosis and initial treatments.
c Hazard ratios are scaled to a 1 standard deviation increase in PM2.5 (equivalent to 5.0 µg/m3)

**eTable 3**. Sensitivity analysis for PM2.5: adjusteda all-cause mortality hazard ratios (and 95% confidence intervals) associated with a standard deviation (SD)b increase in PM2.5 exposure, from models stratified by stage at diagnosis (see Table 2) but with additional stratification by factors of interest.

|  |  |  |  |
| --- | --- | --- | --- |
| Stratifying factor | Local | Regional | Distant |
| Distance to closest air quality monitor |  |  |  |
| <5 km | 1.25 (1.17, 1.34) | 1.12 (1.05, 1.20) | 1.05 (0.97, 1.13) |
| 5-25 km | 1.33 (1.27, 1.38) | 1.22 (1.17, 1.28) | 1.05 (0.99, 1.10) |
| Highest quality geocode match (street-address) | 1.32 (1.27, 1.37) | 1.18 (1.14, 1.23) | 1.05 (1.01, 1.10) |
| Socioeconomic status (SES) |  |  |  |
|  Lowest | 1.39 (1.28, 1.51) | 1.18 (1.09, 1.27) | 1.04 (0.96, 1.13) |
|  Lower-middle | 1.24 (1.16, 1.33) | 1.14 (1.05, 1.24) | 1.06 (0.96, 1.16) |
|  Middle | 1.38 (1.28, 1.48) | 1.18 (1.08, 1.28) | 1.06 (0.96, 1.16) |
|  Higher-middle | 1.22 (1.13, 1.33) | 1.21 (1.10, 1.33) | 1.05 (0.96, 1.16) |
|  Highest | 1.46 (1.32, 1.61) | 1.30 (1.16, 1.45) | 1.02 (0.90, 1.15) |
| Regions in California |  |  |  |
|  Los Angeles County | 2.90 (2.60, 3.23) | 1.88 (1.66, 2.12) | 1.17 (1.02, 1.33) |
|  San Francisco Bay Area | 1.52 (1.21, 1.88) | 1.61 (1.31, 1.97) | 1.08 (0.89, 1.31) |
|  San Diego County | 2.79 (2.15, 3.61) | 2.05 (1.55, 2.70) | 1.35 (0.92, 1.99) |
|  Other  | 1.29 (1.23, 1.35) | 1.13 (1.07, 1.20) | 1.02 (0.96, 1.09) |

a Adjusted for age, sex, race/ethnicity, marital status, socioeconomic status, rural–urban commuting area, distance to primary interstate highway, distance to primary US and state highways, month of diagnosis, year of diagnosis and initial treatments.
b Hazard ratios are scaled to a 1 standard deviation increase in PM2.5 (equivalent to 5.0 µg/m3)

**Information about the socioeconomic variable (SES):**

The area-level SES indicator developed by the CCR1-3 is a composite measure at the census block group level, based on census 2000 data and American Community Survey 2007-2011 5-year estimates, created through a principal components analysis that included the following census variables:

* Proportion with a blue-collar job,
* Proportion older than 16 years in the workforce without a job,
* Median household income,
* Percent below 200% poverty level,
* Median gross rent,
* Median value of owner-occupied houses, and
* Average years of education for individuals 25 years of age and older.

The SES composite score was calculated for all census block groups (n = 22,960) and were then ranked into quintiles resulting in five SES groups.

**REFERENCES**

1. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control 2001;12: 703-11
2. Yin D, Morris C, Allen M, Cress R, Bates J, Liu L. Does socioeconomic disparity in cancer incidence vary across racial/ethnic groups? Cancer Causes Control 2010;21: 1721-30.
3. Yang J, Schupp C, Harrati A, Clarke C, Keegan T, Gomez S, Developing an area-based socioeconomic measure from American Community Survey data. Cancer Prevention Institute of California, 2014.