HHS Public Access

Author manuscript

J Expo Sci Environ Epidemiol. Author manuscript; available in PMC 2024 July 04.

Published in final edited form as:

J Expo Sci Environ Epidemiol. 2024 May ; 34(3): 529–537. doi:10.1038/s41370-023-00607-0.

Application of artificial intelligence in quantifying lung deposition dose of black carbon in people with exposure to ambient combustion particles

Menghui Jiang[1,*], **Chelin Jamie Hu**[2,*], **Cassie L Rowe**[1], **Huining Kang**[1,3], **Xi Gong**[4], **Christopher P Dagucon**[5], **Jialiang Wang**[1], **Yan Lin**[4], **Akshay Sood**[1,6], **Yan Guo**[1,3], **Yiliang Zhu**[1], **Neil E Alexis**[7], **Frank D Gilliland**[8], **Steven A Belinsky**[3,5], **Xiaozhong Yu**[2,†], **Shuguang Leng**[1,3,5,†]

[1.]School of Medicine, University of New Mexico, Albuquerque, NM, USA;

[2.]College of Nursing, University of New Mexico College of Nursing, Albuquerque, NM, USA;

[3.]University of New Mexico Comprehensive Cancer Center, Albuquerque, NM, USA;

[4.]Department of Geography & Environmental Studies, University of New Mexico, Albuquerque, NM, USA;

[5.]Lung Cancer Program, Lovelace Biomedical Research Institute, Albuquerque, NM, USA;

[6.]Miners Colfax Medical Center, Raton, NM, USA;

[7.]Center for Environmental Medicine Asthma and Lung Biology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;

[8.]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

Abstract

Background: Understanding lung deposition dose of black carbon is critical to fully reconcile epidemiological evidence of combustion particles induced health effects and inform the development of air quality metrics concerning black carbon. Macrophage carbon load (MaCL) is a novel cytology method that quantifies lung deposition dose of black carbon, however it has limited feasibility in large-scale epidemiological study due to the labor-intensive manual counting.

Correspondence and requests for reprints should be addressed to Shuguang Leng, Department of Internal Medicine, School of Medicine, University of New Mexico, Albuquerque, USA 87131. sleng@salud.unm.edu; Xiaozhong Yu, College of Nursing, University of New Mexico, Albuquerque, USA 87131. XiYu@salud.unm.edu.

*Contribute equally as first author

†Contribute equally as correspondence author

Authors' contributions

XY and SL conceived of and designed the study; CJH, CLR, CPD, and JW performed the data collection and management; CJH, HK, and SL conducted data analyses and tabulated the results; CJH , HK, and SL interpreted the results and drafted the manuscript; and MJ, XG, YL, AS, YG, YZ, NEL, FDG, SAB, XY and SL critically edited the manuscript. All authors have read the manuscript and approved its submission.

Ethics approval and consent to participate

This study (1054101) was approved by the Western Institutional Review Board and all participants signed consent forms.

Competing interests

The authors declare that they have no competing interests

Objective: To assess the association between MaCL and episodic elevation of combustion particles; to develop artificial intelligence based counting algorithm for MaCL assay.

Methods: Sputum slides were collected during episodic elevation of ambient $PM_{2.5}$ (n=49, daily $PM_{2.5}$ >10 μg/m^3 for over 2 weeks due to wildfire smoke intrusion in summer and local wood burning in winter) and low $PM_{2.5}$ period (n=39, 30-day average $PM_{2.5}$ <4 μg/m^3) from the Lovelace Smokers cohort.

Results: Over 98% individual carbon particles in macrophages had diameter <1 μm. MaCL levels scored manually were highly responsive to episodic elevation of ambient $PM_{2.5}$ and also correlated with lung injury biomarker, plasma CC16. The association with CC16 became more robust when the assessment focused on macrophages with higher carbon load. A Machine-Learning algorithm for Engulfed cArbon Particles (MacLEAP) was developed based on the Mask Region-based Convolutional Neural Network. MacLEAP algorithm yielded excellent correlations with manual counting for number and area of the particles. The algorithm produced associations with ambient $PM_{2.5}$ and plasma CC16 that were nearly identical in magnitude to those obtained through manual counting.

Significance: This study provides support for MaCL as a valuable lung biomarker that effectively indicates episodic increases in exposure to combustion particles. The MacLEAP algorithm demonstrates its robustness in quantifying the lung deposition dose of black carbon, making it suitable for application in large-scale epidemiological studies.

Keywords

combustion-emitted particulate matter; macrophage carbon load; artificial intelligence; lung deposition dose

Introduction

Combustion-emitted particulate matter (CE-PM), formed through incomplete combustion of fossil fuels and biomass from anthropogenic and natural sources, has become a top-priority issue of public health and climate change in the United States (US) and globally[1–3]. CE-PM accounted for 88% of primary $PM_{2.5}$ in ambient air in the US[4] and disproportionately affected older people/children, minorities, rural areas, and high-poverty communities[5–8]. CE-PM was thought to be more harmful to health than PM from non-combustion sources and may be more relevant than $PM_{2.5}$ mass levels for evaluating health risks associated with combustion particles from local sources[9–17]. A US National survey identified that combustion emissions from 7 sectors constituted the largest source of anthropogenic emissions and caused 49,300 to 82,900 premature deaths in 2018[3]. Moreover, CE-PM caused the highest number of premature deaths (46%) among major air pollution species[3], supporting prioritization of CE-PM for advancing air pollution mitigation efforts in the US.

Major CE-PM emissions (i.e., traffic, industry, wildfire, and residential woodsmoke) have temporal and spatial variations, which together with intra- and inter-individual variations for respiratory physiology, affect lung deposition. Understanding lung deposition from CE-PM exposure is critical to fully reconcile epidemiological evidence of air pollution heath effects[18] and is essential for developing air quality metrics concerning black carbon.

The majority of CE-PM is in a nano-scale range and can deposit deep inside the lungs. Black carbon may not be directly toxic, but operates as a universal carrier of a wide variety of combustion-derived chemical constituents[19]. Phagocytosis of black carbon by airway macrophages is a major clearance mechanism in acinar airway and triggers persistent cytokine/chemokine secretion and generation of other mediators for host defense responses[20]. Engulfed black carbon in sputum macrophages can be detected on cytospin slides under the light microscope as "black" particles whose elemental carbon composition has been confirmed using validated methods[21, 22]. Strong evidence suggested that clearance of inhaled environmental PM through airway macrophages is a slow process and may take months to occur, thus forming a mechanistic basis for sustained effects after the cessation of CE-PM exposure, such as wildfire smoke[22–30]. In addition, macrophage carbon load (MaCL) assay can also rapidly detect a mild increase in ambient PM levels[31]. Finally, our studies and others provided extensive evidence supporting the utility of MaCL as a lung deposition biomarker of black carbon-associated PMs for assessing pulmonary and extra-pulmonary outcomes in environmental (low, e.g., traffic related) and occupational (high, e.g., diesel engine testing) exposure settings[22, 24, 27, 28, 31–35].

