

## ORIGINAL ARTICLE

## Air pollution affects lung cancer survival

Sandrah P Eckel,<sup>1</sup> Myles Cockburn,<sup>1</sup> Yu-Hsiang Shu,<sup>1,2</sup> Huiyu Deng,<sup>1</sup> Frederick W Lurmann,<sup>3</sup> Lihua Liu,<sup>1</sup> Frank D Gilliland<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2015-207927>).

<sup>1</sup>Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA

<sup>2</sup>Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, California, USA

<sup>3</sup>Sonoma Technology Inc., Petaluma, California, USA

**Correspondence to**

Dr Sandrah P Eckel, Department of Preventive Medicine, University of Southern California, 2001 N. Soto Street, MC-9234, Los Angeles, CA 90089, USA; [eckel@usc.edu](mailto:eckel@usc.edu)

Received 12 October 2015

Revised 2 May 2016

Accepted 19 May 2016

Published Online First

4 August 2016

**ABSTRACT**

**Rationale** Exposure to ambient air pollutants has been associated with increased lung cancer incidence and mortality, but due to the high case fatality rate, little is known about the impacts of air pollution exposures on survival after diagnosis. This study aimed to determine whether ambient air pollutant exposures are associated with the survival of patients with lung cancer.

**Methods** Participants were 352 053 patients with newly diagnosed lung cancer during 1988–2009 in California, ascertained by the California Cancer Registry. Average residential ambient air pollutant concentrations were estimated for each participant's follow-up period. Cox proportional hazards models were used to estimate HRs relating air pollutant exposures to all-cause mortality overall and stratified by stage (localised only, regional and distant site) and histology (squamous cell carcinoma, adenocarcinoma, small cell carcinoma, large cell carcinoma and others) at diagnosis, adjusting for potential individual and area-level confounders.

**Results** Adjusting for histology and other potential confounders, the HRs associated with 1 SD increases in NO<sub>2</sub>, O<sub>3</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> for patients with localised stage at diagnosis were 1.30 (95% CI 1.28 to 1.32), 1.04 (95% CI 1.02 to 1.05), 1.26 (95% CI 1.25 to 1.28) and 1.38 (95% CI 1.35 to 1.41), respectively. Adjusted HRs were smaller in later stages and varied by histological type within stage ( $p < 0.01$ , except O<sub>3</sub>). The largest associations were for patients with early-stage non-small cell cancers, particularly adenocarcinomas.

**Conclusions** These epidemiological findings support the hypothesis that air pollution exposures after lung cancer diagnosis shorten survival. Future studies should evaluate the impacts of exposure reduction.

**INTRODUCTION**

Lung cancer has been the most common cancer for decades. Worldwide, lung cancer causes nearly one in five cancer deaths, about 1.59 million deaths annually (<http://globocan.iarc.fr>). This heavy burden is largely a result of a high prevalence of cigarette smoking, the leading cause of lung cancer; advanced stage at diagnosis; and poor survival, especially among those with advanced stage disease.<sup>1–2</sup> Accordingly, interventions have focused on the reduction of tobacco use, early-stage diagnosis and improved treatment. Although progress has been made in each area, lung cancer survival remains stubbornly poor, suggesting that novel approaches are needed.<sup>3–6</sup> A promising approach is identifying and intervening on modifiable determinants of survival; however, little research attention has been directed to determinants beyond smoking.

**Key messages****What is the key question?**

- Does exposure to ambient air pollution after diagnosis of lung cancer affect survival?

**What is the bottom line?**

- Patients with lung cancer, with higher average ambient NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> exposures since diagnosis, had shorter survival, with the largest differences in survival for patients with early-stage non-small cell cancers (particularly adenocarcinomas).

**Why read on?**

- This is the first study to link individual-level estimates of air pollution exposures after lung cancer diagnosis to survival, and the study population was the population-based sample of 352 053 patients with newly diagnosed lung cancer during 1988–2009 in California, as ascertained by the California Cancer Registry.

One modifiable determinant of emerging interest is ambient air pollution,<sup>7</sup> which was recently classified as carcinogenic by the International Agency for Research on Cancer (IARC).<sup>8</sup>

A growing body of evidence indicates that ambient air pollutants are associated with lung cancer incidence and mortality.<sup>9–12</sup> However, relatively little is known about air pollution effects on survival after any cancer diagnosis.<sup>13–14</sup> We reasoned that if ambient air pollution is a carcinogen affecting lung cancer development, then inhaled pollutants may also drive tumour progression through the same mechanistic pathways to shorten survival after diagnosis. If ambient air pollution increases both the incidence of lung cancer and shortens survival after diagnosis, then it could have a larger contribution to lung cancer mortality than previously understood.

To determine whether ambient air pollutants are associated with survival in patients with lung cancer, we conducted a population-based cohort study of 352 053 California residents with lung cancer newly diagnosed in 1988–2009. We estimated average residential exposures to nitrogen dioxide, ozone and particulate matter air pollutants from diagnosis to end of follow-up and related these exposures to all-cause mortality and lung cancer-specific mortality by stage and tumour histology at diagnosis.



