SCOPE OF WORK

☒ Contract ☐ Grant

Does this project include Research (as defined in the UTC)? ☒ Yes ☐ No

PI Name: Jason G. Su

Project Title: Impacts of Air Pollution on Life Expectancy across Multiple Generations: Race, Ethnicity, and Vulnerability Perspectives

Project Summary/Abstracta1

It is well established that air pollution is linked to numerous adverse health outcomes. Many studies show disparate air pollution related health impacts in communities due to factors such as proximity to air pollutions sources, age, race/ethnicity, and income. The goal of this study is to build upon previous work by examining statewide air pollution exposure and life expectancy disparities across generations and within communities. The impact of air pollution on life expectancy over two generations in California, especially for vulnerable populations is not well studied. The University of California, Berkeley (UCB) will acquire statewide Medi-Cal population data for the years 1990-2020 for this research. The Medi-Cal population are followed by the California Department of Health Care Services (DHCS) monthly and have the most stable continuous enrollments, generally across multiple generations for a family. The dates of death will be used to estimate the degree to which life expectancies were impacted by air pollution from fine particulate matter ($PM_{2.5}$), separately, for (1) the first and second generations of the entire Medi-Cal population, (2) race-ethnicity and vulnerability subgroups of each generation, and (3) the first and second generations within families. UCB hypothesizes that, due to improvements in air quality, overall air pollution-specific impacts on life expectancy improved from the first to the second generation, however, certain areas and groups may not have experienced the same improvements. Therefore, UCB aims to investigate which communities and groups over two generations continued to experience the greatest disparities in $PM_{2.5}$ exposure and air pollution-related impacts to life expectancy.

To reduce exposure misclassification, UCB will apply daily surfaces of $PM_{2.5}$ at 100 meters (m) spatial resolution that is being developed for the State for the years 1989-2020 (creating 1989 daily surfaces for rolling average exposure for death in 1990). The home address of an enrollee will be assigned daily $PM_{2.5}$ exposure and one-year rolling average exposure to the date of death. The death of an enrollee will be matched through propensity score by two survival enrollees who had similar individual and neighborhood characteristics except air pollution. The same time span and length of exposure will also be estimated for the matched survival enrollees. The Medi-Cal population will be separated into two periods: 1990-2005 and 2006-2020. When comparing the same age group between the two periods, the 1990-2005 population will be identified as first generation and, subsequently, the 2006-2020 population for the second generation.

To identify the impact of air pollution on life expectancy, the years of life expectancy lost due to air pollution will be modeled through differentiating life expectancies estimated with all cause and cause $PM_{2.5}$-eliminated mortality information (i.e., cause $PM_{2.5}$-eliminated mortality = all cause mortality – cause $PM_{2.5}$ mortality). Life expectancies on all cause and cause $PM_{2.5}$ eliminated cohorts will be estimated through the traditional life table method and area under the curve function. The impact of $PM_{2.5}$ on mortality will be modeled through a series of age-group specific logistic regression models that estimate the causal impact of air pollution on mortality. The difference in air pollution specific life expectancy loss between the first and second generation will then be identified. These differences will be identified not only for all the Medi-Cal enrollees, but also for race-ethnicity and vulnerability subgroups.

Numerous studies identified that vulnerable communities historically experienced and continue to experience the highest air pollution and the greatest health burden. The air pollution level from 1989-2020 will be aggregated at the census tract (CT) level to identify if similar exposure trends exist. Air pollution exposure hotspots using the top 25 percent most occurring CTs with the greatest $PM_{2.5}$ for a generation will be created and overlaid with CT level vulnerability (identified through CalEnviroScreen[1]) to identify exposure disparities in vulnerable communities. Further, CT level race-ethnicity composition will be used as weights to adjust statewide life expectancy loss estimated for the race-ethnicity subgroups to create CT level life expectancy loss. Similarly, the CT level life expectancy loss will be overlaid with CT vulnerability to identify life expectancy disparities in vulnerable communities.

The Regional Asthma Management & Prevention (RAMP) will take the lead on the community outreach of the project. RAMP will host two webinars at the beginning of the project to inform stakeholders about the research and to solicit feedback on the scope and direction. RAMP will also host two webinars toward the end of the project to share the research results. UCB will work with RAMP to prepare webinar materials that are in lay language. The community outreach will not only serve the purpose of presenting research ideas and results to communities through accessible languages but will also gain feedback on what areas of the research findings communities are most interested in and what form the research findings should take such as mapping of CT level air pollution exposure and life expectancy disparities in vulnerable communities.

Overall, this study will provide CARB with information on the impacts of $PM_{2.5}$ exposure on life expectancy in the first generation and the second generation not only for California residents overall, but also for race-ethnicity and vulnerability subgroups. The subgroups with the greatest reductions in life expectancy due to increased $PM_{2.5}$ exposure will be identified. This study will also identify at the CT level the communities that consistently experienced both the highest air pollution exposure and had the greatest life expectancy loss due to air pollution. Further, the study will provide research findings on changes in life expectancy from air pollution from the first generation to the second generation within families.

If Third-Party Confidential Information is to be provided by the State:

☐ Performance of the Scope of Work is anticipated to involve use of third-party Confidential Information and is subject to the terms of this Agreement; ***OR***

☒ A separate CNDA between the University and third-party is required by the third-party and is incorporated in this Agreement as Exhibit A7.

Statement of Significance

This project will supplement scientific evidence on the impact of PM exposure on life expectancy by supplying spatially refined air pollution surfaces (100 m) and covering the greatest length of study (30 years). The large population-based Schwartz study[2] used Medicare data across several states in the US for 2000-2013, but relied on annual air pollution exposure data at zip code level. The application of 100 m spatial resolution daily PM data to patient's home address reduces spatiotemporal air pollution exposure misclassification and establishes a more accurate relationship between air pollution and life expectancy. The goal of the project aligns with CARB's work to identify environmental exposure disparities in vulnerable communities and disparities in environment-related adverse health outcomes. The design of two time periods for the study allows for identification of possible improvements in air quality and air pollution-specific impacts on life expectancy from the first to the second generation. The application of Medi-Cal data with its high continued enrollments enables this study to be the first to identify the impact of air pollution on life expectancy across two generations and across two generations within families. Using a series of age-group specific logistic regression models that estimate the causal impact of air pollution on mortality ensures factors other than air pollution such as impact of individual medical conditions are removed. The differentiation of all cause life expectancy from cause $PM_{2.5}$-eliminated life expectancy enables us to identify impact of $PM_{2.5}$ exposure on life expectancy for California residents. This research will also add evidence to the literature on impacts of air pollution on life expectancy in racial/ethnic groups and vulnerable communities across a statewide Medi-Cal population, providing new evidence regarding which race-ethnicity groups and vulnerable populations experienced greater air pollution burden and endured greater air pollution-specific impacts to life expectancy when compared to non-Hispanic Whites and non-vulnerable subgroups. Using within family multiple generation participants, this research will also identify the impacts on life expectancy from air pollution in the second generation compared to the first generation within families. The research will also provide CT level life expectancy measures to allow for analysis of disparities at the neighborhood CT level. Through linkage to CT CalEnviroScreen scores and two generations of air pollution exposure, this research will be the first to identify vulnerable communities that experienced the greatest impacts on life expectancy due to air pollution. The community outreach activities will improve awareness of the trends of air pollution exposure in vulnerable communities over two generations and the impacts of air pollution on life expectancy. Outreach with communities will also open a communication channel for addressing inequalities in environmental exposure and life expectancy to help communities increase resilience. This research will augment the California Air Resources Board's (CARB) work to identify air pollution exposure disparities and associated health impacts in vulnerable communities for its regulations, strategies, and programs.

Scope of Work

UCB will collaborate with the Northern California Institute for Research and Education (NCIRE), to conduct a systematic literature review, using peer-reviewed journal papers to identify effects of air pollution exposure on life expectancy with a focus on exposure to $PM_{2.5}$. During the literature review process, UCB and NCIRE will determine inclusion criteria, publication characteristics, search databases, engines, terms, and selection process.

UCB will generate daily air pollution surfaces of $PM_{2.5}$ at 100 m spatial resolution for the years 1989-2020 and assign one year of daily $PM_{2.5}$ exposure to the residential addresses of all the Medi-Cal enrollees before date of death (i.e., rolling average) between 1990 and 2020. UCB and NCIRE will develop age group specific logistic regression models to determine the impact of $PM_{2.5}$ exposure on mortality, separately, for the first generation (1990-2005) and second generation (2006-2020). UCB and NCIRE will generate life expectancies for the all cause and cause $PM_{2.5}$ eliminated mortality cohorts, with the latter constructed from the impact of $PM_{2.5}$ exposure for a series of age-specific logistic regression models.

UCB and NCIRE will estimate total statewide reductions on life expectancy due to air pollution for each generation and statewide for each major race-ethnicity and vulnerability group and within families for both generations. UCB will identify which statewide race-ethnicity and vulnerable subgroups experienced the greatest air pollution burden and greatest air pollution impacts on life expectancy.

To analyze disparities at the CT level, UCB will aggregate the 100 m daily $PM_{2.5}$ concentrations developed for 1989-2020 to generate CT mean annual $PM_{2.5}$ concentrations. UCB will use CT decennial census data as weights to adjust statewide impacts on life expectancy to create CT level life expectancy measures. Using geographic information system (GIS) software, UCB will map and overlay CT level life expectancy measures and CT level $PM_{2.5}$ exposure for 1990, 2000, 2010, 2020 and the difference from the 1st to 2nd generation. UCB will also map vulnerable CTs identified through CalEnviroScreen[2]. The final GIS maps will allow for identification of communities and vulnerable communities that have the highest magnitude and persistence of $PM_{2.5}$ exposure and impacts to life expectancy due to air pollution.

The RAMP will lead community outreach for the project. Specifically, RAMP will host two webinars at the beginning of the project to inform stakeholders about the research and to solicit feedback about its scope and direction. The community outreach will focus on feedback from communities on what areas of the research findings communities are most interested in and the format of the research findings such as the maps and outreach materials. RAMP will also host two webinars toward the end of the project to share the research results and implications. UCB will work with RAMP to prepare webinar materials, including various maps generated from the project, written in lay person's language. RAMP will also translate the materials into Spanish and will offer simultaneous Spanish/English translation during the webinars.

Project Tasks

Task 1. Conduct a Literature Review

UCB and NCIRE will conduct a systematic literature review, using peer-reviewed journal papers to identify effects of air pollution exposure on life expectancy with a focus on exposure to $PM_{2.5}$. The following inter-connected steps will be used to complete the review:

A. Determine inclusion criteria that will include the:

- Study population of children (1-17) and adults (18-95) for both male and female subjects;
- Study intervention for individuals (1) exposed to air pollution, including nitrogen dioxide (NO_2), fine PM with aerodynamic diameter ≤ 2.5 microns ($PM_{2.5}$) and ozone (O_3);
- Study life expectancy and main factors impacting life expectancy including demographics, socioeconomic status, and chronic diseases such as asthma, cardiovascular disease (CVD), chronic obstructive pulmonary disease (COPD), and diabetes;
- Study design for both randomized trials and observational studies (case-control, longitudinal cohort, and cross-sectional); and
- Statistical modeling for linear and mixed linear models, logistic regressions, machine learning algorithms, such as deletion/substitution/addition and random forest modeling.

B. Identify the publications characteristics for studies:

- Published in peer-reviewed journals;
- Published between January 1, 2000 and current;
- Cited in the background section (of this study) that are not duplicate of above search results; and

C. Written in English.Select the proper search databases and engines, including

- PubMed;

- Medline;
- Web of Science Core Collection; and
- Google Scholar

D. Decide the search terms and selection process.

The search category and search terms are listed in Table 1. The following steps will be used to select scientific publications for the literature review:

1) Use one term from each category and combine together (+) to create integrated search terms using the search databases and engines listed in Item C;
2) Merge together the selected publications and remove the duplicates;

3) Obtain abstracts for the remaining publications selected from Step 2), screen and remove the abstracts that are not related;
4) Obtain full text for the remaining publications selected from Step 3), screen and remove the publications that are not related to the topic; and
5) UCB will conduct the literature review with the final selected ones. UCB will synthesize the Review in identifying the degrees of impact from air pollution on life expectancy for subjects with and without comorbidities, and possibly across more than one generation. UCB will compare the performance of various statistical modeling techniques, including machine learning algorithms. UCB will also identify the studies conducted in California and the United States (U.S) that might have close relevance to the study.

Table 1. Literature review categories and search terms.

Category	Search Terms
Outcome	Life expectancy
Air pollution	Air pollution; NO_2; $PM_{2.5}$; O_3; pollutant mixtures; chemical mixture
Other impact factors	demographics (various race-ethnicities), social economic status, vulnerability, chronic diseases (e.g., asthma; COPD; CVD, diabetes)
Statistical models	Linear; Poisson; negative binomial; logistic regression; machine learning; deletion/substitution/addition; and random forest modeling
Years of publication	2000 – current
Publication type	Peer reviewed journals
Publication language	English

Deliverable: UCB and NCIRE will provide a copy of the systematic literature review findings on the impacts of air pollution on life expectancy to CARB. In addition, a copy of the literature review findings will be included as an appendix in the draft final report.

Task 2. Develop $PM_{2.5}$ Surfaces for the Study Population

Under CARB contract 21RD004, UCB is developing daily air pollution models and surfaces for the State for the years 1990-2020 (30 years). UCB will also generate daily $PM_{2.5}$ surfaces for 1989 to facilitate the requirement of estimating rolling average exposure for enrollees died in 1990. In developing daily land use regression (LUR) models for $PM_{2.5}$, UCB first identifies factors (i.e., source or sink) that might impact $PM_{2.5}$ concentrations and use them as potential predictors. UCB also identifies the optimal distance of impact for a potential predictor and the models should be able to deal with multicollinearity among predictors and can reduce model overfit. Further, UCB aims to avoid excessive number of predictors in the final selected model and will allow a maximum of 15 predictors (in addition to four Seasons) in a LUR model. Due to those considerations, UCB applies the Deletion/Substitution/Addition (D/S/A) algorithm for developing a daily prediction model.[3-5] The modeling process is described in detail below.

Development of Comprehensive Data Sources

UCB developed comprehensive data sources that have potential impact on the concentrations measured at California Environmental Protection Agency (CalEPA) monitoring sites. The data sources include daily traffic data, daily remote sensing data, daily weather data, every two-week vegetation index, one time land use and land cover data, and other potential impact factors. UCB hypothesizes that greater daily traffic is associated with higher $PM_{2.5}$ concentrations. Remote sensing Aerosol Optical Depth (AOD) data is an indirect measure of $PM_{2.5}$ concentrations with greater AOD values being directly associated with higher $PM_{2.5}$ concentrations. Different land use types have different impacts on $PM_{2.5}$ concentrations with, for example, higher industrial and commercial land use being associated with greater concentrations. Similarly, different land cover types can have other impacts on $PM_{2.5}$ concentrations with, for example, high intensity urban developed land cover being associated with greater concentrations but greater vegetation cover (as a sink) being associated with lower concentrations. Further, greener vegetation has a much better air pollutant absorption effect than less green vegetation and thus the former helps reduce concentrations. For weather data, greater wind speed is associated with lower concentrations while lower visibility is associated with higher concentrations. UCB is also collecting daily $PM_{2.5}$ concentrations data at the CARB regulatory monitoring sites (i.e., CalEPA sites) for the years 1989-2020 and will use them as a response variable in generating daily $PM_{2.5}$ concentration models. The potential predictors proposed in the study are listed in Table 2.

Daily traffic data: For daily traffic data, UCB used the data collected by the California Department of Transportation (CalTrans) Performance Measurement System (PeMS) (https://dot.ca.gov/programs/traffic-operations/mpr/pems-source). PeMS data are collected in real-time from nearly 40,000 individual detectors spanning the freeway system across all major metropolitan areas of the State of California and provide an archived data user service that provides over fifteen years of data for historical analysis. PeMS integrates a wide variety of information from Caltrans and other local agency systems including traffic flow, speed, occupancy, incident, toll charge, and other information. UCB used PeMS five-minute road link/segment traffic flow data in the analysis. In PeMS, traffic flow (volume) is a quantity representing the number of vehicles that passed over each detector on the roadway in a given time period (i.e. five-minute flow, hourly flow, etc.). The detector measured traffic flow that covered 12.52 percent highway segments and UCB summed hourly traffic to daily traffic for all the stations across California. The following interconnected stages will be used to derive daily traffic for all the California

Table 2. The potential LUR predictors for the daily LUR model development.