Manual scoring of MaCL slides using ImageJ/NIH Image is labor-intensive and subject to scorer errors and bias, limiting its utility in large-scale epidemiological studies. Artificial intelligence (AI) based algorithms have shown great capacity for defining biologically or clinically interpretable features based on 2-D images. We developed "Machine-Learning algorithm for Engulfed cArbon Particles (MacLEAP)", a very first-of-its kind AI algorithm that automatically identified the macrophages from other cell types on sputum slide images and quantified engulfed black carbon particles. Our initial efforts of algorithm development applied CellProfiler[36] based on 357 bright-field Papanicolaou staining sputum images from 17 Lovelace Smokers cohort (LSC) members[37]. Although CellProfiler was able to segment and classify cells and carbon particles based on staining intensity and extract features, the performance was compromised by blurry background cells in these sputum images as sputum samples were not treated with mucolytic prior to cytospin. Subsequently, the Mask Region-based Convolutional Neural Network (Mask R-CNN)[38, 39], which is the state-of-the-art instance segmentation framework in the current deep-learning field, was adopted to differentiate airway macrophages from other cell types and quantify nano-scale black carbon particles. Mask R-CNN based MacLEAP algorithm yielded excellent correlations with manual count results (R^2=0.99 for the number of macrophage and R^2>0.83 for the number and area of black carbon particles), exhibited greater robustness on challenging images, and required fewer repetitive trials in the model training steps compared to CellProfiler-based algorithm (see Supplemental Materials)[37].

In the past two decades, annual ambient $PM_{2.5}$ levels ranged from 5.2 μg/m^3 to 7.1 μg/m^3 in Albuquerque, New Mexico (NM). Major CE-PM emission sources include year-long traffic emissions, wildfire smoke invasion in summer, and residential wood burning in winter. In this study, we conducted cross-sectional analyses of 88 LSC members to assess whether lung dose of black carbon increased amid episodic elevation of ambient $PM_{2.5}$ levels due to summer wildfire smoke invasion or winter wood burning. Based on our initial version of Mask R-CNN-based MacLEAP algorithm[37], we further enhanced the algorithm using a much larger set of images (n=1043) with more annotation of macrophages and black

carbon particles, and then applied the algorithm in 746 sputum images including 333 freshly-stained sputum images as an indepdedent validation set. The performance of the MacLEAP algorithm was assessed by comparing individual-based MaCL levels between MacLEAP and manual counting. Lastly, we assessed whether association analyses with $PM_{2.5}$ or lung injury biomarker, i.e., Club cell secretory protein (CC16), based on MacLEAP data could reproduce associations seen using manual scoring data.

Materials and Methods

Definition of study periods

Episodic elevation of ambient $PM_{2.5}$ levels was defined as elevation in daily ambient $PM_{2.5}$ (>10 μg/m^3) for over 2 weeks based on Environmental Protection Agency air quality data in the catchment area of the LSC, i.e., Albuquerque[40]. Seven periods were identified prior to 2010 when sputum slides were available for all visits in the LSC and had average $PM_{2.5}$ levels for the entire periods >10 μg/m^3 with peak daily average ranging from 19.8 to 46.8 μg/m^3 (Supplemental Table 1). These seven periods ranged from 14 days to 71 days with an average of 46 days. Analyses of NASA (The National Aeronautics and Space Administration) and NOAA (National Oceanic and Atmospheric Administration) satellite data suggested wildfire smoke intrusion from the southwest of Albuquerque in summer based on timing concordance, smoke plume shapefiles, and major wind directions[41]. Local wood burning for heating during weather inversion most likely contributed to the episodic elevation of ambient $PM_{2.5}$ in winter[42]. We also identified multiple periods with low $PM_{2.5}$ levels (30-day average $PM_{2.5}$ <4 μg/m^3).

Study subjects from the LSC

The LSC was established in 2001 to study sputum and blood biomarkers for risk assessment of lung cancer and chronic obstructive pulmonary disease in current and former smokers from Albuquerque, NM. Recruitment strategy and eligibility have been described for the ~2500 subjects enrolled[43, 44]. At study entry, cohort members completed a health questionnaire, underwent pre- and post- bronchodilator spirometry, and provided blood and induction sputum. Cohort members returned every 18 months to update smoking and medical history, underwent spirometry, and provided biological samples. The LSC followed standard induction procedure to collect sputum samples during each visit. The procedure involved inhaling a saline mist via a nebulizer to stimulate mucus production, followed by a deep cough to expel sputum for analysis[25]. Sputum samples were stored in Saccomanno solution up to three months at controlled room temperature and then washed with Saccomanno solution (without mucolytic treatment) prior to being split into two aliquots: one pelleted for long-term storage at −80°C for genomic analyses and the other spun onto slides using cytospin and stained with Papanicolaou Staining. Saccomanno solution was made in the lab by mixing 20 ml melted polyethylene glycol 1500 (A16241.0B, Thermo Scientific), 480 ml Milli-Q water, and 500 ml ethanol (V1001, Decon labs). Papanicolaou Staining kit contained Cyto-Stain™ (7501R, Epredia™), Hematoxylin 7211 (72–11L, Epredia™), and Bluing Reagent (7301, Epredia™). This study (1054101) was approved by the Western Institutional Review Board and all participants signed consent forms. Sputum samples collected within the seven periods with episodic elevation of

ambient $PM_{2.5}$ levels were selected. Sputum samples collected within the first two weeks of the seven periods were avoided because studies showed two weeks of exposure to a mild increase in ambient PM levels led to a detectable increase in MaCL levels[31]. Sputum slides collected at the end of each clean periods, i.e., 30-day average $PM_{2.5}$ <4 μg/m^3 were selected. A total of 49 and 39 slides collected during episodic elevation and clean periods were chosen, respectively; among which 66 slides that were stained prior to archiving were used in MacLEAP algorithm enhancement and 22 slides freshly stained were used as an independent validation.

Inclusion of current and former smokers

This study included both current and formers smokers from the LSC. This inclusion was supported by fact that 1) the literature reported no significant associations of area of particles (AoP), as the most commonly used MaCL measure, with smoking status (current versus non-current), cigarettes per day, duration of smoking, packyears, exposure to second-hand smoking, or urinary cotinine[28, 32–34, 45], and 2) elemental carbon comprised only <1% of cigarette smoke particles[46, 47] and Saccomanno's solution removed any color from residual tar.

Image acquisition and quantification of MaCL

Fields containing macrophages were randomly selected from each sputum slide and were imaged manually under a 100× oil immersion lens mounted on an Olympus microscope BX43 with a DP28 camera. Through a motorized Z drive controlled by Olympus cellSens software, we acquired Z-stack images with 100 nm as the depth interval to cover the entire cell depth. Stack depth at 100 nm is the finest step size configurable in Olympus cellSens software that controls the motorized Z drive. Increasing the step size was likely to miss some smaller particles or acquire less stack images with black carbon particles. This lowered the quality of projection images for detecting black carbon particles either manually or using MacLEAP algorithm. A flattened image at 3840 × 2160 pixels was generated by stacking all stack images with most contrasted features at each depth projected. Images for 50 macrophages per slide were acquired. A well-trained technician who was blind to the exposure status conducted image acquisition and used ImageJ/NIH Image to quantify MaCL endpoints. Number and area of macrophages were recorded for each macrophage, then median values of number and area of particles for all macrophages scored per individual were calculated as lung dose metric for subsequent analyses.

Mask R-CNN based MacLEAP algorithm

MacLEAP was developed based on a free, open-source deep convolutional neural network Mask R-CNN[38]. We adopted a cascaded approach to first recognize macrophages, and then recognize and quantify black carbon particles in these macrophages (Figure 1). Detailed description of the algorithm development and its application is in the Supplemental Materials.