► <http://dx.doi.org/10.1136/thoraxjnl-2016-208967>



**To cite:** Eckel SP, Cockburn M, Shu Y-H, et al. *Thorax* 2016;**71**:891–898.

## METHODS

### Lung cancer cases and individual-level data

Our study population included lung cancer cases (ICD-O-3 site code C34) diagnosed in 1988–2009 and registered by the California Cancer Registry (CCR), the statewide population-based cancer surveillance system (<http://www.ccrca.org>). The CCR has collected information on all newly diagnosed cancers (except non-melanoma skin cancer) in California since 1988 and has received the highest level of data quality certification from the North American Association of Central Cancer Registries.<sup>15</sup> The CCR gathers individual-level data on demographics (age, sex, marital status, race/ethnicity), date of diagnosis, tumour characteristics at diagnosis (stage, anatomical site, histology), treatment occurring <6 months after diagnosis and patients' vital status (date of death or date last known alive). The CCR routinely updates patients' vital status by linking to the electronic death certificate master file from the California Department of Public Health, recording the underlying cause of death for deceased patients, as defined by the Department of Health Services. After excluding patients with diagnoses of in situ cancer (N=305) and non-carcinoma histology (N=20 964), there were 352 053 cases remaining for analysis with complete information on follow-up, date of birth, date of diagnosis and race/ethnicity. We created standard histology groupings<sup>16</sup> using ICD-O-3 morphology codes for carcinoma (8010–8576): squamous cell carcinoma (8050–8078, 8083–8084), adenocarcinoma (8140, 8211, 8230–8231, 8250–8260, 8323, 8480–8490, 8550–8551, 8570–8574, 8576), small cell carcinoma (8041–8045, 8246), large cell—includes giant cell, clear cell and large cell undifferentiated—carcinoma (8010–8012, 8014–8031, 8035, 8310) and other carcinomas (remaining codes).

No patient contact was conducted for this analysis; so, individual patient-informed consent was not required. The CCR operates under the annual review of the State of California Committee for the Protection of Human Subjects (ie, Institutional Review Board), which provided approval for this analysis.

### Geocoding

We geocoded residential addresses at the date of diagnosis using the Texas A&M geocoding service ([geoservices.tamu.edu](http://geoservices.tamu.edu)). Details of the process, used by cancer registries throughout the USA, are provided elsewhere.<sup>17</sup> Briefly, address records were geocoded to the centroid of the smallest resolvable area based on the address completeness, ranging from tax assessor parcels to state centroid when no address information was available (in <0.1% of cases). In previous work, this method substantially improved spatial resolution.<sup>18</sup>

### Area-level covariates

Area-level covariates were assigned to each patient using the geocodes. Rural–urban commuting area (RUCA) codes, based on data from the 2000 decennial census, classify census tracts into ordinal ranks (1–10, from metropolitan to rural) based on the size and direction of primary commuting flows, using measures of population density, urbanisation and daily commuting (<http://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx>).

Education index was defined as the average years of schooling in the patient's census tract based on a weighted sum of the distribution of years of schooling.<sup>19</sup> Socioeconomic status, at the patient's census block group, was calculated using validated area-level measures from multiple census years.<sup>20</sup>

### Air pollution exposure assignments

California air pollutant data were obtained from the US Environmental Protection Agency's (EPA) Air Quality System (AQS) database.<sup>21</sup> Data were available for nitrogen dioxide (NO<sub>2</sub>, ppb), ozone (O<sub>3</sub>, ppb), particulate matter with diameter <10 µm (PM<sub>10</sub>, µg/m<sup>3</sup>) and 2.5 µm (PM<sub>2.5</sub>, µg/m<sup>3</sup>). Hourly measurements were summarised as 24-hour averages for NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> and average 8 hour daily maximum for O<sub>3</sub>. Monthly average concentrations were spatially interpolated to residence locations from up to four closest air quality monitoring stations within a 50 km radius using the well-established method of inverse distance weighting,<sup>22 23</sup> with the decay parameter equal to the inverse of the square of the distance of the residence from each monitoring site. Interpolation performance is summarised in online supplementary eTable 1. We excluded exposure assignments when the nearest monitor was located >25 km away or a geocode match was unavailable. Residential ambient air pollution exposure assignments were calculated as the average of the patient-level interpolated monthly concentrations from the date of diagnosis to the date of last follow-up or death. PM<sub>2.5</sub> exposure assignments were only available for patients diagnosed in 1998 and later, because routine monitoring did not start until 1998. Our primary goal was to evaluate associations with large-scale regional variation in ambient pollutants; so, to account for potential confounding by local traffic, we calculated and adjusted for the distance from residential address to primary interstate highways and primary US and state highways.

### Outcome

Survival time was calculated from the date of newly diagnosed lung cancer to the date of death from any cause (all-cause mortality). For sensitivity analysis, we assessed time to death from an underlying cause of lung cancer (ICD-9 code 1629 for 1988–1998 deaths and ICD-10 code C349 deaths after 1998). The last day of follow-up was 31 December 2011.

### Statistical analysis

Descriptive statistics were calculated for survival, air pollution exposures and other covariates. Median survival and 5-year survival rates were calculated after stratifying patients by the stage at diagnosis and categorised air pollution exposures (cut-offs: 25th and 75th percentiles, and average of the two). Cox proportional hazards models were used to model time to all-cause mortality, allowing for right censoring due to loss to follow-up (or study end) or, in the case of time to lung cancer mortality, censoring due to death from another underlying cause. Preliminary analyses determined that the following potential confounding variables were predictors of mortality; so, all models adjusted for age, sex, race/ethnicity (non-Hispanic white, Hispanic, non-Hispanic black, other/unknown), marital status (single, married, formerly married, unknown), education index (quartiles), socioeconomic status (quintiles), dichotomised rural–urban community area (metropolitan core (1), non-metropolitan core (>1)), categorised distance to primary interstate highway (<300 m, 300–1500 m, >1500 m, missing), categorised distance to primary US and state highways (<300 m, 300–1500 m, >1500 m, missing), categorised year of diagnosis (1988–1992, 1993–1997, 1998–2002, 2003–2009), calendar month of diagnosis and initial treatment (surgery, radiation and/or chemotherapy vs none).