Category	Variable for Prediction	Resolution	Description
Buffer (50m-5km)	Daily Traffic	vector	California Department of Transportation (CalTrans)
	Land Use	vector	Agricultural, residential, commercial, industrial, government and institutions, open land, parks, and recreational facilities (Parcel data)
	Land Cover	vector	Forest, herbaceous/grassland, shrubland, developed, agriculture, wetlands, water and other (USGS NLCD)
Non-Buffer Remote Sensing Data	Daily GridMET	4 km	Maximum temperature, minimum temperature, precipitation accumulation, downward surface shortwave radiation, wind-velocity, humidity (maximum and minimum relative humidity and specific humidity)
	Two-week Interval Vegetation Index	250 m	Normalized difference vegetation index (NDVI) (NASA MOD13Q1.006 Terra)
	Daily Aerosol Optical Depth (AOD)	1 km	NASA Multiangle Implementation of Atmospheric Correction (MAIAC) algorithm
	Daily Ozone from Ozone Monitoring Instrument (OMI)	27 km	Global for both NO_2 and O_3 measurements for 2004 – current (NASA)
	Annual $PM_{2.5}$	1 km	North America for 1989-2016 (Univ. Washington Randall Martin)
	Annual NO_2	1 km	Global for 1990-2020 (NASA reanalysis)
Other Non-buffer Variables	Digital Elevation Model (DEM)	30 m	U.S. Geological Survey (USGS)
	Distance to Coast	30 m	U.S. Geological Survey (USGS)
	Distance to Roadways	30 m	Environmental Systems Research Institute (ESRI)
	Distance to Ports	30 m	U.S. Geological Survey (USGS)
	Location category	vector	California Department of Transportation (CalTrans)/ESRI

highways for the study period:

1) For a road segment with station traffic measure for a day, use all the station traffic measures on that road segment to generate a daily mean traffic for that road segment for that day.
2) For those road segments without traffic measures for a day, assign them using the assigned segments from step 1 by matching route, county, district, route type and day, and find the one with the smallest distance if having multiple matches. California has 58 counties which are included in one of the 12 CalTrans air districts (1 - Eureka, 2 - Redding, 3 - Marysville / Sacramento, 4 - Bay Area / Oakland, 5 - San Luis Obispo / Santa Barbara, 6 - Fresno / Bakersfield, 7 - Los Angeles, 8 - San Bernardino / Riverside, 9 - Bishop, 10 - Stockton, 11 - San Diego, 12 - Orange County). Highways in California are split into at least four different types of systems: Interstate Highways, U.S. Highways, state highways, and county highways.
3) For those road segments without traffic being assigned from steps 1 & 2, assign them using the assigned segments from steps 1 & 2 by matching route, district, route type and day, and find the one with the smallest distance if having multiple matches. In this step county was not used as a restricting factor in daily traffic assignment.
4) For those road segments without traffic being assigned from the above steps, assign them using the above assigned segments by matching route, county, district and route type, plus at most one day difference in data availability and find the one with the smallest distance if having multiple matches.
5) Identify those not assigned and assign them using the assigned segments from above steps by matching county, district, route type and day and find the one with the smallest distance if having multiple matches. Here the restricting factor of route number is removed.
6) Identify those not assigned and assign them using the assigned segments from the above steps by matching district, route type and day and find the one with the smallest distance if having multiple matches. Here the restricting factors of route number and county are removed.

7a) Identify those not assigned and assign them using the assigned state highway segments from the above steps by matching district and day. Here the restricting factors of route number, route type and county are removed.

7b) Identify those not assigned and assign them using the assigned U.S. highway segments from the above steps by matching district and day. Here the restricting factors of route number, route type and county are removed.

7c) Identify those not assigned and assign them using the assigned interstate highway segments from the above steps by matching district and day. Here the restricting factors of route number, route type and county are removed.

8) Identify those not assigned and assign them using the assigned segments from steps 1-4 by matching district and season to find the one with the smallest distance if having multiple matches. Here route number, county and route type are not required to match.

Table 3 shows the daily traffic assignment statistics for the 12 California districts for years 2012-2019. Overall, 12.52 percent California highways had daily traffic measurements for the study period, with ranges being from 0 percent (district 9) to 38.24percent (district 12). UCB found that the districts with great population (i.e., metropolitan areas) had more roadways and more traffic measures. Those districts thus had smaller proportions of roadways being assigned traffic from greatly relaxed conditions (e.g., by gradually relaxing matching criteria on route, county, district, route type or day). The roadways in the vastly rural districts were the ones with much less proportion of traffic measures. Greater proportion of roadways were thus assigned through greatly relaxed conditions for those rural districts. The CalTrans PeMS traffic data started in 2001. A trend analysis from years 2001-2020 will be used to extend the daily traffic data back to the years between 1989-2001.

Table 3. The daily road traffic assignment statistics for 12 Caltrans districts in California for 2012-2019.

	District #1				District #2				District #3			
Stage	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)
1	34,197	2.93	34,197	2.93	64,284	4.94	64,284	4.94	75,002	4.41	75,002	4.41
2	774	0.07	34,971	3.00	0	0.00	64,284	4.94	142,554	8.38	217,556	12.79
3	686,788	58.91	721,759	61.91	943,806	72.58	1,008,090	77.53	68,950	4.05	286,506	16.85
4	431,122	36.98	1,152,881	98.89	292,200	22.47	1,300,290	100.00	1,548	0.09	288,054	16.94
5	0	0.00	1,152,881	98.89					704,938	41.45	992,992	58.39
6	0	0.00	1,152,881	98.89					503,072	29.58	1,496,064	87.97
7.1	12,997	1.11	1,165,878	100.00					204,540	12.03	1,700,604	100.00
	District #4				District #5				District #6			
Stage	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)
1	360,864	17.08	360,864	17.08	19,666	1.44	19,666	1.44	53,408	3.51	53,408	3.51
2	371,428	17.58	732,292	34.66	83,650	6.14	103,316	7.59	269,068	17.67	322,476	21.18
3	257,311	12.18	989,603	46.84	133,864	9.83	237,180	17.42	107,284	7.05	429,760	28.23
4	2,560	0.12	992,163	46.96	430	0.03	237,610	17.45	552	0.04	430,312	28.27
5	903,900	42.79	1,896,063	89.75	229,642	16.86	467,252	34.32	922,574	60.60	1,352,886	88.87
6	28,870	1.37	1,924,933	91.12	887,904	65.21	1,355,156	99.52	70,128	4.61	1,423,014	93.47
7.1	162,368	7.69	2,087,301	98.8	4,144	0.30	1,359,300	99.83	99,348	6.53	1,522,362	100.00
7.2	0	0.00	2,087,301	98.8	2,352	0.17	1,361,652	100.00				
7.3	0	0.00	2,087,301	98.8								
8	25,305	1.20	2,112,606	100								
	District #7				District #8				District #9			
Stage	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)
1	288,852	25.03	288,852	25.03	68,864	5.82	68,864	5.82	0	0.00	0	0.00
2	315,340	27.32	604,192	52.35	94,562	7.99	163,426	13.81	0	0.00	0	0.00
3	23,360	2.02	627,552	54.37	87,600	7.40	251,026	21.21	198,696	45.95	198,696	45.95
4	466	0.04	628,018	54.41	194	0.02	251,220	21.23	0	0.00	198,696	45.95
5	526,172	45.59	1,154,190	100	867,906	73.34	1,119,126	94.57	0	0.00	198,696	45.95
6					0	0.00	1,119,126	94.57	0	0.00	198,696	45.95
7.1					64,284	5.43	1,183,410	100.00	233,760	54.05	432,456	100.00
	District #10				District #11				District #12			
Stage	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)	RS (#)	RS (%)	Cum RS (#)	Cum RS (%)
1	146,644	9.80	146,644	9.80	241,134	23.85	241,134	23.85	160,898	38.24	160,898	38.24
2	438,638	29.32	585,282	39.12	400,820	39.65	641,954	63.50	139,650	33.19	300,548	71.43
3	352,216	23.54	937,498	62.66	105,120	10.40	747,074	73.89	0	0.00	300,548	71.43
4	2,288	0.15	939,786	62.82	990	0.10	748,064	73.99	290	0.07	300,838	71.50
5	544,392	36.39	1,484,178	99.21	262,948	26.01	1,011,012	100.00	119,930	28.50	420,768	100.00
6	11,886	0.79	1,496,064	100.00								

Note: RS= road segment; Cum RS=cumulative road segments; District 1, 2 and 9 had no traffic station measures and were treated the same as respectively neighboring districts in 4, 3 and 8.

$PM_{2.5}$ remote sensing data: UCB obtained Aerosol Optical Depth (AOD) data from the Moderate Resolution Imaging Spectroradiometer instruments onboard the National Aeronautics and Space Administration Terra and Aqua satellites. The Multiangle Implementation of Atmospheric Correction algorithm was used to derive 1 km resolution AOD surfaces.[6] Due to extensive missing data presented at the 1 km resolution AOD surfaces, UCB will aggregate the daily AOD surfaces into monthly means. The AOD data started in 2000. UCB will use the annual $PM_{2.5}$ data from the Univ. Washington in St Louis provided by Dr. Randall Martin' research group for 1989-2016 to extend the AOD data back to 1989. To do so, UCB built a linear relationship using the annual AOD data and the annual U Washington $PM_{2.5}$ data for 2000-2016 across the 120 PM2.5 monitoring stations with an adjusted R^2 of 0.92. The $PM_{2.5}$ concentration in year (1989-1999) month at location is calculated through:

Where is the predicted annual AOD value for year at location using U Washington data in a linear regression. and refer, respectively, to the mean AOD values at location across 2000-2016 and at location for month (= 1, 2, 3, …, 12).

Parcel-level land use data: UCB acquired statewide parcel data from CARB for 2019 for all the counties in California. The parcel data provides land use information at parcel level, such as agricultural, residential, commercial, industrial, government and institutions, open land, parks, and recreational facilities. For residential land use, the parcel data is further classed into single-family homes, town houses, condominiums, and high-rise apartment buildings. The parcel data also includes building characteristics, including building age, type and existence of fireplace, gas ranges, and other information that can be used to calculate building-specific factors to characterize the indoor infiltration of pollutants.

Land cover data: UCB has acquired the land cover data for year 2016 from the National Land Cover Database (NLCD). The NLCD provides a synoptic nationwide classification of land cover into 16 classes at a spatial resolution of 30 m. The 16 land cover classes were aggregated into eight major land cover types including forest, herbaceous/grassland, shrubland, developed, agriculture, wetlands, water and other, which includes ice/snow, barren areas. UCB also acquired tree canopy and percent impervious surfaces for 2016. These data will be extended to include the data for years 2001, 2004, 2006, 2008, 2011, 2013, 2016, and 2019.

Two-week interval vegetation index: UCB has acquired 16-day interval (23 surfaces for a year) vegetation index surfaces (MOD13Q1.006 Terra Vegetation Indices) for California at a spatial resolution of 250 m for years 2012 to 2019 for the study. This dataset traces back to 2000. UCB will use historical means to estimate vegetation indices for the years 1989-1999.

GridMET meteorological data: UCB has acquired daily high-spatial resolution (~4 km, 1/24th degree) surface meteorological data covering the contiguous U.S. for years 2012-2019. Primary climate variables collected include maximum temperature, minimum temperature, precipitation accumulation, downward surface shortwave radiation, wind-velocity, humidity (maximum and minimum relative humidity and specific humidity). UCB will further acquire the meteorological data for the years 1989-2011 to cover the entire study period.

Digital elevation model (DEM) – in meters: UCB acquired the national elevation dataset for California from the U.S. Geological Survey (USGS) (http://nationalmap.gov and http://seamless.usgs.gov) for 2011. The data included 45 1/3 arc-second (approx. 10 meters) raster DEM and were mosaicked into a single DEM raster for the entire State. Higher elevation is normally associated with lower $PM_{2.5}$ concentrations.

Distance to coast – in meters: The California shoreline was derived from The National Assessment of Shoreline Change: GIS Compilation of Vector Cliff Edges and Associated Cliff Erosion Data for the California Coast (http://pubs.usgs.gov/of/2007/1112). These data are integrated into the GIS mapping tool to produce a geographic view of topographical changes in California's coastline over time. The most recent view was created using data collected between 1998-2002. Greater distance is typically associated with greater $PM_{2.5}$ concentrations.

Distance to roadways – in meters: UCB used Business Analysts 2018 Street Carto map layer provided by the Environmental Systems Research Institute (ESRI in Redlands, CA) to derive distance to nearest highway (defined as feature class classification (FCC) A1 and A2), to nearest major roadway (FCC A3) and to nearest local roadway (FCC A4). Greater distance from roadways is typically associated with lower roadway traffic air pollution.

Location category – unitless: UCB classified the State of California into three exclusive location categories: Goods movement corridor (GMC) - areas within 500 m of truck-permitted freeways and ports, non-goods movement corridor (NGMC) - areas within 500 m of truck-prohibited freeways or 300 m of a connecting roadway, and control areas (CTRL) – locations out of GMC and NGMC. Typically, GMCs have the highest $PM_{2.5}$ concentrations while CTRLs have the lowest $PM_{2.5}$ concentrations. From 2012 to 2019, the number of $PM_{2.5}$ monitoring stations for GMC, NGMC and CTRL was, respectively, 51, 67 and 28. The total number of daily measurements for GMC, NGMC and CTRL for the years 2012-2019 was, respectively, 95113, 147513, and 74107. The $PM_{2.5}$ monitors were successfully deployed to significantly measure its near source impacts (those sites in GMC and NGMC) and also had a fairly number of sites located in the control areas to form a spatial representation of coverage. These statistics will be updated in this new research to include all the days with $PM_{2.5}$ regulatory monitoring.

$PM_{2.5}$ data from CalEPA monitoring: CalEPA started monitoring $PM_{2.5}$ concentrations in 1998. The number of air quality monitors increased substantially from 1998 to the current, with the largest number reaching 120 in 2021 (Figure *1*). The minimum values below detection limit, the mean values close to ten microgram per cubic meter (ug/m^3) and the maximum values over 500 ug/m^3. Though Google Air also measured $PM_{2.5}$ concentrations, they were measured by five binned particle counts, not mass. The Google Air $PM_{2.5}$ measurements will thus not be used in this study.

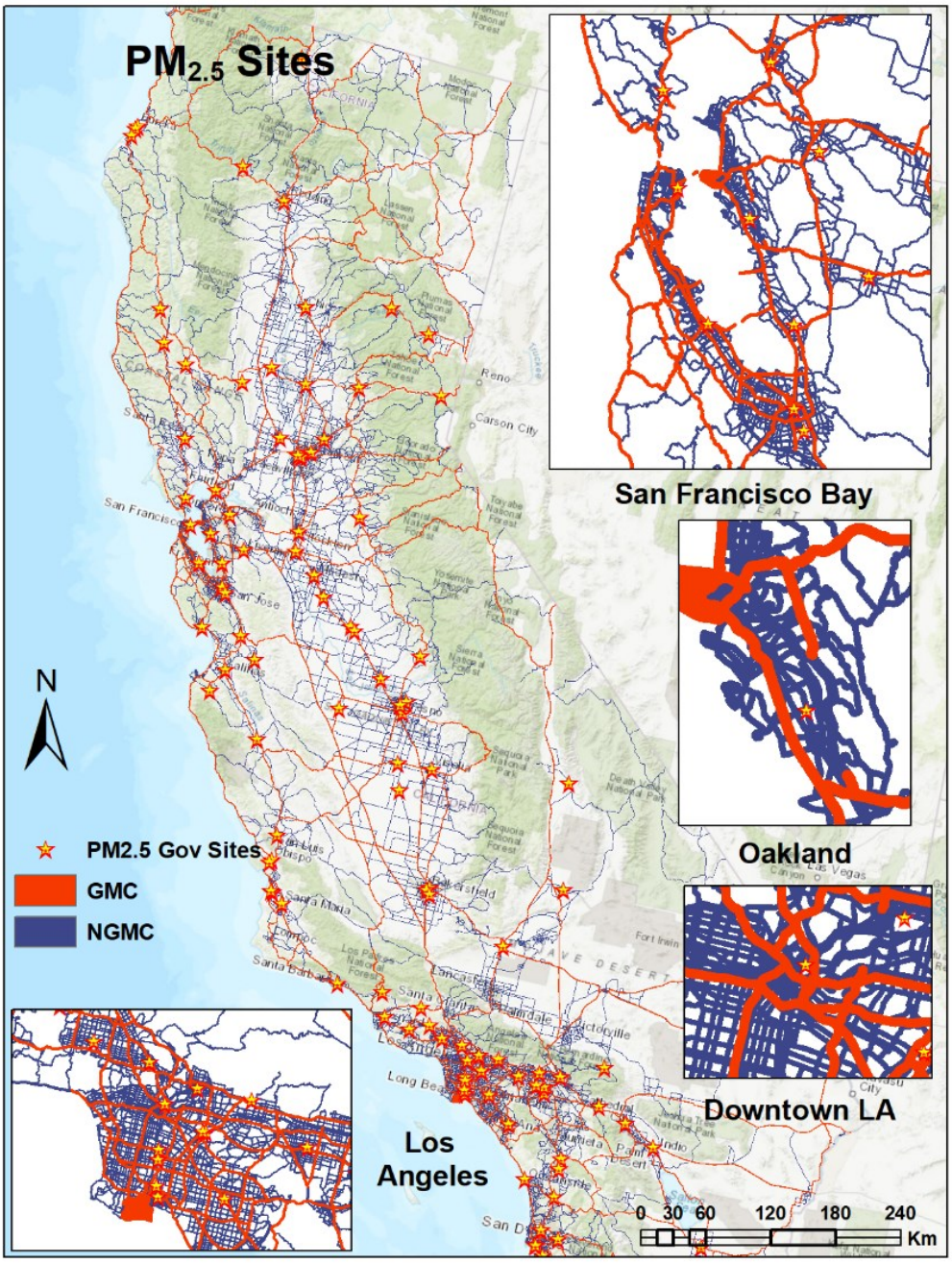


Figure 1. The spatial distribution of the CalEPA $PM_{2.5}$ air quality monitoring stations across California.