Statistical analyses

Variables were summarized by $PM_{2.5}$ exposure status (high versus low) using statistics based on their nature and distribution (Table 1). The associations between $PM_{2.5}$ exposure in different time frames and MaCL measurements were assessed using over-dispersion controlled Poisson (number of particles [NoP]), Gamma (AoP), and Linear (% cells with particles [%CWP]) regressions with adjustment for age, sex, smoking quit time, packyears, overweight or obesity, ethnicity, and ever wood smoke exposure as these variables were reported to affect lung health[48, 49]. Analyses were also conducted without adjustment of these covariates (Supplemental Table 2) and found little difference in associations compared to analyses with covariate adjustment (Table 2). Samples with median AoPs equal to 0 were given a value of 0.0155 μm^2, half of the none-zero minimal in all subjects to allow log transformation in the Gamma regression analyses. Second, Linear regression was used to assess the associations between MaCL levels and a lung injury biomarker (plasma CC16) in a subset of subjects (n=48) with CC16 data available[50]. Third, because NoPs and AoPs showed consistent associations with $PM_{2.5}$ and CC16 based on manual counting, we used these two individual-based MaCL measurements to further evaulate the performance of MacLEAP. All statistical analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Demographics and ambient $PM_{2.5}$ levels

The distribution of demographic variables was similar between high versus low $PM_{2.5}$ exposure groups (Table 1). Large differences in $PM_{2.5}$ levels were observed up to 2 months prior to sputum collection, however differences reduced for 3, 6, 12, and 18 months, consistent with our definition of episodic exposure to ambient combustion particles and low annual $PM_{2.5}$ levels in this area (Figure 2A). It is important to note that even with episodic elevation of $PM_{2.5}$, average $PM_{2.5}$ in the week prior to sputum collection ($PM_{2.5}$-7d) peaked around 23.3 $\mu g/m^3$ (Figure 2A).

Associations between ambient $PM_{2.5}$ and manually counted MaCL

A total of 12,130 carbon particles were identified manually from 4430 macrophages from 88 subjects and were quantified for MaCL endpoints. By observing the 3-D structure of black carbon particles, we identified 1009 macrophages with only one particle engulfed among which 98% of these particles had diameters less than 1 μm (Figure 2C). These individual particles peak around 200–500 nm in diameter. A sharp decline in fractions of particles smaller than 200 nm was mainly due to fact that particles smaller than that size can not be reliably detected under 100× oil immersion objective of numerical aperture 1.25[51]. NoPs and AoPs were significantly higher in high versus low $PM_{2.5}$ exposure groups by at least 44% (Ps <0.05, Table 1). %CWPs were higher in high versus low $PM_{2.5}$ exposure groups but did not reach statistically significance (P =0.11). Because a large variance was observed for $PM_{2.5}$ levels in the high exposure group (Figure 2A), we further include $PM_{2.5}$ levels from different time frame as a continuous variable and assessed their associations with MaCL endpoints. All three MaCL endpoints were associated with $PM_{2.5}$ levels calculated within the time frame of episodic elevation (7 days to 2 months) with most significant associations

observed for 7d average (Table 2, Figure 2B). Each 5 μg/m^3 increase in $PM_{2.5}$-D7 levels was associated with 48%, 31%, and 3.7% increase in NoPs, AoPs, and % CWPs, respectively (Table 2). Moreover, we also identified significant associations of MaCL endpoints with $PM_{2.5}$ levels calculated for longer periods (3 to 18 months, Table 2). Similar associations were identified using long-term $PM_{2.5}$ average calculated by excluding $PM_{2.5}$ data in two months immediately prior to sputum collection to minimize the impact of recent episodic $PM_{2.5}$ elevation on calculating long-term $PM_{2.5}$ average (Supplemental Table 3). Magnitude of associations increased for longer periods (Table 2) because of lower average and standard deviation for $PM_{2.5}$ levels (Figure 2A). For all above analyses, NoPs and AoPs were much more sensitive to ambient $PM_{2.5}$ than %CWPs. Analyses of determinants for MaCL endpoints did not identify significant effects from age, sex, smoking quit-time, packyears, BMI, or ethnicity (Supplemental Table 4). However, we identified a significant association between slef-reported ever woodsmoke exposure and %CWPs (estimate = 8.73%, P = 0.03).

Associations between manually counted MaCL and plasma CC16

CC16 has anti-inflammatory and antioxidant properties in the lung, and its blood levels reduced in cigarette smokers and populations with high $PM_{2.5}$ exposures due to injury of Club cells in the distal airways[52, 53]. Among a subset of LSC subjects (n=48) with plasma CC16 levels available[50], NoPs and AoPs in macrophages were associated with lower plasma CC16 levels with each inter-quartile range associated with up to 19% reduction (Table 3, Figure 3A). However, no significant associations with %CWPs were identified (P=0.33). Additional analyses were conducted to assess whether macrophages with higher carbon content generate higher lung injury using different cut points of percentage of cell area occupied by black carbon. As shown in Table 4, with the increasing cut points for defining macrophages with higher carbon content, stronger associations between %cells with higher carbon load and plasma CC16 were observed and associations reached statistical significance when using 0.46%, the 90^{th} percentile of carbon load, as the cut point. Moreover, a strong linear relationship was identified between thresholds for defining macrophages with higher carbon occupancy and effect sizes (R^2=0.95, Figure 3C).

Performance of MacLEAP algorithm

MacLEAP algorithm successfully recognized 96.3% macrophages from all 1789 sputum images. Visual inspection of images suggested that most missing macrophages occurred in the second batch for independent validation of the MacLEAP algorithm probably due to fresher staining. Slides for independent validation were from the LSC as well, but were stored as unstained in a cold room. Color from first batch for algorithm development may fade to some degree due to long-term storage (>15 years) after staining at room temperature. It is important to note that MacLEAP algorithm also recognized more macrophages (3.5%) and visual inspection of these images suggested that the gain was mainly due to recognition of macrophages with irregular shapes, broken cytoplasm, or not stained perfectly for nuclear material, which were excluded during manual counting. Excellent correlations were identified for mean number (R^2=0.87, Figure 4D) and median area (R^2=0.82, Figure 4A) of particles between manual and AI counting in 66 subjects used for the enhancement of the MacLEAP algorithm. Excellent correlations were also identified for mean number (R^2=0.76, Figure 4E) and median area (R^2=0.91, Figure 4B) of particles between manual and AI

counting in 22 subjects as the independent validation set. There were two and three samples with zero value for both manual and MacLEAP-calculated AoPs in Figure 4A and 4B, respectively. Removing these five samples only slightly reduced the correlation coefficience (R^2=0.82 in Figure 4A and R^2=0.89 in Figure 4B). Both sets combined provided overall R^2 values of 0.86 and 0.83 for mean NoPs (Figure 4F) and median AoPs (Figure 4C) respectively between manual and AI counting in all 88 subjects. The linear regression lines were very close to the 1:1 line for median AoPs between manual and MacLEAP counting (Figure 4C). The linear regression lines deviated slightly from the 1:1 line for NoPs (Figure 4F). The three samples that drove the slight deviation contained macrophages with high carbon load which challenged manual counting. Association analyses with ambient $PM_{2.5}$ and plasma CC16 based on AI measurements identified results (Supplemental Tables 5 and 6, Figure 2 and 3) almost identical to those seen based on manual counting.