Single pollutant models were used to estimate HRs associated with a 1 SD increase in continuous air pollution exposure, after

**Table 1** Demographic, tumour and treatment characteristics of patients with lung cancer in California by stage of diagnosis, 1988–2009

Characteristics (mean±SD or %)	Localised only (n=59 609)	Regional (n=73 513)	Distant site(s) (n=186 496)	Unknown* (n=32 435)	Total (n=352 053)
Age (years)	69.9±10.5	68.8±10.5	68.7±11.3	72.5±10.7	69.3±11.0
% Male	49.8	54.7	56.1	54.8	54.6
Race/ethnicity, %					
Non-Hispanic white	81.0	78.7	75.0	79.2	77.2
Hispanic	6.5	7.1	8.8	7.4	7.9
Non-Hispanic black	6.2	7.2	7.9	6.9	7.4
Other/unknown	6.2	7.0	8.3	6.5	7.5
Marital status, %					
Single	9.4	9.6	11.9	9.6	10.8
Married	56.2	57.1	53.7	47.9	54.3
Formerly married	32.5	31.5	32.1	37.6	32.6
Unknown	1.8	1.8	2.3	4.9	2.4
Education index†, %					
Low	22.3	23.6	25.6	29.6	25.0
Low–medium	24.5	25.0	24.8	27.4	25.0
Medium–high	25.4	25.4	24.9	24.2	25.0
High	27.9	26.1	24.7	18.8	25.0
Rural–urban commuting area (RUCA), %					
Non-metropolitan core	13.8	14.1	13.8	18.7	14.3
Metropolitan core	86.2	85.9	86.2	81.3	85.7
Unknown	<0.1	<0.1	<0.1	<0.1	<0.1
Socioeconomic status (SES), %					
Lowest	14.0	15.0	16.6	18.8	16.1
Lower-middle	20.3	20.7	20.9	23.6	21.0
Middle	21.6	22.1	22.0	22.2	22.0
Higher-middle	21.8	21.3	20.8	19.8	21.0
Highest	21.1	19.6	18.2	15.2	18.7
Unknown	1.3	1.3	1.4	0.5	1.3
Year of diagnosis, %					
1988–1992	22.2	22.4	19.8	38.8	22.5
1993–1997	23.1	22.3	21.6	29.0	22.7
1998–2002	23.0	22.9	23.9	18.7	23.0
2003–2009‡	31.7	32.4	34.7	13.6	31.8
Histology at diagnosis, %					
Squamous cell	26.1	27.2	15.2	22.3	20.2
Adenocarcinoma	45.0	35.4	35.4	23.1	35.9
Small cell	5.7	13.2	18.1	13.8	14.6
Large cell	12.1	14.3	20.8	35.7	19.3
Other	11.2	9.9	10.5	5.2	10.0
Treatment types					
Surgery, %					
No	32.5	59.6	94.6	92.8	76.6
Yes	67.4	40.3	5.2	4.5	23.0
Unknown	0.1	0.1	0.2	2.7	0.4
Radiation, %					
No	81.3	52.1	56.7	72.8	61.4
Yes	18.7	47.9	43.3	26.8	38.6
Unknown	<0.1	<0.1	<0.1	0.5	0.1
Chemotherapy, %					
No	87.0	59.2	55.1	76.2	63.3
Yes	11.3	38.0	41.2	19.8	33.5
Unknown	1.6	2.8	3.7	4.0	3.2
Geocode match quality, %					
Street address match	91.5	91.5	91.5	89.1	91.3
Area-level match	8.5	8.5	8.5	10.8	8.7
Other or missing	<0.1	<0.1	<0.1	<0.1	<0.1
Median survival time (years)	3.6	1.3	0.4	0.7	0.7
Median survival time (years), by histology at diagnosis					
Squamous cell	2.6	1.1	0.4	0.7	0.8

Continued

Table 1 Continued

Characteristics (mean±SD or %)	Localised only (n=59 609)	Regional (n=73 513)	Distant site(s) (n=186 496)	Unknown* (n=32 435)	Total (n=352 053)
Adenocarcinoma	5.1	1.9	0.4	0.8	0.9
Small cell	1.5	1.0	0.4	0.8	0.6
Large cell	1.6	0.8	0.2	0.6	0.4
Other	6.1	1.3	0.4	1.0	0.8

\*Insufficient evidence available to assign a stage (eg, patient dies before workup is complete, patient refuses diagnostic procedure or limited workup is performed due to patient's age or simultaneous contraindicating condition).

†Categorised by quartiles.

‡Most recent year of diagnosis category includes 7 years, while the other categories each include 5 years.

adjusting for the aforementioned covariates. Initial models also adjusted for stage and histology at diagnosis. We then evaluated evidence for modification of air pollution associations by stage and then by histology. Final single pollutant models were fit separately for each stage and histology. Sensitivity analyses were performed by further stratifying stage-specific models (adjusted for histology) by sex, race/ethnicity, year of diagnosis, metropolitan core group and large urban areas (Los Angeles County, Bay Area Counties and San Diego County).

Analyses were performed using SAS V9.4 (SAS Institute). Select graphical displays were created using R V3.1.3 (R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2012). Hypothesis tests were two-sided, with a 0.05 type I error rate.

## RESULTS

Characteristics of the study participants are presented in table 1. Patients were on average 69.3 years old at diagnosis, predominantly non-Hispanic white (77.2%), and most of them lived in a metropolitan core (85.7%). More than half of lung cancers were diagnosed at an advanced stage (53.0% distant site). During the study period, there were 324 266 deaths (92.1% of 352 053 patients). Of these deaths, 78.3% were assigned an underlying cause of lung cancer. Median survival times for localised, regional and distant site diagnoses were 3.6, 1.3 and 0.4 years, respectively. For patients with localised stage at diagnosis, median survival was shortest for patients with small and large cell carcinomas (1.5 and 1.6 years, respectively) and longest for patients with adenocarcinoma (5.1 years). The number of patients with 'unknown' stage at diagnosis decreased from 12 573 in 1988–1992 (5-year period) to 4399 in 2003–2009 (7-year period), likely due to changes in technology,

medical practice and/or coding practices. The highest quality geocode match (street address) was obtained for 91.1% of patients.