Development of Daily $PM_{2.5}$ concentration models

All the data sources of vector shape (e.g., traffic data and parcel level land use data) are converted into rasters with a spatial resolution of 30 m. The following describes a series of interconnected steps to develop a daily $PM_{2.5}$ model through the D/S/A modeling framework:

Generate buffer statistics on 30 m spatial resolution potential predictors: A series of buffer statistics of 50-5000 m at an interval of 50 m are created for the potential spatial predictors with a spatial resolution of 30 m. They include daily traffic data, parcel-level land use data, NLCD land cover data, and NLCD percent impervious and tree-canopy data. For each variable, e.g., industrial land use, a total of 100 buffered statistics (i.e., covariates) are generated. For all the potential predictors, with the inclusion of buffered and non-buffered variables, about 2,500 covariates are identified for the prediction of daily pollutant concentrations. This increases the chance of identifying the optimal distance impact of a predictor and helps improve model performance. However, this also creates high-dimension covariates that are highly correlated. To solve this issue, UCB applies a data reduction strategy to reduce the number of covariates used in predicting a pollutant concentration.

Apply data reduction strategy to reduce the number of predictors: To reduce the number of covariates and avoid high correlations between them for LUR modeling, UCB first creates a correlation coefficient matrix between a pollutant and all the covariates. A covariate of the highest absolute correlation coefficient with the pollutant is maintained. The maintained covariate is then used to calculate correlation coefficients with all the remaining covariates and those with an absolute correlation coefficient greater than 0.9 are removed from inclusion. A second covariate from the remaining covariates with the second highest absolute correlation coefficient with the pollutant is then maintained. Similarly, the second maintained covariate is used to calculate correlation coefficients with all the remaining covariates and those with an absolute correlation coefficient greater than 0.9 are removed from inclusion. This process continues until all the significant covariates are chosen and no two chosen covariates have an absolute correlation coefficient greater than 0.9. After applying the data reduction strategy, the number of predictors maintained in a LUR model is typically less than 100. This process is implemented before a D/S/A is run and it is carried out once for the sole purpose of reducing the number of predictors that might be collinear.

Develop daily LUR models and surfaces for $PM_{2.5}$: LUR modeling is a statistical technique used to estimate the spatial distribution of air pollution concentrations based on land-use characteristics and other variables. It analyzes measurements of air pollution levels taken at specific locations, and then identifies the key factors that influence those levels. To develop daily LUR models, UCB runs the model at 30 m spatial resolution through the D/S/A algorithm from the following line of code: dsa model <- DSA(PM2.5 ~ 1, id = vctrSiteID, data = dataRandom_dsa, usersplits = usrSplits, maxsize = 20, maxsumofpow = 1, maxorderint = 1, vfold = 10).[3,4] The D/S/A algorithm can deal with both linear and non-linear associations. However, for simplicity of model development and for the clear interpretation of the predictors selected for a model, UCB limited the predictors to be only on linear terms (the maximum sum of powers in each variable to be 1) and disallowed any interaction except corridor by year. The D/S/A algorithm is an aggressive model search algorithm, which iteratively generates polynomial generalized linear models based on the existing terms in the current 'best' model and the following three steps: (1) a deletion step, which removes a term from the model, (2) a substitution step, which replaces one term with another, and (3) an addition step, which adds a term to the model. The search for the 'best' estimator starts with the base model specified with 'formula': typically, the intercept model except when the user requires number of terms to be forced in the final model. Before searching through the statistical model space of polynomial functions, the original sample is randomly partitioned into V equal size subsamples. Of the V subsamples, a subsample is retained as the validation data for testing the model, and the remaining V-1 subsamples are used as training data. The cross-validation process is then repeated V times, with each of the V subsamples used exactly once as the validation data. The advantage of this method over the leave-one-out cross-validation technique is that the prediction errors are less impacted by single outliers, and compared to repeated random sub-sampling, all observations in the V-folds are used for both training and validation, and each observation is used for validation once. With each iteration, an independent validation dataset is used to assess the performance of a model built using a training dataset. This technique, therefore, minimizes over-fitting to the data to maximize the probability that the models will predict well at locations that have not been sampled.

During the D/S/A modeling process, UCB classifies the entire dataset into 10-folds. With each iteration, an independent validation dataset in one of the 10-folds is used to assess the performance of the model built using data from the other 9-folds. This process continues for 10 times until every fold of data is used for validation. The mean prediction errors from the validation datasets are averaged across the 10 iterations and compared between a series of built models. The model with the minimum average prediction error is chosen as the final model. During the modeling process, the air quality monitoring data (and associated predictors) for a specific year are equally and randomly distributed into those 10 folds. Because one air quality monitor typically has more than one observation (i.e., multiple days of measurements), a random effect of air quality monitor (in R language this is given by 1|station_ID) is included in the modeling process however, only the fixed effects (i.e., remove the random effect) are used to construct $PM_{2.5}$ surfaces due to the requirement of deriving $PM_{2.5}$ concentrations beyond monitoring stations. The adjusted R^2 for the fixed effects is used as a measure of model performance from the LUR modeling result.

Due to the requirement of more than three gigabytes of storage space for a single statewide raster surface of spatial resolution of 30 m, UCB opt to build daily surfaces of $PM_{2.5}$ concentrations using a spatial resolution of 100 m. Rather than creating a series of daily surfaces using a storage space of 27 terabytes (TBs) for the years 1989-2020 through a spatial resolution of 30 m, the 100 m spatial resolution surfaces will require only a storage space of about 4.5 TBs. The reduced size of the daily $PM_{2.5}$ surfaces will also make exposure assignments more feasible. The 100 m spatial resolution surfaces maintain the ability to identify small area variations of pollutant concentrations, especially those heightened exposures endured by vulnerable communities.

Assessment of model performance and uncertainty

The performance of the D/S/A modeling technique is assessed through an out-of-sample v-fold cross-validation technique, one of the best methods to assess model performance. The modeling process includes both fixed (predictors described in Table 1) and random effects (air quality monitor IDs); however, the adjusted R^2 for the fixed effects is used as a measure of model performance from the LUR modeling result.

Like all other models, the D/S/A LUR determines conditional mean of the response (here $PM_{2.5}$ concentrations) given the values of the explanatory variables. Extreme high and low $PM_{2.5}$ concentrations might not be able to be predicted correctly. This includes, for example, the limited ability to identify the impact of wildfires on $PM_{2.5}$ concentrations though AOD data are part of the predictors. Furthermore, the $PM_{2.5}$ concentrations are developed to represent daily mean concentrations, the D/S/A model proposed in this study cannot model within day variations.

Another uncertainty of the D/S/A modeling technique is that it includes data sources of coarse resolution as predictors (e.g., 1 km resolution AOD and 4 km resolution meteorological conditions). If those data sources of coarse resolution do not have dramatical changes in values between neighboring raster grids, the predicted concentrations will have a continued ‘realistic’ gradient. This is largely true for meteorological conditions. For those models with coarse resolution data sources as predictors and fine resolution data sources (e.g., those of 30m spatial resolution) also having significant impact, the modeling results represent a sub-pixel analysis. Small area variations in $PM_{2.5}$ concentrations can be detected. If for a location with only coarse resolution data impact and fine resolution predictors not presented (e.g., no traffic, no roadway, no variation in land use and land cover), the coarse resolution data source will have predominantly impact on the modeled concentrations. Dramatic edges might show if the coarse resolution data sources have dramatic changes in values between neighboring grids. This will largely happen in locations of remote areas.

Deliverable: UCB will provide daily air pollution surfaces of 100 m resolution for $PM_{2.5}$ for the years 1989- 2020 to CARB plus the modeling techniques used to derive those pollution surfaces.

Task 3. Obtain Medi-Cal Data including Mortality and Covariates

Obtain Medi-Cal data including mortality and individual-level covariates for the years 1990-2020

UCB will apply for Medi-Cal data including all the dates of death from Fee-For-Services and managed care for the years 1990-2020 from the DHCS (Table 4). In 1990, the state had 3.7 million Medi-Cal enrollees and the numbers have been increasing. In 2020, 13.5 million Californians have enrolled in the Medi-Cal program, roughly one-third of the state population. Assuming the same death rate as the general population, it is expected to have 36,000 deaths per year for the Medi-Cal enrollees for the years 1990-2005. That would result in 540,000 Medi-Cal deaths in the first 15-year period. Similarly, for the years 2006-2020, it is expected to have more than 72,000 deaths per year for the Medi-Cal enrollees. The total death from the Medi-Cal group would amount to more than one million for the second 15-year period. It is expected that Medi-Cal enrollees have a much higher death rate than the general population, so the actual number of deaths for the Medi-Cal enrollees should be higher than those estimated from Table 4. However, the relatively stable mortality rate indicates there was not a significant change in population shifts. For Medi-Cal enrollees, they are all low-income populations. When the data are dissected to 19 age groups and five major race-ethnicity groups (Hispanics, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, and Other) (19 *5 = 95), it is expected to have enough sample size on mortality for the most granular analysis and the outcome observations have more than 5 million per year. For the 1990-2005 and 2006-2020 periods, the average number of death is respectively 540,000 / 95 ≈ 5,684 and 1,000,000 / 95 = 10,526 per age group per race-ethnicity. UCB separates data into those two periods for the purpose of identifying the difference between the first and second generation in air pollution exposure and associated life expectancy.

The subject level variables to be collected include age, sex, race-ethnicity, smoking status, primary chronic conditions, comorbidity index, and access to primary care. Subjects' home address data will also be collected for the purpose of identifying their spatiotemporal PM2.5 exposure and vulnerability based on the CT of residence (see section below). Date of death of an enrollee will also be collected to assign one year of daily PM2.5 exposure before date of death (i.e., rolling average). The daily exposures across a generation before death and daily exposure across lifetime available exposure will also be calculated for sensitivity analysis in the life-expectancy analysis. The death of an enrollee will be matched through propensity score by two survival enrollees who had similar individual and neighborhood characteristics except air pollution. The same time span and length of exposure will also be estimated for the matched survival enrollees. UCB will apply for UCB Institutional Review Board (IRB) approval, as well as the IRB approval from the Committee for the Protection of Human Subjects (CPHS) of the California Health and Human Services (CHHS) Agency.

The Medi-Cal data will be used to estimate life expectancies impacted by air pollution from $PM_{2.5}$, separately, for (1) the first and second generations of all the Medi-Cal population, (2) race-ethnicity and vulnerability subgroups in each generation, and (3) the first and second generations within families. Task 4 details how those life-expectancies will be estimated.

Table 4. Annual mortality rates for general population and predicted annual death for Medi-Cal enrollees.

Year	State Pop	Death 1†	Death Rate‡	Medi-Cal Pop	Death 2ξ
1990	29,950,111	214,919	7.18	3,745,552	26,878
1995	31,493,525	224,604	7.13	5,421,262	38,663
2000	33,987,977	230,505	6.78	5,110,057	34,656
2005	35,827,943	237,526	6.63	6,555,369	43,460
2010	37,319,550	234,510	6.28	7,384,904	46,406
2015	38,904,296	260,227	6.69	12,961,488	86,698
2020	39,499,738	320,893	8.12	13,478,851	109,501

†: Death 1 is all cause annual population deaths from the general population.

§: Death 2 is all cause annual Medi-Cal population deaths assuming the same death rate as the general population.

‡: Death rate is the number of deaths per 1,000 people for *the general population.*

Obtain CT level vulnerability data through CalEnviroScreen

The vulnerable communities used in this research is defined by the Senate Bill (SB) 535 designation through CalEnviroScreen[1] at CT level for California. Based on the SB 535 designation, CARB defines the highest scoring 25 percent of CT from CalEnviroScreen as vulnerable communities. Additionally, 22 CTs that score in the highest five percent of CalEnviroScreen's Pollution Burden, but do not have an overall CalEnviroScreen score because of unreliable socioeconomic or health data, are also designated as disadvantaged communities. All the Medi-Cal enrollees will be assigned a CalEnviroScreen score and those living in vulnerable communities will be classified as vulnerable. Similarly, those living in 25 percent CTs of the least CalEnviroScreen score will be classified as non-vulnerable. Due to the potential changes in CT level social economic status and other environmental variables, the top and bottom 25 percent CTs might change for several versions of the CalEnviroScreen scores. UCB will choose the most occurring top and bottom 25 percent CTs as vulnerable and non-vulnerable communities. Loss of life expectancy due to air pollution will be estimated separately for those living in the vulnerable and non-vulnerable communities. Further, vulnerable communities will be overlaid with $PM_{2.5}$ concentrations and life expectancy estimates to identify vulnerable communities that had the highest $PM_{2.5}$ exposure and greatest impacts to life expectancy.

Obtain decennial CT level race-ethnicity and socio-economic status data

UCB will acquire CT level race-ethnicity, socio-economic status, and household characteristics data for decennial years (i.e., 1990, 2000, 2010, and 2020) from the National Historical GIS (NHGIS: see https://www.nhgis.org/). CT level race-ethnicity composition will be used as weights to estimate CT level life-expectancy. These data will also be used in the project for confounding control. A date of death of an enrollee will use the closest decennial data for confounding control. Overall, the confounding control will include those (1) at individual level such as age, sex, race-ethnicity, smoking status, primary chronic conditions, comorbidity index, and access to primary care, and (2) at CT level including race-ethnicity composition, socio-economic status, and household characteristics.

Deliverables: UCB will provide summary statistics of Medi-Cal participants information for the years 1990-2020 to CARB, including age, sex, race-ethnicity, smoking status, primary chronic conditions, comorbidity index, and access to primary care. PM2.5 exposure statistics based on home address will also be included. CT level statistics on vulnerability and race-ethnicity composition will also be provided. These statistics will be shown together and separately for the first and second generations.

Task 4. Calculate PM2.5-specific Reductions in Life Expectancy Across Two Generations and Identify Race-ethnicity and Vulnerability Disparities

It is well known that $PM_{2.5}$ levels in California have been going down for the last 30 years, contributed largely by CARB regulatory actions, however, these statewide reductions have not been equally distributed across communities. Disproportionate pollution exposure and subsequent health impacts are experienced predominately in under-resourced communities and communities of color. UCB and NCIRE hypothesize that overall reductions in statewide pollution exposure will result in differences in $PM_{2.5}$-specific impacts on life expectancy between generations, with the second generation having relatively lower air pollution exposure and smaller air pollution-specific life expectancy impacts. To differentiate effects from the two generations, air pollution-specific life expectancy loss will be modeled for the two generations separately: the life expectancy loss for the years 1990-2005 (first generation) and the life expectancy loss for the years 2006-2020 (second generation). The life expectancy loss will be estimated through differentiation of all cause life expectancy from cause $PM_{2.5}$-eliminated life expectancy. UCB and NCIRE also aim to identify

race-ethnicity subgroups and vulnerable communities that maintained higher air pollution exposure and greater air pollution-specific impacts on life expectancy across both generations. Therefore, air pollution-specific life expectancy loss will be separately estimated for race-ethnicity and vulnerability subgroups for the two generations. Table *5* lists the deliverables to be generated on air pollution exposure and associated life expectancy loss for this research. The following connected steps will be used to model the impact of $PM_{2.5}$ exposure on life expectancy and assess disparities for racial/ethnic and vulnerable subgroups.

Table 5. The deliverables to be generated on air pollution exposure and associated life expectancy loss.