Discussions

This is the first study that quantified airway MaCL using an improved image acquisition system (i.e., projection image) in a US population with relatively low annual ambient $PM_{2.5}$ levels (around 6 μg/m^3) that approach the new standard (5 μg/m^3) of the World Health Organization Air Quality Guidelines 2021[54]. AoPs in all 88 subjects ranged from 0 to 0.83 μm^2 with 0.11 μm^2 as the median, which was lower than that (>0.19 μm^2) seen in European populations with ambient annual $PM_{2.5}$ levels >20 μg/m^3 [22, 26] and much lower than that (>1 μm^2) seen in Asian populations with ambient annual $PM_{2.5}$ levels >40 μg/m^3 [22, 55]. MaCL levels were highly responsive to episodic (2–10 weeks) elevation of ambient $PM_{2.5}$ levels (by 6 μg/m^3 in average) resulting from wildfire smoke invasion in summer or residential wood burning for winter heating. Our findings were supported by a panel study of 20 elderly subjects from Flanders, Belgium with annual PM_{10} at 20–30 μg/m^3, which showed a significant increase of AoPs by 0.54 μm^2 after only 10 days of stay in Milan, Italy with annual PM_{10} at 40–50 μg/m^3 [31]. Moreover, >98% individual black carbon particles in macrophages had diameter <1 μm, suggesting the importance of focusing on PM1-related black carbon levels for health effect studies and development of a new air quality metric for black carbon as prioritized by the WHO Air Quality Guidelines 2021 – “Aiming for Heathier Air for All”[9, 11].

MaCL had a constant declining of 0.006–0.013 μm^2 particle area per day post exposure[22, 28]. Depending on peak exposure levels, estimated clearance half-lives of MaCL levels ranged from months to years[22, 25], which were also affected by lifespan (81 days[56]) and subtype of macrophages, mucociliary clearance, and toxicity of inhalants. MaCL levels also correlated with the 6- to 12-month average of PM_{10} estimated at the participants’ residence in studies conducted in Europe[26, 27]. Our identification of associations of MaCL with long-term (6 to 18 months) $PM_{2.5}$ exposure further supports MaCL as a lung dose biomarker for long-term PM exposure.

Current study identified associations between MaCL levels and the lung injury biomarker CC16 in plasma. Smaller particles are more toxic as they can access and deposit deep in the lungs, however our findings also suggest that engulfed particles may be more toxic with larger aggregate sizes. Phagocytosis is a major clearance mechanism for airway

pathogens, however engulfed carbon particles can not be “digested” and may cumulate in phago-lysosome and trigger persistent inflammatory reactions that mediate pulmonary and extra pulmonary health effects[20]. Thus, MaCL not only reflects lung deposition dose of black carbon, but also serves as a bio-effective biomarker that may be directly involved in the pathogenesis of lung injury[32, 33].

Our current analyses of 88 current and former smokers did not identify statistically significant associations of MaCL endpoints with smoking quit-time and packyears, though people who quit smoking tend to have lower MaCL levels. These findings together with our previous studies and others[28, 32–34, 45–47] do not support the previous practice for excluding current smokers in studies involving MaCL. Instead, smoking history can be either matched at the design stage, or adjusted as covariates or stratified at the data analysis stage. Subjects who were over-weighted or obese had lower MaCL levels, but the difference did not reach statistical significance. Our study also did not find influence of age on MaCL levels. Thus, our studies and others[24, 26, 27, 34] support that the use of MaCL as a lung deposition dose of black carbon is robust and less likely to be confounded by demographic variables including age, cigarette smoking, and BMI. In the LSC, female smokers were initially recruited in 2000, with males added beginning in 2004. The inclusion of only six males in this study was mainly due to the fact that most of the sputum slides pulled for this study were collected prior to 2007.

Instead of detecting and segmenting multiple-class objects (macrophage and carbon) simultaneously from the same images, the two-step detection method extracted the large distinct macrophage cells from the background, and then a second ResNet-backboned Mask R-CNN was dedicated to detect the uniformly dense and often featureless black carbon particles. This allowed for higher precision in instance segmentation and quantification, and ultimately faster for assessing carbon load per macrophage. In addition, extensive masking of features for training in a large number of images is essential for development of a robust AI algorithm. In our study, feature masking was conducted by referring to the detailed scoring sheets that recorded number and locations of macrophages and black carbon particles on each sputum image. Based on our initial version of Mask R-CNN based MacLEAP algorithm[37], we further strengthened the algorithm training using a much larger set of images and then assessed the performance using an independent set of images that were freshly stained. Mask R-CNN-based MacLEAP algorithm yielded excellent correlations with manual counting, a “well-regarded” gold standard for MaCL scoring and generated an almost identical magnitude of associations of MaCL with $PM_{2.5}$ and CC16, as compared to manual counting. Our ongoing studies of residential wood smoke exposure were designed to assess MaCL in spontaneous sputum samples collected in subjects enrolled from Alaska, Montana, Boise, and NM. The MacLEAP algorithm demonstrated high flexibility and openness for additional training using sputum images from different populations and/or collected and stained using different protocols. Future plan to improve the MacLEAP algorithm is to implement more cost-effective AI deep neural network platforms such as PyTorch instead of TensorFlow to handle both analytical and computational challenges.

It is important to note that our improved image acquisition system is semi-automatic and scorer-independent, and greatly reduces the burden associated with image taking as performers only need to define top and bottom of the images under each field and apply universal stack depth (i.e., 100 nm). Thus, our overall MaCL image acquisition and analysis system is estimated to be able to extract data from 50 macrophages on an average-quality slide in less than one hour. This entire process is now scorer-independent and can be done by multiple technicians without lengthy training and tedious quality assurance. All those innovative improvements make MaCL assay ready for large-scale epidemiological studies to integrate the long-missing lung deposition dose component in risk assessment of CE-PM exposure.

Conclusions

In summary, this study added new evidence to support MaCL as a lung bio-effective biomarker for black carbon that was responsive to episodic elevation of CE-PM exposure and also correlated with long-term $PM_{2.5}$ exposure. The influence of age, cigarette smoking, and BMI seemed minimal. The MacLEAP algorithm based on Mask R-CNN in conjunction with our improved image acquisition system made measuring MaCL as lung deposition dose of black carbon feasible in large-scale epidemiological studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Our initial work of establishing the MacLEAP: Machine-Learning algorithm for Engulfed cArbon Particles was presented as an abstract and in details as an ePoster at the Society of Toxicology 61th Annual Meeting in San Diego 2022. The ePoster is attached as a supplemental material for this article.

Funding

Supported by National Cancer Institute grant P30 CA118100 (Leng), National Institute of General Medical Sciences U54 GM104944 (Leng), National Institute of Environmental Health grant NM-INSPIRES P30 ES032755 (Yu) and UNM-METALS P42ES025589 (Yu and Lin), and National Institute on Minority Health and Health Disparities P50MD015706 (Lin).

Availability of data and materials

Sputum images and MaCL measures were submitted to the Lovelace Smokers cohort with Dr. Steven A. Belinsky as the Principal Investigator for the cohort. The Lovelace Respiratory Research Institute owns the cohort and manages all data and tissue request to ensure compliance with the Institutional Rreview Board protocol, consent form, institutional regulations, as well as related National Institute of Health policies. Ms. Maria Picchi as the data manager for the Lovelace Smokers cohort is the contact person for handling any request for data and samples from the Lovelace Smokers cohort. Researchers who would like to request training script and masking images of macrophages and black carbons need to reach out to Ms. Picchi and correspondence authors. Finalized weigh file for MacLEAP algorithm and quantification/application script are now shared via GitHub Platform (https://github.com/yuxz99/MacLEAP).