Average air pollution exposure assignments (average from diagnosis to end of follow-up for each patient  $\leq 25$  km from a monitor) across patients were 21.9 ppb for NO<sub>2</sub>, 40.2 ppb for O<sub>3</sub>, 31.8  $\mu\text{g}/\text{m}^3$  for PM<sub>10</sub> and 13.7  $\mu\text{g}/\text{m}^3$  for PM<sub>2.5</sub> (table 2). As expected, PM<sub>10</sub>, PM<sub>2.5</sub> and NO<sub>2</sub> were more highly correlated (Pearson's R=0.70–0.76) than O<sub>3</sub> and PM<sub>10</sub> (R=0.36), O<sub>3</sub> and NO<sub>2</sub> (R=-0.01) or O<sub>3</sub> and PM<sub>2.5</sub> (R=-0.02). Over the study period, there were long-term downward trends in NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> in California (see online supplementary eFigure 1 and eTable 2). Only 8.7% of patients lived  $< 300$  m from a primary interstate highway, while 45.4% lived  $> 1500$  m (see online supplementary eTable 3).

We observed a pattern of shorter median survival and lower 5-year survival for patients with local or regional stage at diagnosis who had higher categorised NO<sub>2</sub>, PM<sub>10</sub> or PM<sub>2.5</sub> exposures (table 3). For example, median survival for patients with local stage at diagnosis was 2.4 years for those with high PM<sub>2.5</sub> exposure ( $\geq 16 \mu\text{g}/\text{m}^3$ ) and 5.7 years for those with low PM<sub>2.5</sub> exposure ( $< 10 \mu\text{g}/\text{m}^3$ ). Survival for patients with distant stage at diagnosis was poor and showed little variation with air pollution exposure.

After adjusting for important determinants of survival and potential confounders (including stage and histology), the HRs for all-cause mortality associated with a 1 SD increase in each pollutant were 1.13 (95% CI 1.12 to 1.13) for NO<sub>2</sub>, 1.02 (95% CI 1.02 to 1.03) for O<sub>3</sub>, 1.11 (95% CI 1.11 to 1.12) for PM<sub>10</sub> and 1.16 (95% CI 1.16 to 1.17) for PM<sub>2.5</sub> (see online supplementary eTable 4). As shown in table 4, these associations varied by stage at diagnosis (all interaction  $p < 0.01$ ) and were of similar magnitude whether considering all-cause mortality or

Table 2 Air pollution exposure assignments based on residence at diagnosis, by stage at diagnosis

Air pollution exposures (mean±SD or %)	Localised only (n=59 609)	Regional (n=73 513)	Distant site(s) (n=186 496)	Unknown (n=32 435)	Total (n=352 053)
NO <sub>2</sub> (ppb)	20.6±9.3	21.4±9.7	22.0±10.4	24.5±11.3	21.9±10.2
% available*	87.6	86.8	87.3	83.0	86.8
O <sub>3</sub> (ppb)	40.3±9.7	40.3±10.8	40.0±12.8	41.2±12.8	40.2±11.9
% available*	93.6	93.2	93.3	89.9	93.0
PM <sub>10</sub> ( $\mu\text{g}/\text{m}^3$ )	30.5±10.7	31.4±11.6	31.9±12.4	35.0±13.5	31.8±12.1
% available*	91.7	91.3	91.4	88.6	91.2
PM <sub>2.5</sub> ( $\mu\text{g}/\text{m}^3$ )†	13.0±4.5	13.4±4.9	13.9±5.6	14.6±5.7	13.7±5.3
% available*	86.5	84.3	82.6	76.8	83.3

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

†PM<sub>2.5</sub> data are reported only for the subset of patients whose cancer was diagnosed in 1998 or later.

PM<sub>2.5</sub>, particulate matter with diameter  $< 2.5 \mu\text{m}$ ; PM<sub>10</sub>, particulate matter with diameter  $< 10 \mu\text{m}$ .

**Table 3** Median survival and 5-year survival rate, by stage at diagnosis and air pollution exposure

Categorised air pollution exposure	Median survival (years)			Five-year survival rate (%) <sup>*</sup>		
	Localised	Regional	Distant	Localised	Regional	Distant
<b>NO<sub>2</sub> (ppb)</b>						
<14	5.4	1.6	0.3	50	24	4
14–20.5	4.2	1.4	0.4	45	21	4
20.5–27	3.2	1.2	0.4	38	17	3
≥27	2.2	1.0	0.3	30	12	2
<b>O<sub>3</sub> (ppb)</b>						
<32	2.8	1.0	0.3	36	16	2
32–39.5	3.8	1.4	0.5	42	19	3
39.5–47	4.9	1.6	0.5	49	23	4
≥47	2.7	1.1	0.3	35	15	2
<b>PM<sub>10</sub> (μg/m<sup>3</sup>)</b>						
<23	4.7	1.5	0.3	47	23	4
23–30.5	4.4	1.4	0.4	45	20	4
30.5–38	3.7	1.3	0.4	43	19	3
≥38	2.1	1.0	0.3	27	11	2
<b>PM<sub>2.5</sub> (μg/m<sup>3</sup>)<sup>†</sup></b>						
<10	5.7	1.9	0.3	51	27	4
10–13	5.0	1.9	0.5	48	25	5
13–16	4.5	1.5	0.5	46	23	4
≥16	2.4	1.2	0.3	31	14	2

<sup>\*</sup>SEs of all 5-year survival rates are <1%, with calculations based on >5800 patients per group.

<sup>†</sup>PM<sub>2.5</sub> results are only for the subset of patients whose cancer was diagnosed in 1998 or later.