Geographic Scope	Assessment Category		1990-2005 (First Generation)		2006-2019 (Second Generation)		Difference (second vs first)	
			Mean $PM_{2.5}$ Exposure	Life Expectancy Loss	Mean $PM_{2.5}$ Exposure	Life Expectancy Loss	Mean $PM_{2.5}$ Exposure	Life Expectancy Loss
Statewide	Overall Medi-Cal Population							
	Race-ethnicity Sub-groups	Hispanics						
		Non-Hispanic White						
		Non-Hispanic Black						
		Non-Hispanic Asian						
		Other						
	Vulnerability Sub-groups	Vulnerable						
		Marginally vulnerable						
		Marginally non-vulnerable						
		Non-vulnerable						
	Within Family Sub-groups							
CT Level	First Decennial within a Generation (1990 or 2010)							
	Second Decennial within a Generation (2000 or 2020)							

Propensity score matching

A propensity score will be calculated using a logistic regression model and the following covariates: age, sex, race-ethnicity, vulnerability, smoking status, primary chronic conditions, comorbidity index, and access to primary care. Using these propensity scores, each death (case) will be individually matched by those survived (control) using the nearest matching method within a caliper distance, which selects for matching a control subject whose propensity score is closest to that of the case subject (nearest neighbor matching approach) with the further restriction that the absolute difference in the propensity scores of matched subjects must be below some pre-specified threshold (the caliper distance).[7] Thus, participants for whom the propensity score cannot be matched because of a greater caliper distance will be excluded from further analysis. As suggested by Austin,[7] a caliper of width equal to 0.2 of the standard deviation of the logit of the propensity score will be used, as this value minimizes the mean squared error of the estimated treatment effect in several scenarios. To better match cases and controls, a 1:2 ratio matching method will be used. If a case subject cannot be matched to any control subject, then the case subject will be discarded. For each matched control participant, the same time span and length of exposure to the case will also be obtained using daily air pollution surfaces developed for the State. If more than two controls can be identified for a case, the UCB and NCIRE will select controls that have greater difference in air pollution exposure to the case. The propensity score matching process selects an optimal subset of data for health outcome analysis.

Develop model weights through inverse probability weighting (IPW)

Selection bias in pollution exposure and covariate imbalance make traditional air pollution health outcome analysis inaccurate.[8] Further, correlations between exposure of interest and confounding covariates create collinearity between them and bias the effect size of exposure. Causal modeling seeks to make the analysis of observational data mimic a randomized trial as closely as possible. In a randomized trial, the randomization assures that the exposure of interest is independent of the covariates. Application of IPW in the modeling process allows for estimation of the causal effect of increased $PM_{2.5}$ exposure while maintaining independence between $PM_{2.5}$ exposure and confounding (i.e., covariates). IPW is generated by

taking the inverse of the conditional probability of exposure to a given value in the continuum of $PM_{2.5}$ concentrations and stabilized by multiplying these weights by the marginal probability of the level of exposure.[8] For each year of exposure, It can be estimated through a two-step process:

$$E_{siy} = \beta_0 + \beta_1 P_{siy} + \beta_2 N_{siy} + \beta_3 O_{siy} + \gamma_s + \varepsilon_{siy} \quad (1)$$

E_{siy} is the air pollution exposure for subject *s* at age group *i* of year *y* using corresponding annual rolling average (365 days before) $PM_{2.5}$ concentration at each subject's residential address. To maintain statistical power in sample size, Medi-Cal enrollees will be grouped into 19 age groups, including <1, 1-5, 6-11, 12-17, 18-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60, 61-65, 66-70, 71-75, 76-80, 81-85, 86-90, and 91 and over. P_{siy} are the individual characteristics of subject *s* of age group *i* of year *y*. N_{siy} represents neighborhood confounding factors that might impact the relationship between air pollution and health outcome; O_{sty} are other confounding factors like temperature and relative humidity of year *y*. γ_s is the random effect of patient *s* and ε_{siy} is the error term of patient *s* of age group *i* of year *y*. Equation (Eq) (1) includes 19 age group specific models. The modeling results from Eq (1) is then used to create a weight at individual level for each model:

$$W_{siy} = E_{siy} / \widehat{E_{siy}} + (1 - E_{siy})/(1 - \widehat{E_{siy}}) \quad (2)$$

W_{siy} is the modeled weight for subject *s* at age group *i* for year *y* and $\widehat{E_{siy}}$ is the marginal effect estimated from Eq (1) on individuals' potential selective (conscious and unconscious) exposure to $PM_{2.5}$ (i.e., self-selection).

Develop age-group specific logistic regression models of probability of death.

UCB and NCIRE will fit a separate logistic regression mode for each age group using individual deaths as outcome ($d = 1$ for death and $d = 0$ for survival) and obtain an estimate of the probability of failing at age group *i* of date *y* (if no death, year *y)*, conditional on the covariates, and on the person having survived to age group *i*. These estimates make no parametric assumption about the distribution of the survival times. Further, they allow for the effect of both air pollution exposure and covariates to differ by year:

$$logit(d_{siy} = 1) = \beta_0 + \beta_1 E_{siy} + \beta_2 P_{siy} + \beta_3 N_{siy} + \beta_4 O_{siy} + W_{siy} + \gamma_s + \varepsilon_{siy} \quad (3)$$

$\widehat{d_{siy}}$ is the estimated probability that subject *s* dies at age group *i* in year *y*. Note that W_{siy} estimated from Eq (2) is added to Eq (3) to correct for model imbalance and dependence issues. All the other variables have the same definition as those in Eq (1). During the modeling process, UCB and NCIRE will also check for balance issues (due to the wrong specification of model distribution in Eq (1) and deal with them through, for example, doubly robust estimators.[9]

UCB and NCIRE will, for each of the two periods (1990-2005 and 2006-2020), generate 19 logistic models (one model for each age group) for the entire Medi-Cal population. Similar models will be developed separately for each major race-ethnicity (Hispanics, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian and other) and vulnerability subgroups (vulnerable, marginal vulnerable, marginal non-vulnerable and non-vulnerable, respectively, for vulnerability scores of 76-100th, 51-75th, 26-50th and ≤25th).

Develop nonparametric estimates of the distribution of life expectancy loss.

The loss of life expectancy due to air pollution from $PM_{2.5}$ exposure will be calculated by differentiating between two measures of life expectancy: all cause life expectancy at birth (e_0) and life expectancy at birth (e_0') eliminating the impact of $PM_{2.5}$ exposure.

To estimate all cause life expectancy at birth, the standard life table method will be used. First, age group-specific mortality rate (m_i) will be estimated as $m_i = d_i/l_i$, where d_i and l_i are, respectively, the number of death and total population at age group *i* (assume $l_0 = 100{,}000$). The conditional probability of dying from all cause at age group *i* can then be estimated as $q_i = 1 - exp^{-m_i n_i}$ (eq 5 in the Fergany paper[10]). n_i is the length of age interval for age group *i*. Person years of life lived by the cohort within age group *i* is then calculated by $L_i = d_i/m_i$ and the total person years of life contributed by the cohort after

attaining age group *i* can be estimated through $T_i = \sum_{j=i}^{19} L_i$. 19 is the total number of age groups designed for this study. The life expectancy for an individual at age group *i* is then calculated as $e_i = T_i/l_i$. The life expectancy at birth across all the age groups of the population is summarized as $e_0 = \sum_{i=1}^{19} L_i / l_0$. e_0 reflects the average life expectancy of a person from the cohort. The detailed steps in calculating life expectancy is fully documented by Apte et al.[11] and others.[12]

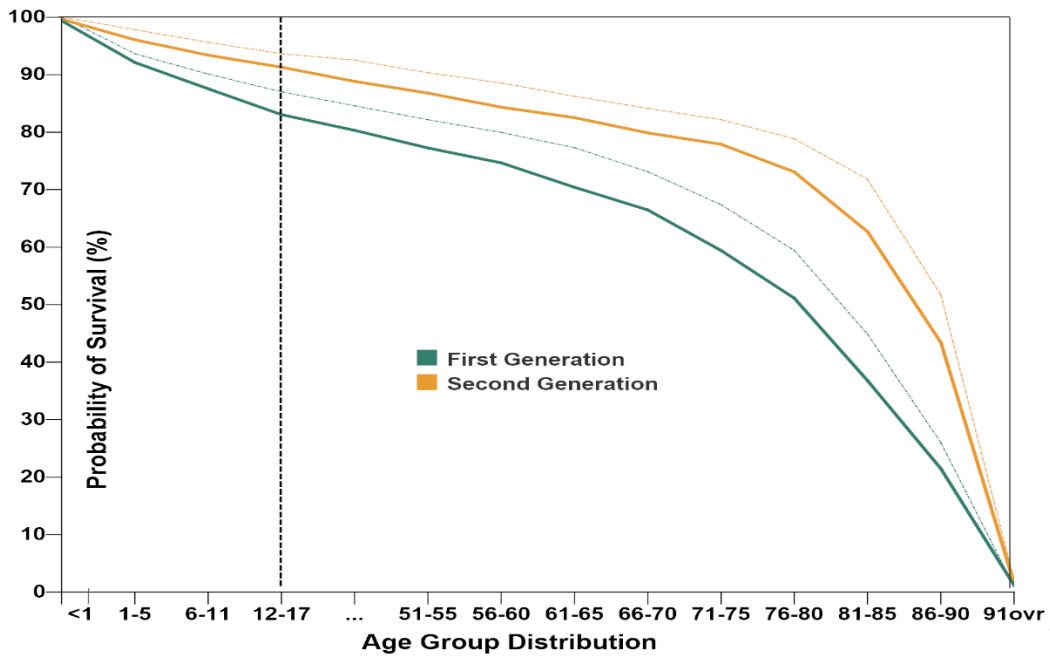
To estimate cause $PM_{2.5}$-eliminated life expectancy,[12,13] the counterfactual conditional probability of death will be estimated as $q_i' = 1 - p_i' = 1 - p_i^{(1-r_i)}$, where p_i is the conditional probability of survival from all cause for age group *i* and $p_i = 1 - q_i$. r_i is the fraction of death that is attributable to $PM_{2.5}$ exposure for age group *i* and $r_i = {q_i^{pm}}/{q_i}$. q_i^{pm} is the conditional probability of death given mean $PM_{2.5}$ exposure for people at age group *i*. It can be estimated through $\exp(\beta_{1i})/(1+\exp(\beta_{1i}))$,[14] where β_{1i} is the model coefficient from $PM_{2.5}$ exposure as in eq (3) from the logistic regression model for age group *i*. Given eq (5) in the Fergany paper,[10] the counterfactual mortality rate for age group *i* can be estimated as: $m_i' = -\frac{\log(1-q_i')}{n_i}$. Similar to estimating life expectancy for all cause, the following interconnected steps will be used to estimate cause $PM_{2.5}$-eliminated life expectancy for individual age groups (e_i') and for the overall life expectancy (e_0'):

$$l_0' = 100{,}000$$
$$d_i' = l_i' * q_i'$$
$$l_{i+1}' = l_i' - d_i'$$
$$L_i' = d_i' / m_i'$$
$$T_i' = \sum_{j=i}^{19} L_i'$$
$$e_i' = T_i' / l_i'$$
$$e_0' = \frac{\sum_{i=1}^{19} L_i'}{l_0'}$$

Life expectancy loss from $PM_{2.5}$ exposure is estimated as: $\Delta LE = e_0' - e_0$. ΔLE will be calculated for the first and second generation, as well as for race-ethnicity and vulnerability subgroups.

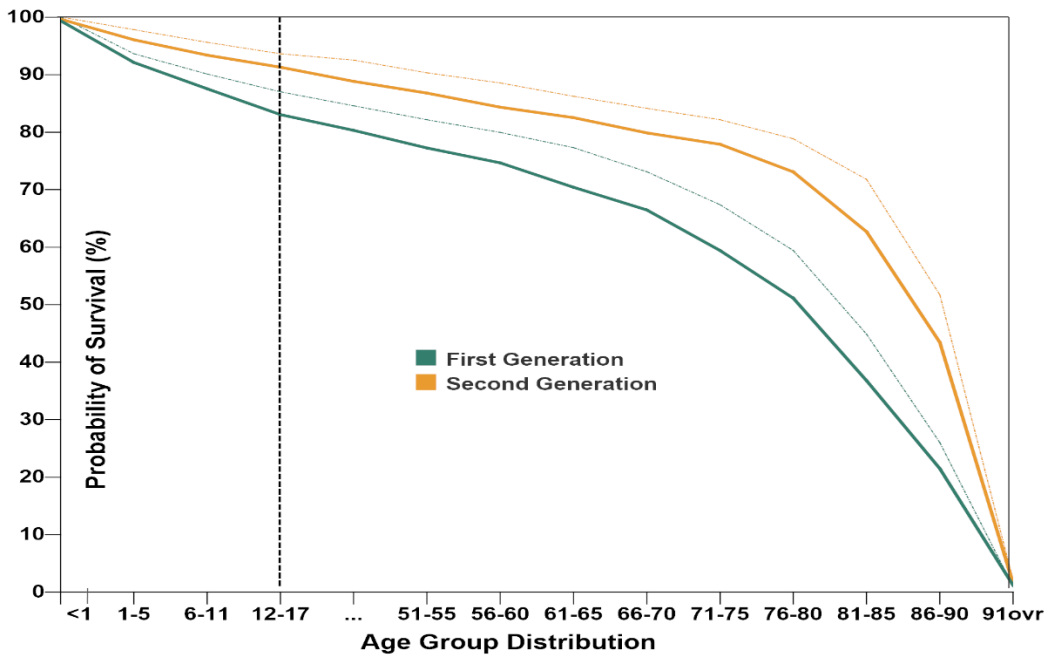
In a graphical representation, life expectancy corresponds to the area under the survival curve (AUC).[15,16] Research found that life expectancy estimated through the life table method (a deterministic value approach, see above) and the lifetime density function (a continuous approach) generated almost identical results with differences being less than 1.1 percent.[17] In this project, UCB and NCIRE will also apply AUC function to estimate life expectancy loss due to $PM_{2.5}$ exposure.

Based on the above estimates, the conditional probabilities of survival from all cause and cause $PM_{2.5}$-eliminated at age group *i* are, respectively, $1 - q_i$ and $1 - q_i'$. The probability survival curves for living up to age group *i* will be generated in a way similar to those presented in



Note: The solid lines represent the survival curves of the all cause cohorts and the dotted lines are for the survival curves of counterfactual PM2.5-eliminated cohorts.

Figure 2. On that figure, each age group is placed (a tick on x-axis) proportionally to the cumulative proportion of all age groups based on its age interval length *n*. The last age group (91 & over) has a standardized value of 1 (i.e., 100 percent), indicating all the deaths in the study. Y-axis is the probability of survival for a specific age group of interest, with the smallest value being zero (all dead, after reaching age group of 91 & over) and the biggest value being one (i.e., 100 percent with all subjects surviving). Assume all the study population are still surviving when reaching the age group 91 & over, the Area Under the Curve (AUC) $\int = 1$ (i.e., life expectancy being maximum: 92 will be used). If all the study population are dead before reaching age group <1, AUC $\int = 0$ (i.e., life expectancy being zero). The life expectancy for a known AUC $\int$ is calculated as 92 * $\int$. This is a nonparametric approach to estimate the distribution of life expectancy and it is particularly advantageous in cohort studies where there is left censoring at age of entry and a skewed distribution of life expectancy.[2] The AUCs will be estimated for both all cause (AUC_{all}) and cause $PM_{2.5}$ eliminated (AUC_{-pm}) using the R Description Tools AUC package through both naive trapezoid and spline interpolation approaches (https://search.r-project.org/CRAN/refmans/DescTools/html/AUC.html). The life expectancy loss due to $PM_{2.5}$ will then be calculated as: $\Delta LE_{auc} = 92 * (AUC_{-pm} - AUC_{all})$. ΔLE_{auc} and will be calculated for the first and second generation, as well as for race-ethnicity and vulnerability subgroups.



Note: The solid lines represent the survival curves of the all cause cohorts and the dotted lines are for the survival curves of counterfactual PM2.5-eliminated cohorts.

Figure 2. Assumptive probability curves of survival for the first and second generations.

Identify changes in the distribution of life expectancy.

UCB and NCIRE will estimate life expectancy loss, respectively, for the first and second generations. Differences in mean $PM_{2.5}$ exposure and life expectancy loss between the first and second generations will then be calculated. Life expectancy loss will also be estimated for children (0-17 years of age) and adults (18 and over years of age) separately when the second generation was compared to the first generation. UCB will also project $PM_{2.5}$ concentrations into future years (e.g., 2030 and 2050) and use the modeling results from eq (3) to derive future life expectancy loss due to air quality.

The changes in the distribution of life expectancy from the first generation to the second generation will be estimated not only for the entire Medi-Cal population but also separately for each major race-ethnicity and vulnerability subgroups using the methods above for each specific subgroup. UCB and NCIRE will identify the race-ethnicity subgroups and those living in the vulnerable communities maintained higher air pollution exposure and had greater air pollution related life expectancy loss when compared to other subgroups for two generations.