Reference

1. Cordero RR, Sepúlveda E, Feron S, Damiani A, Fernandoy F, Neshyba S et al. Black carbon footprint of human presence in Antarctica. Nature Communications 2022; 13.
2. The Lancet Public H Mitigating climate change must be a priority for public health. The Lancet Public Health 2021; 6.
3. Dedoussi IC, Eastham SD, Monier E, Barrett SRH Premature mortality related to United States cross-state air pollution. Nature 2020; 578: 261–265. [PubMed: 32051602]
4. Thurston GD, Ito K, Lall R A source apportionment of U.S. fine particulate matter air pollution. Atmos Environ (1994) 2011; 45: 3924–3936. [PubMed: 24634604]
5. Disparities in the Impact of Air Pollution. In.
6. Rogalsky DK, Mendola P, Metts TA, Martin WJ 2nd, Estimating the number of low-income americans exposed to household air pollution from burning solid fuels. Environ Health Perspect 2014; 122: 806–810. [PubMed: 24833615]
7. Noonan CW, Ward TJ, Semmens EO Estimating the number of vulnerable people in the United States exposed to residential wood smoke. Environ Health Perspect 2015; 123: A30.
8. Rogalsky DK, Mendola P, Metts TA, Martin WJ 2nd, Estimating the number of vulnerable people in the United States exposed to residential wood smoke: Rogalsky et al. Respond. Environ Health Perspect 2015; 123: A30–31.
9. WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. : Geneva: World Health Organization; 2021, 2021.
10. Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H et al. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environ Health Perspect 2011; 119: 1691–1699. [PubMed: 21810552]
11. Grahame TJ, Klemm R, Schlesinger RB Public health and components of particulate matter: the changing assessment of black carbon. J Air Waste Manag Assoc 2014; 64: 620–660. [PubMed: 25039199]
12. Aguilera R, Corringham T, Gershunov A, Benmarhnia T Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. Nat Commun 2021; 12: 1493. [PubMed: 33674571]
13. Thurston GD, Burnett RT, Turner MC, Shi Y, Krewski D, Lall R et al. Ischemic Heart Disease Mortality and Long-Term Exposure to Source-Related Components of U.S. Fine Particle Air Pollution. Environ Health Perspect 2016; 124: 785–794. [PubMed: 26629599]
14. Ostro B, Hu J, Goldberg D, Reynolds P, Hertz A, Bernstein L et al. Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: results from the California Teachers Study Cohort. Environ Health Perspect 2015; 123: 549–556. [PubMed: 25633926]
15. Beelen R, Hoek G, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G et al. Natural-cause mortality and long-term exposure to particle components: an analysis of 19 European cohorts within the multi-center ESCAPE project. Environ Health Perspect 2015; 123: 525–533. [PubMed: 25712504]
16. Hime NJ, Marks GB, Cowie CT A Comparison of the Health Effects of Ambient Particulate Matter Air Pollution from Five Emission Sources. Int J Environ Res Public Health 2018; 15.
17. Martenies SE, Keller JP, WeMott S, Kuiper G, Ross Z, Allshouse WB et al. A Spatiotemporal Prediction Model for Black Carbon in the Denver Metropolitan Area, 2009–2020. Environ Sci Technol 2021; 55: 3112–3123. [PubMed: 33596061]
18. Jung KH, Goodwin KE, Perzanowski MS, Chillrud SN, Perera FP, Miller RL et al. Personal Exposure to Black Carbon at School and Levels of Fractional Exhaled Nitric Oxide in New York City. Environ Health Perspect 2021; 129: 97005. [PubMed: 34495741]
19. Cassee FR, Heroux ME, Gerlofs-Nijland ME, Kelly FJ Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission. Inhal Toxicol 2013; 25: 802–812. [PubMed: 24304307]
20. Nagappan A, Park SB, Lee SJ, Moon Y Mechanistic Implications of Biomass-Derived Particulate Matter for Immunity and Immune Disorders. Toxics 2021; 9.

21. Bunn HJ, Dinsdale D, Smith T, Grigg J Ultrafine particles in alveolar macrophages from normal children. Thorax 2001; 56: 932–934. [PubMed: 11713355]
22. Bai Y, Bove H, Nawrot TS, Nemery B Carbon load in airway macrophages as a biomarker of exposure to particulate air pollution; a longitudinal study of an international Panel. Part Fibre Toxicol 2018; 15: 14. [PubMed: 29540230]
23. Orr A, C ALM, Buford M, Ballou S, Migliaccio CT Sustained Effects on Lung Function in Community Members Following Exposure to Hazardous PM2.5 Levels from Wildfire Smoke. Toxics 2020; 8.
24. Bai Y, Brugha RE, Jacobs L, Grigg J, Nawrot TS, Nemery B Carbon loading in airway macrophages as a biomarker for individual exposure to particulate matter air pollution - A critical review. Environ Int 2015; 74: 32–41. [PubMed: 25318022]
25. Cheng W, Liu Y, Tang J, Duan H, Wei X, Zhang X et al. Carbon content in airway macrophages and genomic instability in Chinese carbon black packers. Arch Toxicol 2020; 94: 761–771. [PubMed: 32076763]
26. Jacobs L, Emmerechts J, Mathieu C, Hoylaerts MF, Fierens F, Hoet PH et al. Air pollution related prothrombotic changes in persons with diabetes. Environ Health Perspect 2010; 118: 191–196. [PubMed: 20123602]
27. Kulkarni N, Pierse N, Rushton L, Grigg J Carbon in airway macrophages and lung function in children. N Engl J Med 2006; 355: 21–30. [PubMed: 16822993]
28. Cheng W, Liu Y, Tang J, Duan H, Wei X, Zhang X et al. Carbon content in airway macrophages and genomic instability in Chinese carbon black packers. Archives of Toxicology 2020; doi 10.1007/s00204-020-02678-6.
29. Eguiluz-Gracia I, Schultz HH, Sikkeland LI, Danilova E, Holm AM, Pronk CJ et al. Long-term persistence of human donor alveolar macrophages in lung transplant recipients. Thorax 2016; 71: 1006–1011. [PubMed: 27329043]
30. Nayak DK, Zhou F, Xu M, Huang J, Tsuji M, Hachem R et al. Long-Term Persistence of Donor Alveolar Macrophages in Human Lung Transplant Recipients That Influences Donor-Specific Immune Responses. Am J Transplant 2016; 16: 2300–2311. [PubMed: 27062199]
31. Bai Y, Casas L, Scheers H, Janssen BG, Nemery B, Nawrot TS Mitochondrial DNA content in blood and carbon load in airway macrophages. A panel study in elderly subjects. Environ Int 2018; 119: 47–53. [PubMed: 29933237]
32. Cao X, Lin L, Sood A, Ma Q, Zhang X, Liu Y et al. Small Airway Wall Thickening Assessed by Computerized Tomography Is Associated With Low Lung Function in Chinese Carbon Black Packers. Toxicol Sci 2020; 178: 26–35. [PubMed: 32818265]
33. Liu H, Li J, Ma Q, Tang J, Jiang M, Cao X et al. Chronic exposure to diesel exhaust may cause small airway wall thickening without lumen narrowing: a quantitative computerized tomography study in Chinese diesel engine testers. Part Fibre Toxicol 2021; 18: 14. [PubMed: 33766066]
34. Miri M, Rezaei H, Momtaz SM, Najafi ML, Adli A, Pajohanfar NS et al. Determinants of carbon load in airway macrophages in pregnant women. Environ Pollut 2022; 297: 118765. [PubMed: 34973383]
35. Kulkarni NS, Prudon B, Panditi SL, Abebe Y, Grigg J Carbon loading of alveolar macrophages in adults and children exposed to biomass smoke particles. Sci Total Environ 2005; 345: 23–30. [PubMed: 15919524]
36. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics 2021; 22: 433. [PubMed: 34507520]
37. MacLEAP: Machine-Learning approach for recognition and quantification of Carbon Content in Airway Macrophages Society of Toxicology 61th Annual Meeting San DIego, 2022.
38. He K GG, Dollár P and Girshick R Mask R-CNN. In: IEEE International Conference on Computer Vision (ICCV), 2017. pp 2980–2988.
39. W A. Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. GitHub repository. In, 2017.
40. Download Daily Data. In.