PM<sub>2.5</sub>, particulate matter with diameter <2.5 μm; PM<sub>10</sub>, particulate matter with diameter <10 μm.

lung cancer-specific mortality. For each pollutant, adjusted HRs were larger for patients diagnosed at early stages. After stratifying by stage, we found additional variation in the associations by histology (all interaction  $p < 0.01$ , except O<sub>3</sub>). After stratifying by stage and histology, exposure to NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> remained strongly associated with all-cause mortality, with the largest magnitude-adjusted HR for local stage (figure 1). The adjusted HRs for NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> were generally smaller for patients with small cell carcinoma and larger for patients with adenocarcinoma (eg, local stage HR for PM<sub>10</sub>: 1.16 (95% CI 1.11 to 1.21) vs 1.30 (95% CI 1.28 to 1.33), respectively). O<sub>3</sub> was not statistically significantly associated with all-cause mortality for patients with small and large cell cancer, but was modestly associated for patients with squamous cell carcinoma and adenocarcinoma (local stage adjusted HR of 1.04 (95% CI 1.02 to 1.07) and 1.03 (95% CI 1.01 to 1.05), respectively). Dose–response associations were evaluated in adjusted Cox models with categorised air pollution exposures, stratified by stage (data not shown). Results were qualitatively similar to the unadjusted associations in table 3.

In sensitivity analyses, no substantial heterogeneity in stage-specific adjusted HR was found by sex, race/ethnicity or distance to air quality monitors (see online supplementary eTables 5a, b). There was modest heterogeneity by year of diagnosis, particularly for NO<sub>2</sub> and PM<sub>10</sub>, but the patterns of larger HR for patients diagnosed at earlier stages remained consistent. Patients with local stage at diagnosis living in a metropolitan core had slightly higher HR for PM<sub>10</sub> and PM<sub>2.5</sub> than those living in non-metropolitan core areas (eg, PM<sub>2.5</sub> HR of 1.40 vs 1.25), a pattern that was also observed in the subsets of patients

diagnosed in Los Angeles County, the San Francisco Bay area or San Diego County. These findings merit further study.

## DISCUSSION

While ambient air pollutants have been associated with lung cancer incidence and mortality,<sup>7–11</sup> their impacts on survival after diagnosis have yet to be fully assessed.<sup>14</sup> In a population-based study of 352 053 patients with newly diagnosed lung cancer in California, we observed reduced survival associated with higher average NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> exposure over the follow-up period after diagnosis. HRs associated with these pollutants were largest for early-stage cancers and varied by histology, with the largest HR in early-stage non-small cell cancers, particularly adenocarcinoma.

A growing number of large cohort studies have found evidence for associations between air pollution exposures and lung cancer mortality using either incident lung cancer or death from lung cancer.<sup>9–12</sup> Meta-analysis estimates of the relative risk of lung cancer incidence/death (not stratified by stage) were slightly smaller than those observed in our study (1.04 (95% CI 1.01 to 1.08) for a 10 ppb increase in NO<sub>2</sub>,<sup>11</sup> 1.08 (95% CI 1.00 to 1.17) for a 10 μg/m<sup>3</sup> increase in PM<sub>10</sub><sup>10</sup> and 1.04 (95% CI 1.02 to 1.07) for a 5 μg/m<sup>3</sup> increase in PM<sub>2.5</sub>) and showed some evidence for heterogeneity by histology.<sup>10</sup> For the two most common histologies, relative risks associated with a 5 μg/m<sup>3</sup> increase in PM<sub>2.5</sub> were 1.18 (95% CI 1.03 to 1.35) for adenocarcinoma and 1.05 (95% CI 0.85 to 1.31) for squamous cell carcinoma.<sup>10</sup>

Few studies have attempted to disentangle determinants of lung cancer incidence from determinants of lung cancer survival due to the high case fatality rate.<sup>10–24</sup> To our knowledge, only one study has related air pollution exposures to survival in patients diagnosed with lung cancer.<sup>14</sup> Xu *et al*<sup>14</sup> studied white patients with respiratory cancer in Honolulu and Los Angeles between 1992–2008 and found adjusted HR slightly larger than that we observed for all-cause mortality (1.48 (95% CI 1.44 to 1.52) for a 10 μg/m<sup>3</sup> increase in PM<sub>10</sub>; 1.57 (95% CI 1.53 to 1.61) for a 5 μg/m<sup>3</sup> increase in PM<sub>2.5</sub>; 1.04 (95% CI 1.03 to 1.06) for a 10 ppb change in O<sub>3</sub>) and slightly larger PM associations when restricting the analysis to Los Angeles cases only. Key differences include that we interpolated ambient exposures to residence locations (rather than using county-level exposures) and that we considered only lung cancer cases stratified by stage and histology. Xu *et al* considered all respiratory cancer cases and adjusted for primary cancer site and stage. By fully conditioning on disease type and severity at diagnosis, we more effectively target inference about air pollution exposure associations with survival after diagnosis by limiting carryover effects from differences at diagnosis potentially caused by earlier air pollution exposures.

Our observed associations were clinically significant (≤38% increased risk of death depending on stage and pollutant), suggesting that reductions in exposure have the potential to improve lung cancer survival. As expected, we observed a substantially larger association with survival in local compared with distant stage at diagnosis. As lung cancer screening becomes widely implemented, a shift to diagnosis at earlier stages is likely to occur. This is the stage at which air pollutants appear to have the most impact on survival. To maximise the effectiveness of lung cancer screening, interventions targeting modifiable determinants of survival for early-stage diagnoses are needed. Our findings suggest that future work should investigate the impact of interventions to reduce air pollution exposures (eg, avoidance, relocation, home filtration systems) on lung cancer survival.