Sensitivity analysis

UCB and NCIRE will also generate life expectancies using families with more than one generation in the cohort. Here UCB and NCIRE will remove Medi-Cal enrollees without more than one generation participating in the program from analysis. For the two generation families, parents will be classified into the first generation and children into the second generation. If more than two generations exist, the grandparents will be classified into the first generation. The parents will be included in the second generation (in comparison to grandparents) and in the first generation (in comparison to children). The children will be included in the second generation. An interaction term between air pollution and family identification (e.g., insurer's name or home address) will be generated to identify if the second generation had reduced life expectancy loss. It is hypothesized that the second generation had relatively lower $PM_{2.5}$ exposure than the corresponding first generation and had smaller air pollution

related life expectancy loss. The within family analyses will be dependent on the number of family members identified from Medi-Cal enrollees. Information on families from the Medi-Cal enrollees such as race/ethnicity, total number, and generation of family members will not be known until the Medi-Cal data is acquired.

Another sensitivity analysis will use a subject's lifetime exposure before death for life expectancy modeling. A lifetime exposure is defined in two ways: (1) all exposure windows from the earliest days of available exposure (e.g., 1990) to death/year of assessment and (2) all exposure windows within a generation (1990-2005 or 2006-2020) before death/year of assessment. The matched participants will also be estimated for air pollution exposure during the same exposure window.

A third sensitivity analysis will use (1) exposure to NO_2 as the main air pollution impact in analyzing life expectancy and (2) simultaneous exposures to both NO_2 and $PM_{2.5}$ as the main air pollution impact factors. Use of NO_2 for air pollution related life expectancy loss estimates could present increased differences in air pollution and life expectancy loss between the race-ethnicity and vulnerability subgroups due to the increased spatial gradients in NO_2 exposure.

Deliverables: UCB and NCIRE will provide modeled Medi-Cal population mortality and life expectancy estimates due to air pollution from $PM_{2.5}$. The life expectancy estimates will include all cause life expectancy, cause $PM_{2.5}$-eliminated life expectancy and life expectancy loss from $PM_{2.5}$ exposure for (1) the first and second generations, (2) the race-ethnicity and vulnerability subgroups, and (3) the first and second generations within families. A summary of the methods and findings will be submitted to CARB.

Task 5. Create GIS Maps for the Study Results

UCB will aggregate the 100 m spatial resolution daily $PM_{2.5}$ concentrations developed through Task 2 for the years 1990-2020 to CT level means to generate annual $PM_{2.5}$ concentrations. Trends in annual changes in $PM_{2.5}$ exposure at CT level from 1990 to 2020 will be identified. $PM_{2.5}$ hotspots will be generated using top 25 percent CTs with the greatest $PM_{2.5}$ exposure. UCB will identify whether the vulnerable communities (identified through CalEnviroScreen) coincides with $PM_{2.5}$ hotspots through overlaying those two layers. UCB will also identify whether vulnerable communities experienced greater reductions in $PM_{2.5}$ exposure. The high $PM_{2.5}$ exposure and high vulnerability CTs will be mapped and overlaid with CTs of race-ethnicity composition to identify potential exposure disparities over race-ethnicity.

All life expectancy losses are estimated at State level. UCB will use CT level race-ethnicity composition as weights to adjust statewide life expectancy losses estimated for the race-ethnicity subgroups to create CT level life expectancy measures. They will be estimated through the following equation:

$$V_{jk} = \sum_{r=1}^{n} (V_{rk} * C_{rj}) \quad (5)$$

Where V_{jk} is the estimated life expectancy loss for CT *j* at generation *k* (*k*=1 for first and 2 for second). V_{rk} is the estimated life expectancy loss for race/ethnicity *r* of generation *k* and C_{rj} is the proportion of population of race/ethnicity *r* (*r* = Hispanics, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian and other) for CT *j*. Life expectancy loss V_{rk} values range from 0 to 95 and race-ethnicity composition C_{rj} values range from 0 to 1 (1 for 100 percent). Here *n* is equal to 5. Eq (5) also allows us to estimate life expectancy improved at CT level from the first to the second generation using $V_{j2} - V_{j1}$. For the first generation, UCB will estimate CT level life expectancies for decennial years 1990 and 2000. For the second generation, those life expectancies will be estimated for years 2010 and 2020. Those four decennial years have observed and most accurate CT level race-ethnicity composition data. Eq (5) will also be used to estimate CT level life expectancy improvements for two generations being family pairs (i.e., within families). The CT level life expectancy loss will also be

overlaid with CT level PM2.5 exposure and vulnerability to identify CTs with high PM2.5 exposure, vulnerability, and life expectancy loss and determine whether those disparities persisted through four decennial years (1990, 2000, 2010 and 2020).

Summary of the GIS maps/layers:

- CT level annual $PM_{2.5}$ concentrations for 1990, 2000, 2010, and 2020.
- $PM_{2.5}$ hotspots of top 25 percent CTs with greatest exposure for 1990, 2000, 2010, and 2020.
- CalEnviroScreen disadvantaged communities (CTs with top 25 percent of overall CalEnviroScreen scores).
- CT level life expectancy loss for 1990, 2000, 2010, and 2020.
- Change in CT level life expectancy loss from the first to the second generation and within families.
- Change in CT level $PM_{2.5}$ concentrations from the first to the second generation.

The final format and layers for the GIS maps will be determined after community feedback from Task 6.

Deliverables: UCB will provide statewide GIS maps at CT level to CARB in an online and publicly accessible format. They include (1) maps with $PM_{2.5}$ hotspots being overlaid with social vulnerabilities, (2) maps with life expectancies being overlaid with vulnerable communities, and (3) maps of trends of CT level exposure and life expectancy.

Task 6. Address Impacts in Communities

UCB will present the study design, research efforts and findings to local communities and receive feedback through an environmental equity approach that recognizes disparities in the distribution of environmental exposure and life expectancy within and across communities and seeks to identify individual and community needs. The RAMP will take the lead on the community outreach of the project. RAMP will host two webinars at the beginning of the project to inform stakeholders about the research and to solicit feedback about its scope and direction. Efforts will be made to get feedback from communities on what results communities are most interested in and what form the research findings used for outreach, such as the factsheets and maps of the findings should take. RAMP will also host two webinars toward the end of the project to share the research results and implications.

To reach relevant stakeholders, particularly those of vulnerable communities and communities of color, UCB will invite webinar participants through RAMP's various networks and e-newsletter (which reaches over 1,200 individuals, most of whom are in California). Likely webinar participants include representatives from a wide range of organizations and coalitions that provide services to people with asthma, particularly in low-income communities and communities of color. These representatives also shape policies and programs designed to address the environmental and social inequities (e.g., air pollution, substandard house, insufficient health care) that contribute to asthma disparities.

UCB will work with RAMP to prepare webinar materials, including summaries of results and impact of the project and various maps generated from Task 5, written in lay person's language. RAMP will also translate the materials into Spanish and will offer simultaneous Spanish/English translation during the webinars.

Deliverables: RAMP and UCB will provide four webinar recordings, summary reports of community feedback from webinars, English and Spanish presentation materials used for webinars, including lay-oriented summaries of project and results to CARB.

Task 7. Reporting and Data Delivery

UCB will meet with CARB staff quarterly and submit quarterly progress reports using the

CARB-designated template, and an invoice for the same period will accompany each progress report. Six months prior to the end of the study, UCB will submit a draft final report (DFR), which will include the results of the study and the additional deliverables identified in the Schedule of Deliverables. The DFR will be submitted in accordance with the Final Report format and will be reviewed by CARB staff. CARB's comments will be sent to UCB and after receiving the reviewer's comments, UCB shall modify and resubmit the modified DFR to the CARB contract manager. The modified DFR will be subject to formal review by the Research Screening Committee (RSC). Once accepted by the RSC, UCB will revise the modified draft final report addressing the RSC comments and any remaining concerns from CARB staff and will submit the revised final report to CARB. If CARB has additional comments on the report, UCB will be notified so appropriate changes can be made; otherwise, CARB will accept the revised final report as the final. UCB will submit the final report in an American with Disabilities compliant format. A notation in the Final Report task should denote that the University will incorporate a one-page Public Outreach Document into the Final Report, that will be widely used to communicate, in clear and direct terms, the key research findings from the study to the public. The format for the Public Outreach Document is outlined in Exhibit A1, Section 2.

In addition, UCB will present study findings at a CARB research seminar and provide electronic data and analyses to CARB, as identified in Exhibit A1 – Schedule of Deliverables.

Summary of Main Deliverables

PM2.5 Exposure			
1. 100m and Census Tract level Annual $PM_{2.5}$ Concentrations for years 1989-2020			
	2. Mean Annual $PM_{2.5}$ Exposure:		
	First Generation (1990-2005)	**Second Generation (2006-2020)**	**Difference (2nd gen - 1st gen)**
Statewide	Overall Medi-Cal population	Overall Medi-Cal population	Overall Medi-Cal population
	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other
	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)
	Within Families	Within Families	Within Families
PM2.5-specific life expectancy loss			
	First Generation (1990-2005)	**Second Generation (2006-2020)**	**Difference (2nd gen - 1st gen)**
Statewide	Overall Medi-Cal population	Overall Medi-Cal population	Overall Medi-Cal population
	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other	Race/ethnicity subgroups: • Hispanic • White • Black • Asian • Other
	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)	CalEnvioScreen Vulnerability subgroups: • Vulnerable (≥76th) • Marginally vulnerable (75-51th) • Marginally non-vulnerable (51-26th) • Non-vulnerable (≤25th)
	Within Families	Within Families	Within Families
Census Tract (CT) Level	Overall life expectancy loss for 1990 and 2000	Overall life expectancy loss for 2010 and 2020	Overall life expectancy loss (2010+2020) - (1990+2000)
	Life expectancy loss within families for 1990 and 2000	Life expectancy loss families for 2010 and 2020	Within family life expectancy loss (2010+2020) - (1990+2000)

Final GIS Maps/Layers
CT level annual $PM_{2.5}$ concentrations for 1990, 2000, 2010, and 2020.
$PM_{2.5}$ hotspots of top 25 percent CTs with greatest exposure for 1990, 2000, 2010, 2020.
CalEnviroScreen disadvantaged communities (CTs with top 25 percent of overall CalEnviroScreen scores)
CT level air pollution-specific life expectancy loss for 1990, 2000, 2010, 2020
Change in CT level air pollution-specific life expectancy loss from the first to the second generation and within families
Change in CT level $PM_{2.5}$ concentrations from the first to the second generation.

Project Schedule

The project will be completed in 24 months from the start date and the timeline is listed below.

	Months																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Task 1	■	■	■	■	■	■	■	■	■	■	■	■												
Task 2	■	■	■	■	■	■																		
Task 3	■	■	■	■	■	■	■	■	■	■	■	■												
Task 4							■	■	■	■	■	■	■	■	■	■	■	■						
Task 5													■	■	■	■	■	■						
Task 6	■	■	■													■	■	■						
Task 7	m		mp			mp			mp			mp			mp			md			mp			Fr
Note:	m = meeting with CARB staff; p = quarterly progress report; d = deliver draft final report; f = deliver final report; r = research seminar																							

Meetings

A. Initial meeting. Before work on the contract begins, the Principal Investigator and key personnel will meet with the CARB Contract Project Manager and other staff to discuss the overall plan, details of performing the tasks, the project schedule, items related to personnel or changes in personnel, and any issues that may need to be resolved before work can begin.

B. Progress review meetings. The Principal Investigator and appropriate members of his or her staff will meet with CARB's Contract Project Manager at quarterly intervals to discuss the progress of the project. This meeting may be conducted by phone.

C. Technical Seminar. The Contractor will present the results of the project to CARB staff and a possible webcast at a seminar at CARB facilities in Sacramento or El Monte.

CONFIDENTIAL HEALTH DATA AND PERSONAL INFORMATION

CARB will not be provided access to and will not receive any confidential health data or other confidential personal information under this contract. Further, CARB will have no ownership of confidential health data or other confidential personal information used in connection with this contract. The entities conducting the research in this contract will follow all applicable rules and regulations regarding access to and the use of confidential health data and personal information, including the Health Insurance Portability and Accountability Act (HIPAA) and requirements related to the Institutional Review Board (IRB) process. CARB will not be a listed entity with authorized access to confidential information pursuant to the IRB process for this contract.

HEALTH AND SAFETY

Contractors are required to, at their own expense, comply with all applicable health and safety laws and regulations. Upon notice, Contractors are also required to comply with the state agency's specific health

and safety requirements and policies. Contractors agree to include in any subcontract related to performance of this Agreement, a requirement that the subcontractor comply with all applicable health and safety laws and regulations, and upon notice, the state agency's specific health and safety requirements and policies.

Project Management Plan

The project will be led by UCB, in collaboration with NCIRE, University of Southern California (USC), and RAMP for a 24-month period and the project management plan is as follows (Figure 3 diagram below):

- Dr. Jason Su, Principal Investigator, will act as project director and the UCB-CARB contact and will be responsible for carrying out all seven tasks. Dr. Su will be working with the UCB Undergraduate Research Apprentice Program (URAP) research assistants to complete those seven tasks. Dr. Su and his URAP students will also draft quarterly progress reports, the draft final report, and the final report.

- Dr. Sei Lee, NCIRE, will act as subcontractor project manager for NCIRE. Dr. Lee will use his expertise in life expectancy modeling to help with the study design on acquisition of Medi-Cal data including determining cofounding factors and on identification of the impacts of air pollution on life-expectancy for the modeled proposed in this research. Further, Dr. Lee's research team will assist in literature review on impacts of air pollution on life expectancy.

- RAMP will lead the community outreach on the project that includes hosting two webinars at the beginning of the project to inform stakeholders about the research and to solicit feedback about the scope and two webinars toward the end of the project to share the research results. RAMP will also work with UCB to prepare webinar materials. (Task 6).

- Dr. Rob McConnell, USC, will provide advice on all aspects of the project, especially on life expectancy and health impacts analyses. Dr. McConnell is not budgeted on the project and will participate as a volunteer and a senior advisor.

The project team members, except Dr. McConnell, will have weekly conference calls to discuss issues associated with implementation of the project tasks, solutions for the tasks, and progress of the project. Dr. McConnell will attend the meetings at least once per month or at critical moments when advice is needed for moving the project forward. Dr. Su will track project progress as it relates to budgeted amounts in order to ensure that the project meets the targets.

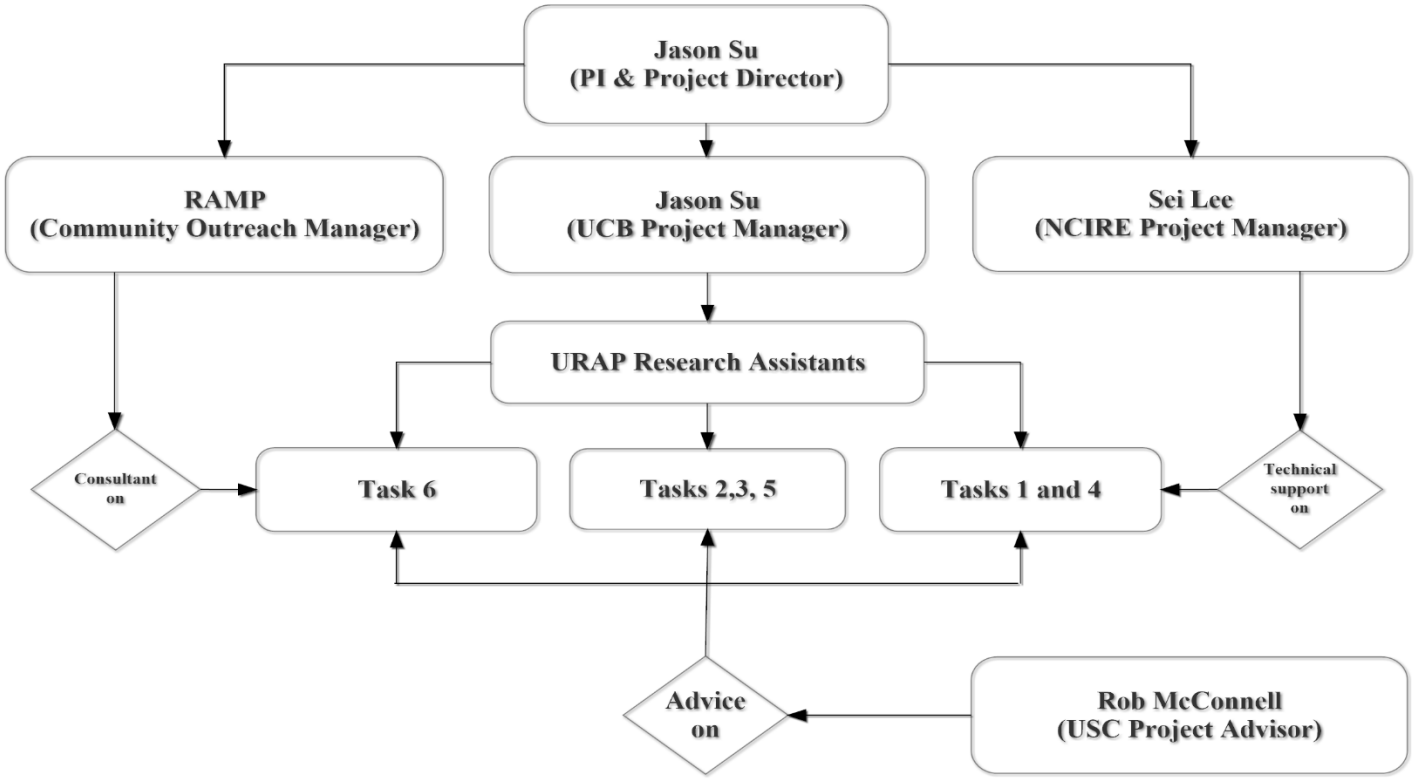


Figure 3. Project organizational chart.