41. Martenies SE, Hoskovec L, Wilson A, Allshouse WB, Adgate JL, Dabelea D et al. Assessing the Impact of Wildfires on the Use of Black Carbon as an Indicator of Traffic Exposures in Environmental Epidemiology Studies. Geohealth 2021; 5: e2020GH000347.
42. Glojek K, Močnik G, Alas HDC, Cuesta-Mosquera A, Drinovec L, Gregorič A et al. The impact of temperature inversions on black carbon and particle mass concentrations in a mountainous area. Atmospheric Chemistry and Physics 2022; 22: 5577–5601.
43. Stidley CA, Picchi MA, Leng S, Willink R, Crowell RE, Flores KG et al. Multivitamins, folate, and green vegetables protect against gene promoter methylation in the aerodigestive tract of smokers. Cancer Res 2010; 70: 568–574. [PubMed: 20068159]
44. Leng S, Stidley CA, Willink R, Bernauer A, Do K, Picchi MA et al. Double-strand break damage and associated DNA repair genes predispose smokers to gene methylation. Cancer Res 2008; 68: 3049–3056. [PubMed: 18413776]
45. Tejwani V, Woo H, Liu C, Tillery AK, Gassett AJ, Kanner RE et al. Black carbon content in airway macrophages is associated with increased severe exacerbations and worse COPD morbidity in SPIROMICS. Respir Res 2022; 23: 310. [PubMed: 36376879]
46. Daisey JM, Mahanama KR, Hodgson AT Toxic volatile organic compounds in simulated environmental tobacco smoke: emission factors for exposure assessment. J Expo Anal Environ Epidemiol 1998; 8: 313–334. [PubMed: 9679214]
47. Hildemann LM, Markowski GR, Cass GR Chemical composition of emissions from urban sources of fine organic aerosol. Environmental Science & Technology 1991; 25: 744–759.
48. Bruse S, Sood A, Petersen H, Liu Y, Leng S, Celedon JC et al. New Mexican Hispanic smokers have lower odds of chronic obstructive pulmonary disease and less decline in lung function than non-Hispanic whites. Am J Respir Crit Care Med 2011; 184: 1254–1260. [PubMed: 21908412]
49. Leng S, Picchi MA, Meek PM, Jiang M, Bayliss SH, Zhai T et al. Wood smoke exposure affects lung aging, quality of life, and all-cause mortality in New Mexican smokers. Respir Res 2022; 23: 236. [PubMed: 36076291]
50. Petersen H, Leng S, Belinsky SA, Miller BE, Tal-Singer R, Owen CA et al. Low plasma CC16 levels in smokers are associated with a higher risk for chronic bronchitis. Eur Respir J 2015; 46: 1501–1503. [PubMed: 26160869]
51. Silfies JS SS, Davidson MW. The Diffraction Barrier in Optical Microscopy. In.
52. Wang H, Duan H, Meng T, Yang M, Cui L, Bin P et al. Local and Systemic Inflammation May Mediate Diesel Engine Exhaust-Induced Lung Function Impairment in a Chinese Occupational Cohort. Toxicol Sci 2018; 162: 372–382. [PubMed: 29186624]
53. Shijubo N, Itoh Y, Yamaguchi T, Shibuya Y, Morita Y, Hirasawa M et al. Serum and BAL Clara cell 10 kDa protein (CC10) levels and CC10-positive bronchiolar cells are decreased in smokers. Eur Respir J 1997; 10: 1108–1114. [PubMed: 9163654]
54. Hoffmann B, Boogaard H, de Nazelle A, Andersen ZJ, Abramson M, Brauer M et al. WHO Air Quality Guidelines 2021-Aiming for Healthier Air for all: A Joint Statement by Medical, Public Health, Scientific Societies and Patient Representative Organisations. Int J Public Health 2021; 66: 1604465. [PubMed: 34630006]
55. Yin G, Wu X, Wu Y, Li H, Gao L, Zhu X et al. Evaluating carbon content in airway macrophages as a biomarker of personal exposure to fine particulate matter and its acute respiratory effects. Chemosphere 2021; 283: 131179. [PubMed: 34146873]
56. Thomas ED, Ramberg RE, Sale GE, Sparkes RS, Golde DW Direct evidence for a bone marrow origin of the alveolar macrophage in man. Science 1976; 192: 1016–1018. [PubMed: 775638]

Impact statement

Understanding lung black carbon deposition is crucial for comprehending health effects of combustion particles. We developed “Machine-Learning algorithm for Engulfed cArbon Particles (MacLEAP)”, the first artificial intelligence algorithm for quantifying airway macrophage black carbon. Our study bolstered the algorithm with more training images and its first use in air pollution epidemiology. We revealed macrophage carbon load as a sensitive biomarker for heightened ambient combustion particles due to wildfires and residential wood burning.

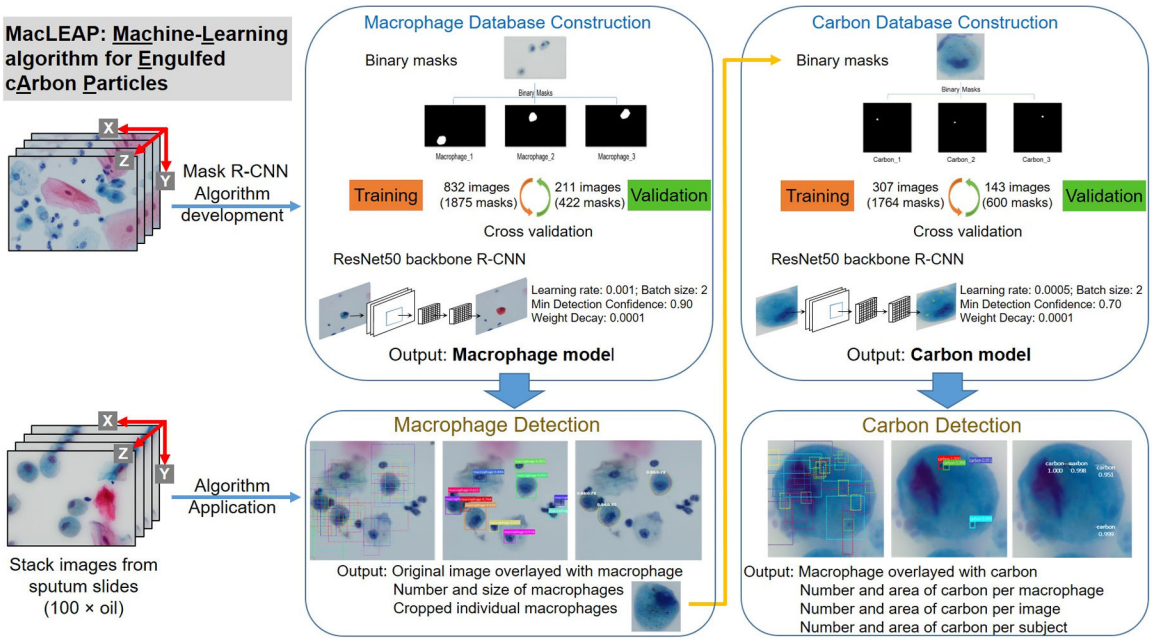

Figure 1. Pipeline for the Development of the Machine-Learning algorithm for Engulfed cArbon Particles (MacLEAP) algorithm.

A cascaded approach was adopted to develop macrophage and carbon particle models using the Mask R-CNN. A subset (n=1043) of sputum images from 66 LSC subejcts were used to improve the MacLEAP algorithm. A total of 2297 macrophages were masked manually on these 1043 projection images to enhance the macrophage model. A total of 2364 black carbon particles were masked manually on 450 cropped macrophages to develop the black carbon model. The optimized algorithm was then applied to the rest 413 sputum images from the 66 subjects and all 333 sputum images from the 22 subjects.