**Table 4** Adjusted\* HRs (95% CI) for all-cause and lung cancer mortality associated with 1 SD increase in air pollutant exposure,† stratified by stage at diagnosis

Air pollutant	Stage at diagnosis	Sample size	All-cause mortality HR (95% CI)	Lung cancer mortality HR (95% CI)
NO <sub>2</sub>	Localised only	52 223	1.30 (1.28 to 1.32)	1.31 (1.29 to 1.33)
	Regional	63 777	1.18 (1.17 to 1.20)	1.18 (1.16 to 1.19)
	Distant site(s)	162 816	1.07 (1.07 to 1.08)‡	1.07 (1.06 to 1.08)
	Overall§	305 721	1.13 (1.12 to 1.13)‡	1.12 (1.11 to 1.12)‡
O <sub>3</sub>	Localised only	55 823	1.04 (1.02 to 1.05)	1.05 (1.04 to 1.07)
	Regional	68 504	1.03 (1.02 to 1.04)	1.03 (1.02 to 1.05)
	Distant site(s)	174 022	1.01 (1.01 to 1.02)‡	1.02 (1.01 to 1.02)‡
	Overall§	327 513	1.02 (1.02 to 1.03)‡	1.03 (1.02 to 1.03)‡
PM <sub>10</sub>	Localised only	54 671	1.26 (1.25 to 1.28)	1.27 (1.25 to 1.29)
	Regional	67 108	1.16 (1.15 to 1.17)	1.15 (1.14 to 1.17)
	Distant site(s)	170 415	1.07 (1.06 to 1.07)‡	1.07 (1.06 to 1.07)‡
	Overall§	320 940	1.11 (1.11 to 1.12)‡	1.11 (1.10 to 1.11)‡
PM <sub>2.5</sub> ¶	Localised only	28 212	1.38 (1.35 to 1.41)	1.39 (1.36 to 1.43)
	Regional	34 223	1.26 (1.24 to 1.28)	1.24 (1.22 to 1.27)
	Distant site(s)	90 243	1.10 (1.09 to 1.11)‡	1.10 (1.09 to 1.11)
	Overall§	160 707	1.16 (1.16 to 1.17)‡	1.15 (1.14 to 1.16)‡

\*Adjusted for age, sex, race/ethnicity, marital status, education index, socioeconomic status, rural–urban commuting area, distance to primary interstate highway, distance to primary US and state highways, histology at diagnosis, month of diagnosis, year of diagnosis and initial treatment.

†SD values: 10.2 ppb for NO<sub>2</sub>, 11.9 ppb for O<sub>3</sub>, 12.1 µg/m<sup>3</sup> for PM<sub>10</sub> and 5.3 µg/m<sup>3</sup> for PM<sub>2.5</sub>.

‡Estimate and CI bounds appear identical due to rounding.

§Overall analyses do not stratify by stage, but adjust for stage and include patients with unknown stage.

¶PM<sub>2.5</sub> results are only for the subset of patients whose cancer was diagnosed in 1998 or later.

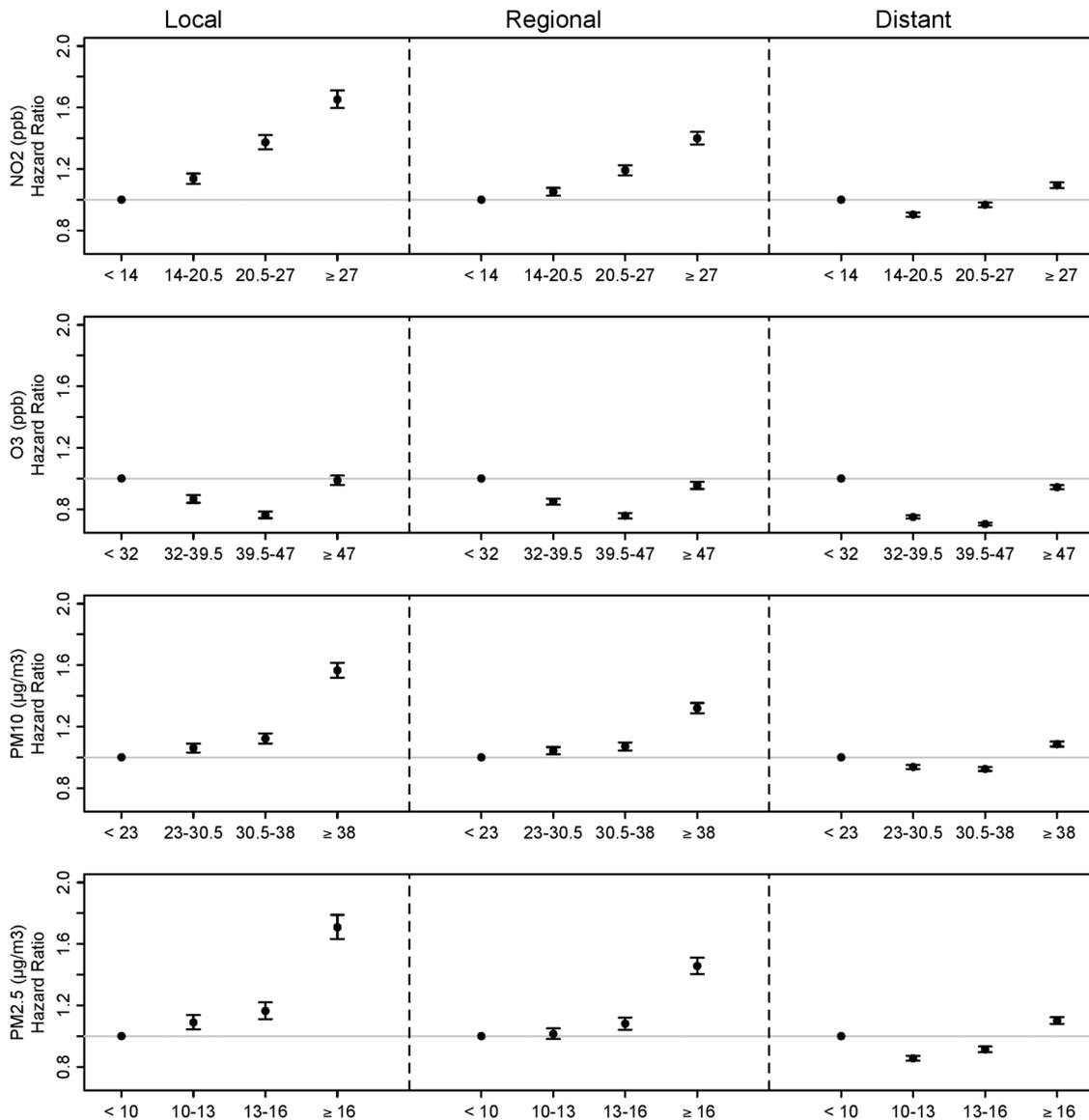
PM<sub>2.5</sub>, particulate matter with diameter <2.5 µm; PM<sub>10</sub>, particulate matter with diameter <10 µm.