References

1 August, L., Bangia, K., Faust, J., Slocombe, A. & Wieland, W. in *APHA's 2018 Annual Meeting & Expo (Nov. 10-Nov. 14).* (APHA).

2 Schwartz, J. D. *et al.* Estimating the effects of PM 2.5 on life expectancy using causal modeling methods. *Environ Health Persp* **126**, 127002 (2018).

3 Beckerman, B. S. *et al.* A Hybrid Approach to Estimating National Scale Spatiotemporal Variability of PM2.5 in the Contiguous United States. *Environ Sci Technol* **47**, 7233-7241 (2013). https://doi.org:10.1021/es400039u

4 Su, J. G., Jerrett, M., Meng, Y. Y., Pickett, M. & Ritz, B. Integrating smart-phone based momentary location tracking with fixed site air quality monitoring for personal exposure assessment. *Sci Total Environ* **506**, 518-526 (2015). https://doi.org:10.1016/j.scitotenv.2014.11.022

5 Su, J. G. *et al.* Predicting differential improvements in annual pollutant concentrations and exposures for regulatory policy assessment. *Environ Int* **143**, 105942 (2020).

6 Kloog, I. *et al.* A new hybrid spatio-temporal model for estimating daily multi-year PM2. 5 concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmospheric Environment* **95**, 581-590 (2014).

7 Austin, P. C. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res* **46**, 399-424 (2011).

8 Higbee, J. D. *et al.* Estimating long-term pollution exposure effects through inverse probability weighting methods with Cox proportional hazards models. *Environmental Epidemiology* **4** (2020).

9 Funk, M. J. *et al.* Doubly robust estimation of causal effects. *American journal of epidemiology* **173**, 761-767 (2011).

10 Fergany, N. On the human survivorship function and life table construction. *Demography* **8**, 331-334 (1971).

11 Apte, J. S., Brauer, M., Cohen, A. J., Ezzati, M. & Pope III, C. A. Ambient PM2. 5 reduces global and regional life expectancy. Environmental Science & Technology Letters 5, 546-551 (2018).

12 Arias, E., Heron, M. P. & Tejada-Vera, B. United States life tables eliminating certain causes of death, 1999-2001. (2013).

13 Tsai, S. P., Lee, E. S. & Hardy, R. J. The effect of a reduction in leading causes of death: potential gains in life expectancy. *Am J Public Health* **68**, 966-971 (1978).

14 Dayton, C. M. Logistic regression analysis. *Stat* **474**, 574 (1992).

15 Modig, K., Rau, R. & Ahlbom, A. Life expectancy: what does it measure? *BMJ open* **10**, e035932 (2020).

16 Day, S. M., Reynolds, R. J. & Kush, S. J. Extrapolating published survival curves to obtain evidence-based estimates of life expectancy in cerebral palsy. *Developmental Medicine & Child Neurology* **57**, 1105-1118 (2015).

17 Román, R., Comas, M., Hoffmeister, L. & Castells, X. Determining the lifetime density function using a continuous approach. *Journal of Epidemiology & Community Health* **61**, 923-925 (2007).

SCHEDULE OF DELIVERABLES

If use of any Deliverable is restricted or is anticipated to contain preexisting Intellectual Property with any restricted use, it will be clearly identified in Exhibit A4, Use of Preexisting Intellectual Property & Data.

Unless otherwise directed by the State, the University Principal Investigator shall submit all deliverables to State Contract Project Manager, identified in Exhibit A3, Authorized Representatives.

Deliverable	Description	Due Date
Initial Meeting	Principal Investigator and key personnel will meet with CARB Contract Project Manager and other staff to discuss the overall plan, details of performing the tasks, project schedule, items related to personnel or changes in personnel, and any issues that may need to be resolved before work can begin.	Month 1
Progress Reports & Meetings	Quarterly progress reports and meetings throughout the agreement term, to coincide with work completed in quarterly invoices.	Quarterly
Task 1 - Literature Review Findings	A copy of the systematic literature review findings on the impacts of air pollution on life expectancy to CARB. In addition, a copy of the literature review findings will be included as an appendix in the draft final report.	Month 12
Task 2 –$PM_{2.5}$ surfaces for the study population	Daily air pollution surfaces of 100 m resolution for PM2.5 for the years 1989- 2020 plus the modeling techniques used to derive those pollution surfaces.	Month 6
Task 3 – Obtain Medi-Cal Data including Mortality and Covariates	Summary statistics of mortality for the Medi-Cal population for the years 1990- 2020, including age, sex, race-ethnicity, vulnerability, smoking status, primary chronic conditions, comorbidity index, and access to primary care. PM2.5 exposure statistics based on home address will also be included. CT level statistics on social vulnerability and race-ethnicity composition will also be provided. These statistics will be shown together and separately for the first and second generations.	Month 12
Task 4 - Calculate PM2.5-specific Reductions in Life Expectancy Across Two Generations and Identify Race-ethnicity and Vulnerability Disparities	Modeled Medi-Cal population mortality and life expectancy estimates due to air pollution from PM2.5. The life expectancy estimates will include all cause life expectancy, cause $PM_{2.5}$-eliminated life expectancy and life expectancy loss from $PM_{2.5}$ exposure for (1) the first and second generations, (2) the race-ethnicity and vulnerability subgroups, and (3) the first and second generations for within families. A summary of the methods and findings will be submitted to CARB.	Month 18
Task 5 - GIS Maps for the Study Results	Statewide GIS maps at CT level in an online and publicly available format to CARB. They include (1) maps with PM2.5 hotspots being overlaid with social vulnerabilities, (2) maps with life expectancies being overlaid with vulnerable	Month 18

	communities, and (3) maps of trends of CT level exposure and life expectancy.	
Task 6 - Address Impacts in Communities	Four webinar recordings, summary reports of community feedback from webinars, English and Spanish presentation materials used for webinars, including lay-oriented summaries of project and results.	Months 3 and 18
Draft Final Report	Draft version of the Final Report detailing the purpose and scope of the work undertaken, the work performed, and the results obtained and conclusions.	Six (6) months prior to agreement end date.
Data	Data compilations first produced in the performance of this Agreement by the Principal investigator or the University's project personnel.	Two (2) weeks prior to agreement end date.
Technical Seminar	Presentation of the results of the project to CARB staff and a possible webcast at a seminar at CARB facilities in Sacramento or El Monte.	On or before agreement end date.
The following Deliverables are subject to paragraph 19. Copyrights, paragraph B of Exhibit C		
Final Report	Written record of the project and its results. The Final Report shall be submitted in an Americans with Disabilities Act compliant format. The Public Outreach Document, as described in Exhibit A1, Section 2, shall be incorporated into the Final Report.	Two (2) weeks prior to agreement end date.

KEY PERSONNEL

Last Name, First Name	Institutional Affiliation	Role on Project
Principal Investigator (PI):		
Su, Jason, PhD	The Regents of the University of California (UCB)	Principal Investigator
Other Key Personnel:		
Lee, Sei	Geriatrics, UCSF Department of Medicine; NCIRE	Subcontractor
McConnell, Rob	University of Southern California	Project Advisor (No Salary Requested)

RÉSUMÉ / BIOSKETCH

BIOGRAPHICAL SKETCH

NAME: Su, Guangquan (Jason)

eRA COMMONS USER NAME (credential, e.g., agency login): jasonsu

POSITION TITLE: Associate Researcher & Principal Investigator of Public Health

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

INSTITUTION AND LOCATION	DEGREE	COMPLETION	FIELD OF STUDY
Xiangtan Normal University, Huna, China	B.S	07/1989	Geography
Shaanxi Normal University, Shaanxi, China	M.S.	07/1996	Regional Geography (minor GIS)
Chinese Academy of Sciences, Beijing, China	Ph.D.	07/1999	Human and Economic Geography (minor GIS)
University of Alberta, Alberta, Canada	Ph.D.	06/2004	Rangeland and Wildlife Resources Manag. (minor Remote Sensing of Environment)
University of British Columbia, Vancouver, Canada	Postdoc	07/2007	Air Pollution Exposure Assessment and Environmental Health
University of California at Berkeley, Berkeley, USA	Postdoc	06/2010	Air Pollution Exposure Assessment and Environmental Health

A. Personal Statement

I have the expertise, leadership, training, expertise and motivation necessary to successfully carry out the proposed research project. With a background in Geographic Information Sciences (GIS) (first PhD), Remote Sensing (second PhD) and Microsoft-certified application development, I built a strong foundation on application of information technology for large scale data processing and statistical analysis. After two postdoctoral research experiences in environmental epidemiology, I advanced my career to machine learning modeling and bio-statistics analysis in environmental epidemiology. To facilitate identification of linkage between environmental exposure and health, I also developed strong skills in applying mathematical and machine learning algorithms for environmental exposure modeling and assessment. As a Principal Investigator, I have been very active in seeking funding to advance my research through the roles of contact PI and sub-award PI. Throughout the years, I have developed a collaborative network with other data science experts, environmental epidemiologists and bio- statisticians.

1. Williams AM, Phaneuf DJ, Barrett MA, Su JG. (2018). Short-term impact of $PM_{2.5}$ on contemporaneous asthma medication use: Behavior and the value of pollution reductions. Proceedings of the National Academy of Sciences 201805647.
2. Su JG, Barrett MA, Henderson K, Humblet O, Smith T, Sublett JW, et al. (2016). Feasibility of deploying inhaler sensors to identify the impacts of environmental triggers and built environment factors on asthma short-acting bronchodilator use. Environmental Health Perspectives 125(2):254- 61.

3. Su JG, Hopke PK, Tian Y, Baldwin N, Thurston SW, Evans K, et al. (2015). Modeling particulate matter concentrations measured through mobile monitoring in a deletion/substitution/addition approach. Atmospheric Environment 122:477-83.
4. Bork E and Su JG. (2007). Integrating LIDAR data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis. Remote Sensing of Environment 111(1):
   11– 24.

B. Contribution to Science

1. People with respiratory symptoms use rescue medication for immediate relief of acute symptoms. Latest advancements have medical sensors fit onto rescue inhalers to passively and objectively monitor the use of inhaled medications, capturing date and time and location of use of short-acting beta agonist. My research interests focus on applications of machine learning and bio-statistical modeling techniques to identify personal environmental triggers for people with respiratory symptoms, including those of asthma and COPD. Analyzing those data provides a driving force for my research and the analysis results are fed back into the individual-based healthcare system to promote self-management. I also collaborate extensively with other data science experts, environmental epidemiologists and bio-statisticians to conduct environmental health-related research.
   a. Casey JA, Su JG, Henneman LR, Zigler C, Neophytou AM, Catalano R, et al. (2020). Improved asthma outcomes observed in the vicinity of coal power plant retirement, retrofit and conversion to natural gas. Nature Energy:1-11.
   b. Pepper JR, Barrett MA, Su JG, Merchant R, Henderson K, Van Sickle D, et al. (2020). Geospatial-temporal analysis of the impact of ozone on asthma rescue inhaler use. Environment International 136:105331.
   c. Williams AM, Phaneuf DJ, Barrett MA, Su JG (2018). Short-term impact of $PM_{2.5}$ on contemporaneous asthma medication use: Behavior and the value of pollution reductions. Proceedings of the National Academy of Sciences:201805647.
   d. Su JG, Barrett MA, Henderson K, Humblet O, Smith T, Sublett JW, et al. (2017). Feasibility of deploying inhaler sensors to identify the impacts of environmental triggers and built environment factors on asthma short-acting bronchodilator use. Environmental Health Perspectives 125:254- 261.
2. In addition to the contributions described above, I also devote significant efforts to model small area variations in air pollutant concentrations and environmental exposures through land use regression (LUR) and other modeling techniques. They include applications of machine learning algorithms (e.g., Deletion / Substitution / Addition V-fold cross-validation and random forest modeling) and mathematical models (e.g., distance decay regression selection, fusion of multiple types of measurements into a single modeling framework, mobile data topic models and cumulative inequality models) for environmental exposure assessment and modeling. In related efforts, I also design fixed site and mobile saturation air quality monitoring. The fixed site and mobile saturation monitoring can complement existing spatially sparse networks from government continuous monitoring to uncover greater spatiotemporal variations.
   a. Su JG, Meng YY, Xiao C, Molitor J, Yue D, Jerrett M. (2020). Predicting differential improvements in annual pollutant concentrations and exposures for regulatory policy assessment. Environment International 143:105942.
   b. Su JG, Meng Y-Y, Pickett M, Seto E, Ritz B, Jerrett M. (2016). Identification of effects of regulatory actions on air quality in Goods Movement Corridors in California. Environmental Science & Technology 50(16):8687-96.

c. Su JG, Jerrett M, Beckerman B, Wilhelm M, Ghosh JK, Ritz B. (2009). Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. Environmental Research 109:657-70. PMCID: PMC3656661.

d. Su JG, Morello-Frosch R, Jesdale BM, Kyle AD, Shamasunder B, Jerrett M. (2009) An index for assessing demographic inequalities in cumulative environmental hazards with application to Los Angeles, California. Environmental Science & Technology 43:7626-34. PMID: 19921871.

3. To provide data for exposure assessment and environmental health study, I also apply my GIS, Remote Sensing and programming skills to acquire, process and generate related data. They include (1) acquire, process and model Remote Sensing data such as applications of high spatial resolution Landsat data to derive green spaces (e.g., parks, forest, grasslands and shrubs), calculate vegetation indices (e.g., normalized difference vegetation index) and derive urban land cover classification; (2) acquire, process and model spatial data such as conversion of attribute data into spatial data, various vector and raster spatial operations and application of numpy computing capabilities to handle large numeric operations; (3) develop dynamic systems for use by public health.
   a. Su JG, Dadvand P, Nieuwenhuijsen MJ, Bartoll X, Jerrett M. (2019). Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions. Environment international 126:162-170.
   b. Guo Y, Su JG, Dong Y, Wolch J. (2019). Application of land use regression techniques for urban greening: An analysis of tianjin, china. Urban Forestry & Urban Greening 38:11-21.
   c. Su JG. (2016). An online tool for obesity intervention and public health. BMC public health 16 (1): 136.

Complete List of Published Work in Google Scholar:
https://scholar.google.com/citations?hl=en&user=64EdGIQAAAAJ&view_op=list_works&sortby=pubdate

OMB No. 0925-0001 and 0925-0002 (Rev. 10/2021 Approved Through 09/30/2024)

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Lee, Sei Jae

eRA COMMONS USER NAME (credential, e.g., agency login): seilee

POSITION TITLE: Professor of Medicine, UCSF/NCIRE; Staff Physician, San Francisco VA Medical Center

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE	Completion Date	FIELD OF STUDY
University of Chicago	B.A.	1987-1991	Chemistry
University of Mississippi	-	1991	Education
University of California, Berkeley	-	1994-1995	Pre-Medicine
University of Illinois, College of Medicine	M.D.	1995-1999	Medicine
University of California, San Francisco	-	1999-2002	Internal Medicine
San Francisco VA Medical Center	-	2004-2006	Healthcare Quality
University of California, San Francisco	M.A.S.	2004-2006	Clinical Research
University of California, San Francisco	-	2006-2007	Geriatrics

A. Personal Statement

I am a Professor in the UCSF Division of Geriatrics and an physician-researcher at the San Francisco VA Medical Center. I am a geriatrician with advanced training in research methods (Master's degree in Clinical Research) and healthcare quality improvement (2-year VA Quality Scholars Fellowship). My long-term focus has been on how best to individualize treatment and prevention in older adults. The conceptual framework that underpins my work is that prevention (and treatment for asymptomatic conditions) exposes older patients to immediate burdens and risks (e.g., perforation during colonoscopy) for the promise of benefits in the future (e.g., decreased colorectal cancer mortality). For older patients with a limited life expectancy, prevention may expose them to risks with little chance that they would survive to benefit. To advance this work, I have led numerous studies focusing on predicting mortality risk and life expectancy in older adults, including several large national mortality prediction studies funded by National Institutes of Health and the Veterans Affairs Health Services Research and Development.