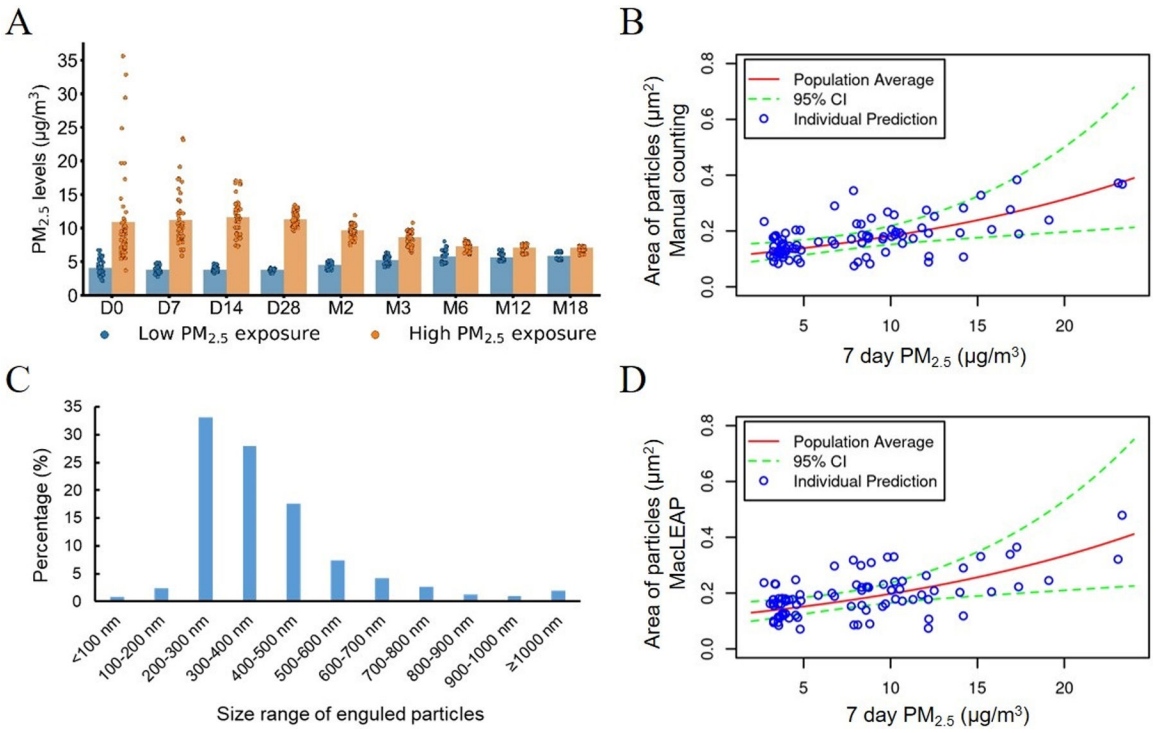

Figure 2. Episodic elevation of $PM_{2.5}$ and areas of the particles in macrophages.

Large differences in $PM_{2.5}$ levels were observed up to 2 months prior to sputum collection, however differences reduced for 3 months and beyond (**A**). Over 98% of engulfed individual particles had diameters less than 1 μm with 78.6% between 200 and 500 nm (**C**). 7d average of $PM_{2.5}$ was less than 5 μg/m^3 for low $PM_{2.5}$ group, but ranged from 6 to 25 μg/m^3 for high $PM_{2.5}$ group. Each 5 μg/m^3 increase in 7d $PM_{2.5}$ average was associated with 48% and 30% increase in AoPs scored manually (**B**, P=0.0009) or by MacLEAP (**D**, P=0.0024), respectively. D: day; M: month.

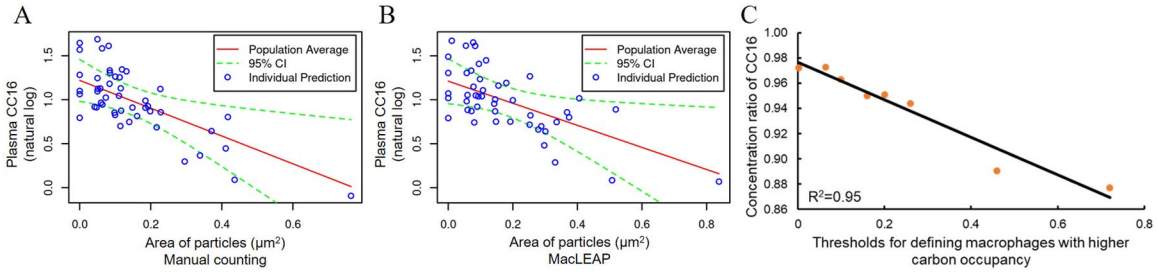

Figure 3. MaCL and lung injury biomarker CC16.

Each inter-quartile range increase in AoPs scored manually (**A**, 0.134 μm^2) or by MacLEAP (**B**, 0.195 μm^2) was associated with 19% and 22% reduction in plasma CC16 levels, respectively. Higher cut points for defining macrophages with large carbon occupancy resulted in stronger associations between %cells with higher carbon load and plasma CC16 levels (**C**).

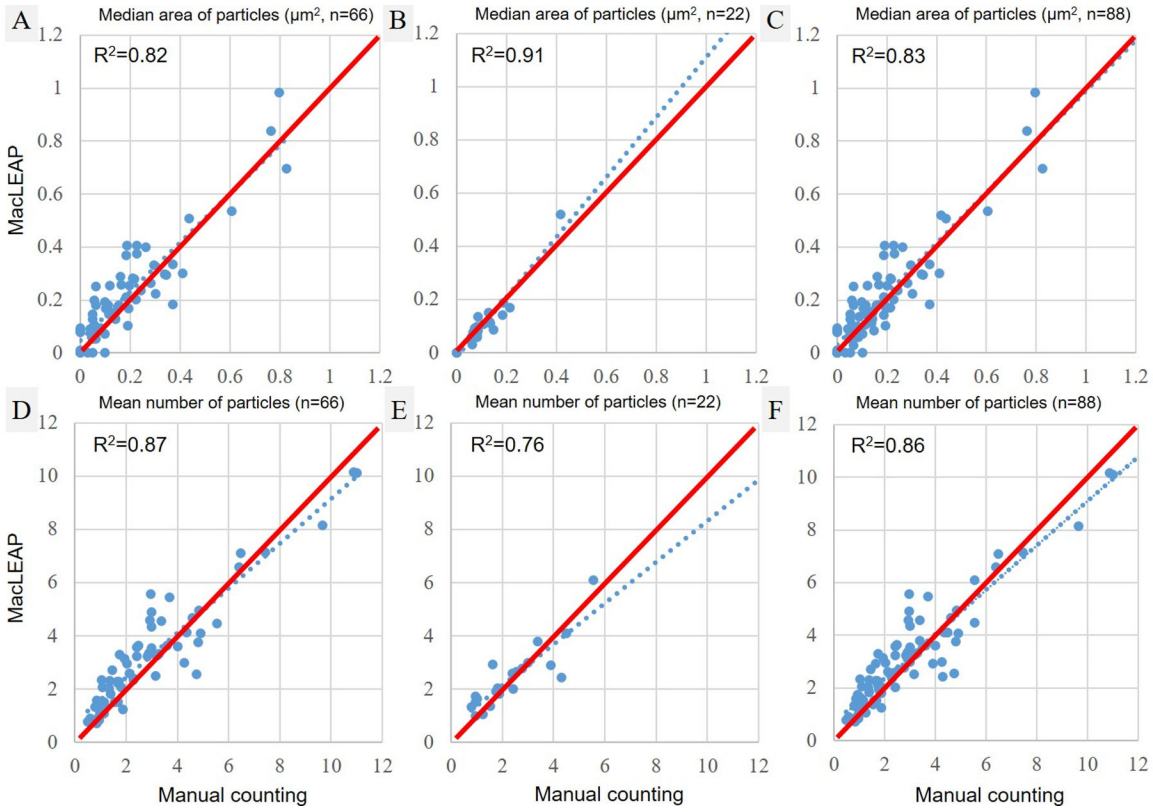

Figure 4. Correlations between manual and MacLEAP counting for NoPs and AoPs.