The pathophysiological mechanism underlying the relationship between NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> and lung cancer survival is uncertain. Ambient air pollution has been classified as a carcinogen and therefore may affect cancer progression after diagnosis via the same well-described pathways including oxidative stress, DNA damage, cell proliferation or epigenetic modifications. We observed some of the largest air pollution HRs for adenocarcinoma, the only common histological subtype of lung cancer that develops in a significant number of non-smokers.<sup>25 26</sup> More generally, air pollution may reduce survival in the susceptible subpopulation of patients with cancer, for example, by impairing respiratory function.

Strengths of our study include the population-based, large sample size drawn from all cases diagnosed in California, minimising selection bias and avoiding the survivorship bias in standard cohort studies. Using standardised methods, the CCR collects detailed clinical data and individual-level information on important determinants of survival (histology, stage, age and year of diagnosis; first course of treatment, sex, race/ethnicity and marital status). Our study focused on California, which has one of the most extensive and longest running air quality monitoring networks in the USA.

Several limitations of our study should be considered. The CCR collects information only on first course treatments, but residual confounding by subsequent treatments is unlikely since treatment is determined primarily by stage at diagnosis, and we stratify by stage. Follow-up in the CCR is passive, but nearly complete (>95%) for cancers with short survival. Individualised residential ambient air pollution exposure assignments offer a refinement over area-level exposure assignments (eg, reducing spatial exposure misclassification, which can attenuate associations),<sup>10 27 28</sup> but are subject to standard limitations, including inability to account for individual behaviour (eg, patients with cancer may spend even more time indoors than the general population), changes of residence or potentially long periods of time at medical facilities located in an area with different air pollution levels. We focused on air pollution exposures with

large-scale regional variability using spatial interpolation of air quality monitoring data, which does not capture the effects of traffic-related pollution (TRP) that varies over a finer spatial scale. We accounted for potential confounding by a crude measure of local traffic (distance to highways). Future investigation of the effects of TRP on lung cancer survival requires the development of high spatially resolved TRP exposure metrics (eg, using land-use regression or line-source dispersion models) to directly evaluate TRP associations. The air pollution monitoring network is less dense in rural areas; so, exclusion of patients living >25 km from a monitor differentially excludes patients in rural areas. Long-term downward trends in NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> in California during the study period have been recognised previously.<sup>29</sup> The lack of consistent long-term temporal trend for O<sub>3</sub> likely reduced the variability in O<sub>3</sub> exposure across participants. Note that because survival is relatively short in patients with lung cancer, we expected short-term (seasonal) variability to dominate long-term variability during each patient's follow-up period. We adjusted for month of diagnosis in our models to account for potential confounding by short-term temporal factors. Results were robust to sensitivity analyses stratifying by categorised year of diagnosis, suggesting that long-term trends did not induce spurious associations (particularly of concern for early-stage diagnosis adenocarcinoma cases with longer median survival). Finally, we lacked individual-level data on important potential confounders/effect modifiers and risk factors (eg, smoking, diet, alcohol use, education, access to care, obesity, previous lung disease and occupational exposures). These omitted factors could have spuriously induced the observed associations only if they were strongly associated with the spatiotemporal distribution of ambient air pollution exposures, which seems unlikely. Previous studies have suggested that non-smokers may be at greater risk for air pollution-related lung cancer incidence/mortality than current smokers.<sup>10</sup> While smoking is an important risk factor, previous data suggest that, at diagnosis, only 39% of patients with lung cancer are current smokers (drops to 14% at 5 months after diagnosis).<sup>30</sup>



**Figure 1** Adjusted<sup>a</sup> HRs and 95% CIs for all-cause mortality associated with 1 SD increase in air pollutant exposure,<sup>b,c</sup> stratified by stage and histology at diagnosis. <sup>a</sup>Adjusted for age, sex, race/ethnicity, marital status, education index, 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 treatment. <sup>b</sup>SD values: 10.2 ppb for NO<sub>2</sub>, 11.9 ppb for O<sub>3</sub>, 12.1 µg/m<sup>3</sup> for PM<sub>10</sub> (particulate matter with diameter <10 µm) and 5.3 µg/m<sup>3</sup> for PM<sub>2.5</sub> (particulate matter with diameter <2.5 µm). <sup>c</sup>PM<sub>2.5</sub> results are only for the subset of patients whose cancer was diagnosed in 1998 or later.

In summary, we found evidence for associations between all-cause and lung cancer-specific mortalities and NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>, robust to a number of sensitivity analyses. Future studies should evaluate the impacts of ambient air pollution exposure reduction, since controlling patients' exposures could offer a novel approach to improve lung cancer outcomes, especially among patients diagnosed at early stages.