I am thrilled to work with Dr. Jason Su as a consultant on the proposed CARB supported projects focusing on life expectancy in race/ethnicity and vulnerability subgroups across generations of Medi-Cal beneficiaries. The proposed study will address important questions regarding the magnitude of benefit in terms of additional life among Medi-Cal beneficiaries due to the work of CARB in decreasing $PM_{2.5}$ levels. I will provide my perspective as a geriatrician to ensure that important associations between age, chronic conditions and mortality are modeled appropriately. I will provide my expertise as a mortality prediction expert to ensure that the proposed projects are completed at the highest levels of methodologic rigor.

Ongoing and recently completed projects that I would like to highlight include:

01/14/2026
NIH/NIA
Improving Care for Nursing Home Residents with Diabetes and Cognitive Impairment
The goal of the project is to engage mentees to examine current diabetes care practices in NH residents across the spectrum of cognitive impairment to determine which practices are associated with better (and worse) outcomes. These studies will help develop the aging research careers of junior investigators and will build the evidence base for diabetes care for nursing home residents with diabetes and cognitive impairment.

Role: Principal Investigator

R01 AG067427-01 (Dublin/Barnes) **07/15/2020 – 04/30/2024**
NIH/NIA

Identifying and supporting patients with undiagnosed dementia using the EHR Risk of Alzheimer's and Dementia Assessment Rule (eRADAR): a pilot clinical trial. This study will develop and test a new approach for identifying and reaching out to patients at risk of having undiagnosed dementia based on information in their electronic health records, aiming to maximize benefits while reducing harms.
Role: Co-Investigator.

R01AG057751 (Lee & Smith) **08/15/2018 – 04/30/2022**
NIH/NIA
Prognostic calculators for patients with Alzheimers disease and related dementias
The goal of this project is to create prognostic tools for estimating life expectancy, time to the nursing home placement, and 6-month mortality for persons with Alzheimer's disease and related dementias.
Role: Co-Principal Investigator

KL2TR001870 (Bauer) **07/18/2016 – 06/30/2022**

NIH/NCATS
Clinical and Translational Science Institute
The UCSF Clinical and Translational Science Institute (CTSI) at the University of California, San Francisco (UCSF) was established in 2006 to accelerate the pace of research that improves the health of the public. The overall mission of CTSI is to improve and transform clinical and translational research infrastructure and training at UCSF and partner institutions.
Role: Co-Investigator, K-Grant Writing Workshop Co-Director

RF1AG062568 (Odden) **05/15/2019 – 03/31/2024**
NIH/NIA
Management of Hypertension Among Persons with and without Dementia in Long-Term Care
Recent guidelines have identified a lack of information on how to best manage hypertension among older adults in long-term care and those with dementia. This research will evaluate the benefits and harms of treatment for high blood pressure in this population.
Role: Co-Investigator

Selected Completed Research Support
VA HSR&D IIR 15-434 (Lee) **02/01/2017 – 09/30/2021**
Department of Veterans Affairs/Veterans Health Administration

Development and Validation of 10-Year Life Expectancy Calculators to Individualize Veterans' Prevention Decisions

The goal of this project is to develop and validate 10-year life expectancy calculators using VA electronic medical record data so that these automated calculators can be linked with existing colorectal (CRC) screening clinical reminders so that 1) reminders are suppressed for Veterans with limited LE (even if they fall within the recommended age for screening) and 2) reminders are triggered for Veterans with an extended LE (even if they are beyond the recommended age range for screening).
Role: Principal Investigator

R01AG052041 (Steinman) 08/01/2016 – 04/30/2020
NIH/NIA
Secondary Analyses of Existing Data Sets to Address Clinical Aging Research Questions

The primary aims of this study are (1) To develop and validate claims-based measures of multimorbidity that predict functiona outcomesl; (2) Using an expanded range claims data, to develop and validate measures of multimorbidity that predict hospitalization and death; (3) To compare the predictive validity of our measures for functional decline, hospitalization, and death with existing measures of multimorbidity.
Role: Co-Investigator

R01AG047897 (Smith & Lee) 05/15/2015 – 01/31/2020
NIH/NIA
Developing prognostic models for life expectancy and geriatric outcomes
The objective of this project is to create prognostic tools for estimating life expectancy and time to the onset of disability, difficulty managing finances or medications, and mobility impairment.
Role: Co-Principal Investigator

Relevant Publications
a. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4- year mortality in older adults. JAMA. 2006 Feb 15;295(7):801-8.
b. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic Indices for Older Adults: A Systematic Review. JAMA. 2012 Jan 11;307(2):182-92.
c. Cruz ML, Covinsky KE, Widera EW, Stijacic Cenzer I, Lee SJ. Predicting 10-Year Mortality in Older Adults. JAMA. 2013 Mar 6;309(9):874-6.
d. Lee SJ, Boscardin WJ, Kirby K, Covinsky KE. Individualizing Life Expectancy Estimates for Older Adults Using the Gompertz Law of Human Mortality. PLOS One. 2014 Sep 29;9(9):e108540.

B. Positions, Scientific Appointments,

and Honors Positions and Employment
2019- Professor of Medicine, UCSF and Staff Physician, SFVAMC
2013-2019 Associate Professor of Medicine, UCSF and Staff Physician, SFVAMC
2012- Senior Scholar, VA Quality Scholars, VA Medical Center, San Francisco, CA
2007-2013 Assistant Professor of Medicine, UCSF and Staff Physician, SFVAMC
2007-2012 Associate Director, VA Quality Scholars, VA Medical Center, San Francisco, CA
2006-2007 Clinical Geriatrics Fellow, UCSF and SFVAMC
2004-2006 Research Fellow, VA Quality Scholars Fellowship, SFVAMC and UCSF
2002-2004 Hospitalist Physician, Kaiser Foundation Hospital, South San Francisco, CA
1999-2002 Primary Care Internal Medicine Resident, UCSF

Other Experience and Professional Memberships
2022-present VA/DoD Diabetes Mellitus Clinical Practice Guideline Workgroup 2021-present VA HSR&D MRA0 study section

2019-present Founding board member, International Geriatric Diabetes Society 2015-2020 VA HSR&D HSR2 study section
2012-2015 AGS Choosing Wisely Workgroup member and panelist
2012 Invited Presenter, National Advisory Council on Aging (National Institute on Aging)
2011-present AGS Quality and Performance Measurement Committee member (Vice-Chair 2016-19, Chair 2019-present)
2011-present American Colleges of Physicians 2010-present American Medical Association
2008-2012 AHRQ Development of Measures for Home and Community Based Population, Consultant 2005-present American Geriatrics Society
2002-2007 Society of Hospital Medicine
2002-present Society of General Internal Medicine
1991-1992 High School Chemistry and Physics teacher, Clarksdale, MS

Honors

2017 AGS Yoshikawa Award for Clinical Investigation 2015 UCSF Mentor Development Program
2014 AGS Outstanding Committee Service Award
2013 POGOe Product of the Year Awardee, Best POGOe Educational Material of the Year (for ePrognosis: Cancer Screening)
2012 AGS Clinical Decision Making, Best Paper Award 2012 Merck/AGS New Investigator Award
2010 Annals of Internal Medicine Top Grade Reviewer 2009 JGIM Distinguished Reviewer
2008 Hartford Geriatrics Health Outcomes Research Scholar 1999 Medical School Commencement Speaker
1999 Thomas Reeves Award (chosen by medical school peers as the person they would want caring for themselves or their family)
1999 King's College (London) Scholarship 1997 Alpha Omega Alpha

C. Contributions to Science

1) Improving the accuracy of mortality prediction in older adults. My research team has focused on how best to improve the accuracy of mortality prediction indices. In my first paper, we showed that incorporating functional measures improved the accuracy of 4-year mortality prediction. In follow-up papers we have shown that the importance of function for predicting mortality increases with increasing age, and validated our index for 10 year mortality. We have also developed models to predict mortality using EHR data.

a. Lee SJ, Moody-Ayers SY, Landefeld CS, Walter LC, Lindquist K, Segal MR, Covinsky KE. The Relationship Between Self-Rated Health and Mortality in Black and While Americans. J Am Geriatr Soc. 2007 Aug 14; 55(10): 1624-9
b. Lee SJ, Go AS, Lindquist K, Bertenthal DB, Covinsky KE. Chronic Conditions and Mortality Among the Oldest Old. Am J Public Health. 2008 Jul; 98(7): 1209-14.
c. Cruz M, Covinsky KE, Widera EW, Cenzer IS, Lee SJ. Predicting 10-Year Mortality for Older Adults. JAMA. 2013 Mar 6;309(9):874-6. doi: 10.1001/jama.2013.1184.
d. Lee AK, Jing B, Jeon SY, Boscardin WJ, Lee SJ. Predicting Life Expectancy to Target Cancer Screening Using Electronic Health Record Clinical Data. J Gen Intern Med. 2021 Jul 29. doi: 10.1007/s11606-021-07018-7.

2) Improving the usability of prediction indices. We conducted a systematic review to identify mortality indices for older adults and then transformed these indices into online calculators available at www.eprognosis.ucsf.edu to help busy clinicians use these indices. We then studied how these online calculators were being used and developed an iPhone/iPad app which links mortality prediction information with time to benefit to guide clinicians in determining whether breast or colorectal cancer screening is indicated in a given patient. Since life expectancy (e.g., 6.5 years) is easier to interpret than mortality risk prediction (e.g., 40% risk at 5 years) for both clinicians and patients, we developed a methodology that leverages the Gompertz Law of Human Mortality to transform develop life expectancy calculators. We've also developed methods to make prediction models easier to use, including accounting for the time required to ascertain predictor variable (Time Cost Information Criterion), developing optimal sets for predictors when predicting several outcomes (e.g. death and disability) and how to allow models to predict when the user is unable to provide several predictors.

a. Website: www.eprognosis.ucsf.edu
b. McClymont K, Lee SJ, Schonberg M, Widera E, Miao Y, Smith AK. Usefulness and Impact of Online Prognostic Calculators. J Am Geriatr Soc. 2014 Dec;62(12):2444-5.
c. Lee SJ, Smith AK, Diaz-Ramirez LG, Covinsky KE, Gan S, Chen CL, Boscardin WJ. A Novel Metric for Devloping Easy-to-Use and Accurate Clinical Prediction Models: The Time-Cost Information Criterion. Med Care. 2021 May 1;59(5):418-24. doi: 10.1097/MLR.0000000000001510.
d. Aliberti MJR, Kotwal AA, Smith AK, Lee SJ, Banda S, Boscardin WJ. Pre-Estimating Subsets: A new Approach for Unavailable Predictors in Prognostic Modeling. J Am Geriatr Soc. 2021 Sep;69(9):2675-
78. Alzheimer's Research and Therapy. 2014 Feb;10(6):646-55. doi: 10.1111/jgs.17278

3) Determining the time to benefit for preventive interventions. Most randomized trials are designed to determine whether an intervention works (e.g., p-value) and then to determine how well the intervention works (e.g., effect size). However, for older adults, the question of "when will it help?" is critical in determining who should receive the intervention. Since the question of "when will it help?" was generally not addressed by high-quality trials, my team and I developed a survival meta-analysis methodology that combined data from trials to quantify when benefits were seen. We measured the time to benefit for flexible sigmoidoscopy screening to be 9.4 years, but substantially shorter times to benefit for statins for primary prevention (2.5 years) and bisphosphonates (12.4 months).

a. Lee SJ, Leipzig RM, Walter LC. "When Will it Help?" Incorporating Lagtime to Benefit into Prevention Decisions for Older Adults. JAMA. 2013 Dec 25;310(24):2609-10.
b. Tang V, Boscardin WJ, Stijacic-Cenzer I, Lee SJ. Time to Benefit for Colorectal Cancer Screening: Survival Meta-Analysis of Flexible Sigmoidoscopy Trials. BMJ. 2015 Apr 16;350:h1662.
c. Yourman LC, Cenzer IC, Boscardin WJ, Nguyen BT, Smith AK, Schonberg MA, Schoenborn NL, Widera EW, Orkaby A, Rodriguez A, Lee SJ. Evaluation of Time to Benefit for the Primary Prevention of Cardiovascular Events in Adults Aged 50 to 75 Years: A Meta-Analysis. JAMA Intern Med. 2021 Feb 1;181(2):179-185. doi: 10.1001/jamainternmed.2020.6084.
d. Deardorff WJ, Cenzer I, Nguyen B, Lee SJ. Time to Benefit of Bisphosphonate Therapy for the Prevention of Fractures Among Postmenopausal Women with Osteoporosis: A Meta-Analysis of Randomized Clinical Trials. JAMA Intern Med. 2022 Jan 1;182(1):33-41. doi: 10.1001/jamainternmed.2021.6745.

4) Glycemic control in Older Adults with Diabetes. Although older adults are a large and rapidly growing proportion of the diabetes population, very little trial evidence is available to inform

decision making. Further, diabetes trials have often ignored outcomes most important to older adults, such as incontinence and functional decline. Thus, my team and I have published a series of studies examining outcomes in older adults associated with how varying glycemic treatment strategies.

a. Lee SJ, Eng C. Goals of Glycemic Control in Frail Older Patients with Diabetes. JAMA. 2011 Apr 6;305(13):1350-1.
b. Lipska KJ, Krumholz H, Soones T, Lee SJ. Polypharmacy in the Aging Patient: A Review of Glycemic Control in Older Adults with Type 2 Diabetes. JAMA. 2016 Mar 8;315(10):1034-45. doi: 10.1001/jama.2016.0299
c. Petrillo LA, Gan S, Jing B, Lang-Brown S, Boscardin WJ, Lee SJ. Hypoglycemia in Hospice Patients with Type 2 Diabetes in a National Sample of Nursing Homes. JAMA Intern Med. 2018 May 1;178(5):713-15. doi: 10.1001/jamainternmed.2017.7744
d. Jeon SY, Shi Y, Lee AK, Hunt L, Lipska K, Boscardin J, Lee S. Fingerstick Glucose Monitoring in Veterans Affairs Nursing Home Residents with Diabetes Mellitus. J Am Geriatr Soc. 2021 Feb;69(2):424-31. doi: 10.1111/jgs.16880

5) Improving Quality Indicators and Clinical Guidelines for Older Adults. After my training in the VA Quality Scholars fellowship and in Geriatrics, one of my areas of interest has been how quality indicators and clinical guidelines affect care for older adults. I co-mentored a fellow focusing on whether a widely used hospital quality score (Leapfrog Safe Practices Score) was associated with mortality. Additionally, we studied how the implementation of the guideline to keep hemoglobin A1c <8% in frail older adults led to increased rates of severe hypoglycemia. I have written about how to develop quality indicators for older adults that could avoid unintended harms as well as recommendations for how to craft performance measures for older adults with limited life expectancy.

a. Kernisan LP, Lee SJ, Boscardin WJ, Landefeld CS, Dudley RA. Association between hospital- reported Leapfrog Safe Practices Scores and Inpatient Mortality. JAMA. 2009 Apr 1;301(13):1341-8. PMC2851624
b. Lee SJ, Boscardin WJ, Cenzer IS, Huang ES, Rice-Trumble K, Eng C. The Risks and Benefits of Implementing Glycemic Control Guidelines in Frail Elders with Diabetes. J Am Geriatr Soc. 2011 Apr;59(4):666-72.
c. Lee SJ, Walter LC. Quality Indicators for Older Adults: Preventing Unintended Harms. JAMA. 2011 Oct 5;306(13):1481-2.

Complete List of Published Work in MyBibliography:
http://www.ncbi.nlm.nih.gov/sites/myncbi/sei.lee.1/bibliography/42323297/public/?sort=date&direction=ascen ding.

NAME: McConnell, Rob S.

eRA COMMONS USER NAME (credential, e.g., agency login):

rmcconne POSITION TITLE: Professor

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

INSTITUTION AND LOCATION	DEGREE (if applicable)	Completion Date MM/YYYY	FIELD OF STUDY
Stanford University, Palo Alto, CA	OTH	06/1976	Urban Studies/Latin American Studies
University of California, San Francisco, CA	MD	06/1980	Medicine

A. Personal Statement

I am a Professor of Preventive Medicine and the current Director of an NIEHS-supported P30 Infrastructure Center and of a new Children's Environmental Health Science Research Translation (CEHRT) Center. I lead the Climate and Health Initiative in the P30. The CEHRT Center theme is "Imagining L.A. without fossil fuels". I have developed innovative approaches to estimating the burden, cost and public health impact of air pollution exposure including in the context of climate change, focused on policy-relevant outcomes at a local scale.