A and **D** summarized the correlations of AoPs and NoPs between manual and MacLEAP counting for 1456 sputum images from the 66 subjects. The optimized algorithm was then applied to the 333 sputum images from 22 subjects as an independent validation set (**B** and **E**). **C** and **F** demonstrated the correlations of AoPs and NoPs between manual and MacLEAP counting for all 1789 sputum images from all 88 subjects. Performance was evaluated using individual-based NoPs and AoPs. Red lines represented the 1:1 identity line. Blue dotted lines were the linear regression lines.

Table 1.

Demographics and macrophage carbon load measurements in high and low $PM_{2.5}$ exposure groups

Variable	High $PM_{2.5}$ exposure	Low $PM_{2.5}$ exposure	P
n	49	39	
Age (yr, mean ± SD)	55.1 ± 8.7	55.7 ± 7.5	0.74 *
Male (n, %)	4, 8.2	2, 5.1	0.57 †
Ethnicity			0.90 †
Non-Hispanic white (n, %)	41, 83.7	33, 84.6	
Others (n, %)	8, 16.3	6, 15.4	
Current smoker (n, %)	32, 65.3	26, 66.7	0.89 †
Packyears (mean ± SD)	41.2 ± 27.8	45.0 ± 25.1	
BMI (mean ± SD)	29.2 ± 5.9	28.1 ± 6.7	0.43 *
BMI>25 (n, %)	37, 75.5	24, 61.5	0.16 †
Ever woodsmoke exposure (n, %)	11, 22.5	10, 25.6	0.73 †
MaCL measurements			
Nops (mean ± SD)	2.02 ± 2.05	1.36 ± 1.06	0.035 ‡
AoPs (μm^2, median, Q1–Q3)	0.119, 0.058 – 0.217	0.088, 0.060 – 0.188	0.041 ‡
%CWP (%, mean ± SD)	69.9 ± 14.5	64.5 ± 16.4	0.11 ‡

NoP: Number of particle; AoP: Area of particle; %CWP: % cells with particles; SD: standard deviation; BMI: body mass index; Q: quartile

* Student t test

† Chi square test

‡ Poisson, Gamma, and Linear regressions were used for assessing associations between combustion particle exposure (high versus low) and NoP (count ratio=1.49, 95%CI=1.02–2.16), AoP (area ratio=1.44, 95%CI=1.01–2.05), and %CWP (estimate = 5.4% with 3.3% as standard error), respectively. NoPs and AoPs are median values for all macrophages scored per individual.

Table 2.

Associations between ambient $PM_{2.5}$ levels (per 5 μg/m^3 increase) and macrophage carbon load endpoints in all study subjects (n=88)

MaCL	$PM_{2.5}$-D0	$PM_{2.5}$-D7	$PM_{2.5}$-D14	$PM_{2.5}$-D28	$PM_{2.5}$-M2	$PM_{2.5}$-M3	$PM_{2.5}$-M6	$PM_{2.5}$-M12	$PM_{2.5}$-M18
NoP									
Ratio	1.23	1.48	1.46	1.41	1.59	1.96	3.68	4.53	6.87
95%CI	(1.11 – 1.37)	(1.28 – 1.71)	(1.21 – 1.77)	(1.11 – 1.77)	(1.15 – 2.21)	(1.23 – 3.11)	(1.56 – 8.72)	(1.70 – 12.11)	(1.99 – 23.77)
P	0.0001	<0.0001	<0.0001	0.0041	0.0056	0.0045	0.003	0.0026	0.0023
AoP									
Ratio	1.15	1.31	1.32	1.29	1.43	1.69	2.49	2.99	3.93
95%CI	(1.00 – 1.32)	(1.12 – 1.54)	(1.10 – 1.59)	(1.03 – 1.61)	(1.04 – 1.96)	(1.06 – 2.68)	(1.12 – 5.52)	(1.14 – 7.82)	(1.27 – 12.18)
P	0.055	0.0009	0.003	0.024	0.026	0.027	0.025	0.026	0.018
%CWP									
Est	1.05	3.70	4.05	3.90	6.35	9.82	15.55	17.87	21.34
SE	1.35	1.75	1.90	2.15	3.00	4.30	7.73	9.11	10.85
P	0.44	0.038	0.035	0.075	0.039	0.026	0.048	0.053	0.053

NoP: Number of particle; AoP: Area of particle; %CWP: % cells with particles; CI: confidence interval; D: day; M: month

Poisson, Gamma, and Linear regressions with adjustment for age, sex, quit time, packyears, overweight or obesity, ethnicity, and ever wood smoke exposure were used for assessing associations between $PM_{2.5}$ levels and NoPs, AoPs, and %CWP, respectively. NoPs and AoPs are median values for all macrophages scored per individual.

Table 3.

Associations between macrophage carbon load levels and plasma CC16 in a subset of study subjects (n=48)

MaCL variable	Unit	IQR	CC16 concentration ratio	P
NoP	count	1	0.84 (0.73 – 0.96)	0.011
AoP	μm^2	0.134	0.81 (0.69 – 0.95)	0.011
%CWP	%	20.4	0.89 (0.70 – 1.13)	0.33

NoP: Number of particle; AoP: Area of particle; %CWP: % cells with particles; IQR: inter-quartile range

Linear models were used to analyze the associations between MaCL endpoints and plasma CC16 (natural log transformed) with adjustment for age, sex, quit time, packyears, overweight or obesity, and ethnicity. CC16 concentration ratios were calculated as the exponential of estimates for each IQR change of the MaCL endpoints.

Table 4.

Associations between %cell with higher carbon load and plasma CC16 (n=48)

PCOC threshold	Mean ± SD	CC16 concentration ratio	P
Minimal (0.0025%)	67.5 ± 15.5	0.97 (0.92 – 1.03)	0.33
Median (0.064%)	46.4 ± 21.2	0.97 (0.93 – 1.02)	0.22
60th percentile (0.10%)	37.1 ± 20.3	0.96 (0.92 – 1.01)	0.10
70th percentile (0.16%)	27.6 ± 18.1	0.95 (0.90 – 1.00)	0.046
75th percentile (0.20%)	23.3 ± 17.0	0.95 (0.90 – 1.00)	0.064
80th percentile (0.26%)	18.6 ± 15.2	0.94 (0.89 – 1.00)	0.059
90th percentile (0.46%)	9.4 ± 9.5	0.89 (0.81 – 0.98)	0.018
95th percentile (0.72%)	5.0 ± 5.5	0.88 (0.74 – 1.04)	0.13

PCOC: percentage of cell area occupied by carbon particles

Linear models were used to analyze the associations between %cells with large P and plasma CC16 (natural log transformed) with adjustment for age, sex, quit time, packyears, overweight or obesity, and ethnicity. CC16 concentration ratios were calculated as the exponential of estimates for each 5% increase in % cells with large P to ensure comparability of magnitude of associations across models.