**Contributors** Study concept, design and obtained funding: MC and FDG; acquisition, analysis or interpretation of data and critical revision of the manuscript for important intellectual content: all authors; drafting of the manuscript: SPE, MC, FDG and Y-HS; statistical analysis: Y-HS and HD; administrative, technical or material support: MC, LL and FWL; study supervision: MC, FDG and SPE.

**Funding** This work was supported by the Southern California Environmental Health Sciences Center (grant 5P30ES007048) funded by the National Institute of Environmental Health Sciences; the Hastings Foundation; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California and

contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement U58DP003862-01 awarded to the California Department of Public Health.

**Disclaimer** The funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication. The ideas and opinions expressed herein are those of the author(s), and endorsement by the State of California, Department of Public Health, the National Cancer Institute and the Centers for Disease Control and Prevention or their contractors and subcontractors is neither intended nor should be inferred.

**Competing interests** None declared.

**Ethics approval** State of California Committee for the Protection of Human Subjects.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Schottenfeld D, Fraumeni J. *Cancer epidemiology and prevention*. New York: Oxford University Press, 2006.

- 2 Alberg AJ, Brock MV, Samet JM. Epidemiology of lung cancer: looking to the future. *J Clin Oncol* 2005;23:3175–85.
- 3 Kanodra NM, Silvestri GA, Tanner NT. Screening and early detection efforts in lung cancer. *Cancer* 2015;121:1347–56.
- 4 Pilkington G, Boland A, Brown T, et al. A systematic review of the clinical effectiveness of first-line chemotherapy for adult patients with locally advanced or metastatic non-small cell lung cancer. *Thorax* 2015;70:359–67.
- 5 Rosell R, Karachaliou N. Lung cancer in 2014: optimizing lung cancer treatment approaches. *Nat Rev Clin Oncol* 2015;12:75–6.
- 6 Tanoue LT, Tanner NT, Gould MK, et al. Lung cancer screening. *Am J Respir Crit Care Med* 2015;191:19–33.
- 7 Fajersztajn L, Veras M, Barrozo LV, et al. Air pollution: a potentially modifiable risk factor for lung cancer. *Nat Rev Cancer* 2013;13:674–8.
- 8 Loomis D, Huang W, Chen G. The International Agency for Research on Cancer (IARC) evaluation of the carcinogenicity of outdoor air pollution: focus on China. *Chin J Cancer* 2014;33:189–96.
- 9 Burnett RT, Pope CA III, Ezzati M, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ Health Perspect* 2014;122:397–403.
- 10 Hamra GB, Guha N, Cohen A, et al. Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environ Health Perspect* 2014;122:906–11.
- 11 Hamra GB, Laden F, Cohen AJ, et al. Lung cancer and exposure to nitrogen dioxide and traffic: a systematic review and meta-analysis. *Environ Health Perspect* 2015;123:1107–12.
- 12 Fischer PH, Marra M, Ameling CB, et al. Air pollution and mortality in seven million adults: the Dutch Environmental Longitudinal Study (DUELS). *Environ Health Perspect* 2015;123:697–704.
- 13 Hu H, Dailey AB, Kan H, et al. The effect of atmospheric particulate matter on survival of breast cancer among US females. *Breast Cancer Res Treat* 2013;139:217–26.
- 14 Xu X, Ha S, Kan H, et al. Health effects of air pollution on length of respiratory cancer survival. *BMC Public Health* 2013;13:800.
- 15 Tucker T, Howe H, Weir H. Certification for population-based cancer registries. *J Registry Manag* 1999;26:24–7.
- 16 Egevad L, Heanue M, Berney D. Histological groups. In: Curado M, Edwards B, Shin H, et al., eds. *Cancer incidence in five continents. Volume IX. IARC Scientific Publications No. 160*. Lyon, France: IARC Press, 2007:61–6.
- 17 Goldberg DW, Cockburn MG. Improving geocode accuracy with candidate selection criteria. *Trans GIS* 2010;14(Suppl 1):149–76.
- 18 Goldberg DW, Cockburn MG. The effect of administrative boundaries and geocoding error on cancer rates in California. *Spat Spatiotemporal Epidemiol* 2012;3:39–54.
- 19 Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). *Cancer Causes Control* 1998;9:369–80.
- 20 Yost K, Perkins C, Cohen R, et al. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control* 2001;12:703–11.
- 21 US Environmental Protection Agency. Air Quality System Data Mart [internet database]. <http://www.epa.gov/ttn/airs/aqsdatamart> (accessed 1 Oct 2012).
- 22 Wong DW, Yuan L, Perlin SA. Comparison of spatial interpolation methods for the estimation of air quality data. *J Expo Sci Environ Epidemiol* 2004;14:404–15.
- 23 Rivera-González LO, Zhang Z, Sánchez BN, et al. An assessment of air pollutant exposure methods in Mexico City, Mexico. *J Air Waste Manag Assoc* 2015;65:581–91.
- 24 Turner MC, Krewski D, Pope III CA, et al. Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. *Am J Respir Crit Care Med* 2011;184:1374–81.
- 25 Raaschou-Nielsen O, Andersen ZJ, Beelen R, et al. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet Oncol* 2013;14:813–22.
- 26 Schuller HM. Mechanisms of smoking-related lung and pancreatic adenocarcinoma development. *Nat Rev Cancer* 2002;2:455–63.
- 27 Jerrett M, Burnett R, Beckerman BS, et al. Spatial analysis of air pollution and mortality in California. *Am J Respir Crit Care Med* 2013;188:593.
- 28 Zeger SL, Thomas D, Dominici F, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000;108:419.
- 29 Lurmann F, Avol E, Gilliland F. Emissions reduction policies and recent trends in Southern California's ambient air quality. *J Air Waste Manage* 2015;65:324–35.
- 30 Park ER, Japuntich SJ, Rigotti NA, et al. A snapshot of smokers after lung and colorectal cancer diagnosis. *Cancer* 2012;118:3153–64.