I was also the director of the NIEHS-supported Southern California Children's Environmental Health Center during most of the last 7 years, and I have been the PI or project director for several large R01s or Centers leveraging the Southern California Children's Health Study, a large, ongoing longitudinal cohort study that has made important contributions to understanding the role of genetics and air pollution in childhood origins of respiratory and cardiometabolic health and obesity, and more recently to understanding the evolving epidemic of e-cigarette use. My research interests include, in addition, environmental contributions to autism in children, novel methods for assessment of environmental exposure, and understanding childhood susceptibility to the effects of air pollution related to psychosocial stress and social factors, exercise, and environmental disparities. I have estimated the burden of disease associated with near-roadway air pollution as a local tool for risk communication with communities and policy makers. I also served on the immediate past EPA Clean Air Science Advisory Committee on $PM_{2.5}$ before it was disbanded under the last administration. I have extensive experience in organizing and promoting research collaborations, including with other NIEHS Children's and P30 Centers. I am on the governing Council of the International Society for Children's Health and Environment, a policy-focused organization. I also have experience developing large interdisciplinary translational projects prior to coming to USC. As the director of a World Health Organization (WHO) regional center for environmental health for Latin America and the Caribbean, I was a member of advisory committees to the Ministries of Health in the Americas and of the senior management team advising the WHO Regional Director for the Americas. I have presented extensively in policy forums, including to Congress and to state, county and local government. I also developed and co-direct the Career Development Program for the USC NIEHS Centers, a role in which I have acquired extensive experience as formal and informal mentor to junior faculty and postdoctoral students. I am the MPI for the USC Environmental Genomics T32 training grant. In summary, I have a demonstrated record of successful and productive research, leadership in organizing research translational collaborations, and in mentoring junior investigators, skills that are relevant to this application.

1. **McConnell R**, Berhane K, Gilliland F, London SJ, Islam T, Gauderman WJ, Avol E, Margolis HG, Peters JM. (2002). Asthma in exercising children exposed to ozone: a cohort study. Lancet, 359(9304):386-91. PubMed PMID: 11844508.
2. **McConnell R**, Berhane K, Gilliland F, Molitor J, Thomas D, Lurmann F, Avol E, Gauderman WJ, Peters JM. (2003). Prospective study of air pollution and bronchitic symptoms in children with asthma. Am J Respir Crit Care Med, 168(7):790-7. PubMed PMID: 12893648.
3. *Shankardass K, **McConnell R**, Jerrett M, Milam J, Richardson J, Berhane K. (2009). Parental stress increases the effect of traffic-related air pollution on childhood asthma incidence. Proc Natl Acad Sci U S A, 106(30):12406-11. PubMed PMID: 19620729; PubMed Central PMCID: PMC2718368.
4. **Garcia E, Urman R, Berhane K, **McConnell R**, Gilliland F. Effects of policy-driven hypothetical air pollutant interventions on childhood asthma incidence in southern California. *Proc Natl Acad Sci U S A.* 2019;116:15883-8. Epub 2019/07/25. doi: 10.1073/pnas.1815678116. PubMed PMID: 31332016; PMCID: PMC6689942

*Student or junior faculty mentored by McConnell

A complete list of over 230 peer reviewed publications is available at: http://www.ncbi.nlm.nih.gov/sites/myncbi/rob.mcconnell.1/bibliography/40704438/public/?sort=date&direction = descending

B. Positions and

Honors

Positions and

Employment

1980 - 1983	Residency in Internal Medicine, Montefiore Hospital, Bronx, NY
1983 - 1985	Epidemic Intelligence Service, Centers for Disease Control
1983 - 1986	Medical Officer, Hazard Evaluations and Technical Assistance Branch, National Institute for Occupational Safety and Health
1984 - 1985	Residency in Occupational Health, National Institute for Occupational Safety and Health
1986 - 1989	Epidemiologist, CARE International and American Friends Service Committee, Leon
1989 - 1990	Environmental medicine fellowship, Charles A. Dana Foundation
1989 - 1993	Assistant Professor (1991-1993), Instructor (1989-1991), Division of Environmental and Occupational Medicine, Dept of Community Medicine, New York, NY
1990 - 1992	Clinical Fellow in Environmental Medicine, U.S. Agency for Toxic Substances and Disease Registry
1993 - 1994	Deputy Director, World Health Organization Collaborating Center in Environmental Epidemiology, Mount Sinai School of Medicine
1994 - 1997	Director, Pan American Center for Human Ecology and Health, Pan American Health Organization/World Health Organization
1997 - 2005	Associate Professor, Division of Occupational and Environmental Health, Department of Preventive Medicine
2005 -	Professor, Division of Environmental Health, Department of Preventive Medicine

2009 - 2019 Deputy Director, Southern California Environmental Health Sciences Center

2013 - Director, Southern California Children's Environmental Health Center

2019 - Director, Southern California Environmental Health Sciences Center

Other Experience and Professional Memberships

1983 - Diplomate, American Board of

Internal Medicine Honors

1976 Phi Beta Kappa, Stanford University

2014 Fellow, American Association for the Advancement of Science

C. Contribution to Science

1. I have addressed gaps in our understanding of the burden and financial cost of disease associated with air pollution and climate change mitigation, using novel methods on a scale relevant to local policy makers.
    a. *Ghosh R, Lurmann F, Perez L, Penfold B, Brandt S, Wilson J, Milet M, Künzli N, **McConnell R** (2016). Near-Roadway Air Pollution and Coronary Heart Disease: Burden of Disease and Potential Impact of a Greenhouse Gas Reduction Strategy in Southern California. Environ Health Perspect 124:193-200. PubMed PMID: 26149207
    b. Brandt S, Perez L, Kunzil, N, Lurman F, Wilson J, Pastor M, **McConnell, R**. (2014). Cost of near-roadway and regional air pollution–attributable childhood asthma in Los Angeles County. J Allergy Clin Immunology, 134: 1028-1035. PubMed PMID: 25439228; PMCID: PMC4257136.
    c. Perez L, Lurmann F, Wilson J, Pastor M, Brandt SJ, Künzli N, **McConnell R** (2012). Near- roadway pollution and childhood asthma: implications for developing "win-win" compact urban development and clean vehicle strategies. Environ Health Perspect. 120:1619-26. PubMed PMID: 23008270
    d. Hu H, Cohen G, Sharma B, Yin H, **McConnell R.** Sustainability in Health Care". Annual Review of Environment and Resources (in press).
2. I have advanced our understanding of the impact of air pollution on childhood respiratory disease through seminal epidemiological studies of lung function and asthma, and local health impact assessment.
    a. Gauderman WJ, Avol E, Gilliland F, Vora H, Thomas D, Berhane K, **McConnell R**, Kuenzli N, Lurmann F, Rappaport E, Margolis H, Bates D, Peters J. (2004). The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med, 351(11):1057-67. PubMed PMID: 15356303.
    b. Gauderman WJ, Vora H, **McConnell R**, Berhane K, Gilliland F, Thomas D, Lurmann F, Avol E, Kunzli N, Jerrett M, Peters J. (2007). Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study. Lancet, 369(9561):571-7. PubMed PMID: 17307103.
    c. **McConnell R**, Islam T, Shankardass K, Jerrett M, Lurmann F, Gilliland F, Gauderman J, Avol E, Künzli N, Yao L, Peters J, Berhane K. (2010). Childhood incident asthma and traffic-related air pollution at home and school. Environ Health Perspect, 118(7):1021-6. PubMed PMID: 20371422; PubMed Central PMCID: PMC2920902.
    d. *Urman R, **McConnell R**, Islam T, Avol EL, Lurmann FW, Vora H, Linn WS, Rappaport EB, Gilliland FD, Gauderman WJ (2014). Associations of children's lung function with ambient air pollution: joint effects of regional and near-roadway pollutants. Thorax. 69:540-7. Pubmied PMID: 24253832
3. I have developed studies of novel risk factors for respiratory disease and their

interactions with air pollution that may provide clues to relevant biological pathways.

e. **McConnell R**, Berhane K, Molitor J, Gilliland F, Künzli N, Thorne PS, Thomas D, Gauderman WJ, Avol E, Lurmann F, Rappaport E, Jerrett M, Peters JM. (2006). Dog ownership enhances symptomatic responses to air pollution in children with asthma. Environ Health Perspect, 114(12):1910-5. PubMed PMID: 17185284; PubMed Central PMCID: PMC1764158.
f. *Islam T, **McConnell R**, Gauderman WJ, Avol E, Peters JM, Gilliland FD. (2008). Ozone, oxidant defense genes and risk of asthma during adolescence. Am J Respir Crit Care Med, 177(4): 388-395. PubMed PMID: 18048809; PubMed Central PMCID: PMC2258440.
g. *Islam T, **McConnell R**, Gauderman WJ, Berhane K, Avol E, Peters JM,Gilliland FD. (2009). Glutathione-S-Transferase (GST) P1, exercise, ozone and asthma incidence in school children. Thorax, 64(3):197-202. PubMed PMID: 18988661; PubMed Central PMCID: PMC2738935.
h. *Islam T, Urman R, Gauderman WJ, Milam J, Lurmann F, Shankardass K, Avol E, Gilliland F,
McConnell R. (2011). Parental Stress Increases the Detrimental Effect of Traffic Exposure on Children's Lung Function. Am J Respir Crit Care Med, 184(7):822-7. PubMed PMID: 21700914; PMCID: PMC3208647.

4. I have contributed to our knowledge of neurological effects of diverse environment exposures in studies of children and workers.
   a. Rosenstock L, Keifer M, Daniell WE, **McConnell R**, Claypoole K. (1991). Chronic central nervous system effects of acute organophosphate pesticide intoxication. The Pesticide Health Effects Study Group. Lancet, 338(8761):223-7. PubMed PMID: 1676786.
   b. *Volk HE, Lurmann F, Penfold B, Hertz-Picciotto I, **McConnell R**. (2013). Traffic-related air pollution, particulate matter, and autism. JAMA Psychiatry, 70(1):71-7. PubMed PMID: 23404082; PubMed Central PMCID: PMC4019010.
   c. *Volk HE, Kerin T, Lurmann F, Hertz-Picciotto I, **McConnell R**, Campbell DB. (2014). Autism spectrum disorder: interaction of air pollution with the MET receptor tyrosine kinase gene. Epidemiology, 25(1):44-7. PubMed PMID: 24240654; PubMed Central PMCID: PMC4019012.
   d. *Jo H, Eckel SP, Wang X, Chen JC, Cockburn M, Martinez MP, Chow T, Molshatzki N, Lurmann FW, Funk WE, Xiang AH, **McConnell R**. (2019). Sex-specific associations of autism spectrum disorder with residential air pollution exposure in a large Southern California pregnancy cohort. Environ Pollut, 254(Pt A):113010. PubMed PMID: 31554142; PubMed Central PMCID: PMC6764604.
5. I have contributed to studies examining associations of air pollution in children with obesogenic and cardiometabolic outcomes.
   a. Breton CV, Wang X, Mack WJ, Berhane K, Lopez M, Islam TS, Feng M, Lurmann F, **McConnell R**, Hodis HN, Künzli N, Avol E. (2012). Childhood air pollutant exposure and carotid artery intima-media thickness in young adults. Circulation, 126(13):1614-20. PubMed PMID: 22896588; PubMed Central PMCID: PMC3474843.
   b. Jerrett M, **McConnell R**, Wolch J, Chang R, Lam C, Dunton G, Gilliland F, Lurmann F, Islam T, Berhane K. (2014). Traffic-related air pollution and obesity formation in children: a longitudinal, multilevel analysis. Environ Health, 13:49. PubMed PMID: 24913018; PubMed Central PMCID: PMC4106205.
   c. **McConnell R**, Gilliland FD, Goran M, Allayee H, Hricko A, Mittelman S. (2015). Does near- roadway air pollution contribute to childhood obesity?. Pediatr Obes, 11(1):1-3. PubMed PMID: 25820202; PMCID: PMC4821543
   d. **McConnell R**, Shen E, Gilliland FD, Jerrett M, Wolch J, Chang CC, Lurmann F,

Berhane K. (2015). A longitudinal cohort study of body mass index and childhood exposure to secondhand tobacco smoke and air pollution: the Southern California Children's Health Study. Environ Health Perspect, 123(4):360-6. PubMed PMID: 25389275; PubMed Central PMCID: PMC4384197.

*Junior investigator mentored by McConnell.

CURRENT & PENDING SUPPORT

PI: Su, Jason					
Status (currently active or pending approval)	**Award #** (if available)	**Source** (name of the sponsor)	**Project Title**	**Start Date**	**End Date**
Pending Approval	22RD011	California Air Resources Board (CARB)	Impacts of Air Pollution on Life Expectancy across Multiple Generations: Race, Ethnicity, and Vulnerability Perspectives	TBD	TBD
Currently Active	22RD010	California Air Resources Board (CARB)	Impact of Air Pollution Exposure on Metabolic Health Outcomes for California Residents	5/1/2023	4/30/2025
Currently Active	21RD004	California Air Resources Board (CARB)	Preterm Birth, Low Birth Weight, Childhood Autism, Parkinson's and Alzheimer Disease and Air Pollution – California Studies	5/1/2022	4/30/2024
Currently Active	R840239	US EPA STAR Program	Participatory Design of Effective Risk Communication about Wildfire Smoke for Hard-to-Reach Populations	9/01/2021	2/29/2024
Currently Active	20RD016	California Air Resources Board (CARB)	Impacts of Train and Port Pollution and Air Toxics on Respiratory Symptoms and ED Visits Within Vulnerable Communities in Southern California	5/20/2021	5/19/2023
Currently Active		Health Effects Institute	Ambient air pollution and COVID-19 in California	5/15/2021	5/14/2024

Consultant: Lee, Sei					
Status (currently active or pending approval)	**Award #** (if available)	**Source** (name of the sponsor)	**Project Title**	**Start Date**	**End Date**
Pending Approval	22RD011	California Air Resources Board (CARB)	Impacts of Air Pollution on Life Expectancy across Multiple Generations: Race, Ethnicity and Vulnerability Perspectives	TBD	TBD
Currently Active	R01AG057751	NIH/NIA	Prognostic calculators for patients with Alzheimers disease and related dementias	8/15/2018	4/30/2023
Currently Active	RF1AG062568	NIH/NIA	Management of Hypertension among Persons with and without Dementia in Long-Term Care	5/5/2019	3/31/2024

Currently Active	K24AG066998	NIH/NIA	Improving Care for Nursing Home Residents with Diabetes and Cognitive Impairment	1/15/2021	12/31/2025
Currently Active	R01AG067427	NIH/NIA	Identifying and supporting patients with undiagnosed dementia using the EHR Risk of Alzheimer's and Dementia Assessment Rule (eRADAR): a pilot clinical trial	7/15/2020	4/30/2024
Currently Active	P01AG066605	NIH/NIA	Deploying High Value Longitudinal Population-Based Data in Dementia Research (DEVELOP AD esearch)	9/30/2020	5/31/2025
Currently Active		The Donaghue Foundation	RAFT COVID: Recovery After LTACH Transfer for Older Adults	12/1/2021	3/31/2023
Currently Active		NIH/NIA	Predicting long-term cognitive outcomes and Alzheimer's disease and related dementias after major noncardiac surgery for older adults	9/1/2022	8/31/2027

Project Advisor (No Salary Requested): McConnell, Rob

Status (currently active or pending approval)	**Award #** (if available)	**Source** (name of the sponsor)	**Project Title**	**Start Date**	**End Date**
Pending Approval	22RD011	California Air Resources Board (CARB)	Impacts of Air Pollution on Life Expectancy across Multiple Generations: Race, Ethnicity, and Vulnerability Perspectives	TBD	TBD
Currently Active	21RD010	California Air Resources Board (CARB)	Impact of Air Pollution Exposure on Metabolic Health Outcomes for California Residents	5/1/2023	4/30/2025
Currently Active	NIH/NIEHS	R01ES03069 1	Hepatotoxic effects of perfluoroalkyl substances: a new epidemiological approach for studying environmental	5/1/2020	4/30/2023
Currently Active	NIH/NIEHS	R01ES030364	Effects of DDE exposure on adipose tissue function, weight loss and metabolic improvement after bariatric surgery: A new paradigm for study of lipophilic chemicals	2/1/2020	11/30/2024
Currently Active	NIH	U54CA180905	Tobacco Regulatory Science Investigating the Intersections of Products with Diverse Populations	9/19/2013	8/31/2023

Currently Active	NIH/NIEHS	R01ES029963	Particulate Air Pollutants and Autism Risk: Exposure Characteristics, Indicators of Susceptibility, and Mechanistic Pathways	8/1/2019	5/31/2023
Currently Active	NIH/NIEHS	P30ES007048	Environmental Exposures, Host Factors, and Human Disease	6/1/2021	2/28/2026
Currently Active	NIH	UH3OD023287	Lifecourse Approach to Developmental Repercussions of Environmental Agents on Metabolic and Respiratory health (LA DREAMERs)	9/21/2016	8/31/2023
Currently Active	NIH	R01ES031247	Prenatal and Early Postnatal Lead Exposure on Childhood and Adolescent Brain, Cognitive and Behavioral Development	8/17/2020	6/30/2025
Currently Active	NIH/NIEHS	R01ES031074	Urban air pollution and neurobehavioral trajectories in the ABCD study	9/1/2020	6/30/2025
Currently Active	NIH/NIEHS	P2CES033433	Southern California Center for Children's Environmental Health Translational Research	12/9/2021	11/30/2026