Integrated assessment of exposure to PM$_{2.5}$ in South India and its relation with cardiovascular risk: design of the CHAI observational cohort study

Cathryn Tonne$^{1,2}$, Maëlle Salmon$^{1,2}$, Margaux Sanchez$^{1,2,3}$, V Sreekanth$^4$, Santhi Bhogadi$^5$, Sankar Sambandam$^6$, Kalpana Balakrishnan$^6$, Sanjay Kinra$^7$, Julian D. Marshall$^4$

$^1$Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
$^2$Universitat Pompeu Fabra (UPF), Barcelona, Spain
$^3$CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
$^4$Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, United States
$^5$Public Health Foundation of India, New Delhi, India
$^6$Department of Environmental Health Engineering, Sri Ramachandra University, Chennai, India
$^7$Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

Corresponding author:

Cathryn Tonne

Cathryn.tonne@isglobal.org

+34 93 214 7361

Barcelona Institute for Global Health (ISGlobal)

Doctor Aiguader, 88

08003 Barcelona, Spain

Funding: This work was supported by the European Research Council under ERC Grant Agreement number 336167 for the CHAI Project.
Abstract

While there is convincing evidence that fine particulate matter causes cardiovascular mortality and morbidity, little of the evidence is based on populations outside of high income countries, leaving large uncertainties at high exposures. India is an attractive setting for investigating the cardiovascular risk of particles across a wide concentration range, including concentrations for which there is the largest uncertainty in the exposure-response relationship.

CHAI is a European Research Council funded project that investigates the relationship between particulate air pollution from outdoor and household sources with markers of atherosclerosis, an important cardiovascular pathology. The project aims to 1) characterize the exposure of a cohort of adults to particulate air pollution from household and outdoor sources 2) integrate information from GPS, wearable cameras, and continuous measurements of personal exposure to particles to understand where and through which activities people are most exposed and 3) quantify the association between particles and markers of atherosclerosis. CHAI has the potential to make important methodological contributions to modeling air pollution exposure integrating outdoor and household sources as well as in the application of wearable camera data in environmental exposure assessment.

**Keywords:** Air pollution; particulate matter; black carbon, biomass, atherosclerosis; cardiovascular disease; India; South Asia; mobility; time-activity; epidemiology

**Highlights:**

- Provides integrated assessment of exposure to particulate air pollution from both household and outdoor sources in rural South India
- Advances methods for application of wearable camera technology to environmental exposure assessment
- Investigates association between particulate air pollution and markers of atherosclerosis across a wide exposure range
Introduction

**Gaps in existing research** While there is convincing evidence that fine particulate matter (<2.5 micrometers in aerodynamic diameter, PM$_{2.5}$) causes cardiovascular mortality and morbidity (Brook et al., 2010; Newby et al., 2014), very little of this evidence is based on populations outside of high income countries. Findings from high income countries may have limited generalisability to low and middle income countries due to: 1) PM exposure levels are typically outside of the range observed in high income countries; 2) differences in particle composition and toxicity; and 3) differences in susceptibility to air pollution related to baseline health status and material deprivation.

Furthermore, there is large uncertainty in the relationship between exposure to combustion particles and cardiovascular risk for concentrations higher than those typical of outdoor concentrations in urban areas of high income countries and lower than active smoking (Smith and Peel, 2010; Pope et al., 2009). Exposures in this range are common in households which rely on solid fuel for cooking under poor combustion and ventilation conditions and outdoors in urban areas of low and middle income countries (Smith and Peel, 2010). The overall shape of the exposure-response relationship between long-term exposure to combustion particles and cardiovascular risk is likely to be non-linear and has important implications for which policies would deliver the largest health benefits (Smith and Peel, 2010).

India is an attractive setting for investigating the cardiovascular risk of particles across a wide exposure range, including the range for which there is the largest uncertainty in the exposure-response relationship. High particle concentrations have been measured outdoors in Indian cities, within vehicles, and in households using solid fuel for cooking (Pant et al., 2016). Similar to other settings outside high income countries, there is limited evidence regarding the health effects of such exposures in India, particularly for cardiovascular disease, although the estimated health burden is large. Air pollution is now estimated to have the largest burden of death and disability in India out of all modifiable risk factors (IHME).

**Cardiovascular disease burden in India** Cardiovascular diseases (CVD) are now the leading cause of mortality in India, largely due to ischemic heart disease (IHME; Prabhakaran et al., 2016). Compared to populations of European ancestry, CVD affects Indians at least one decade earlier: 52% of CVD deaths in India occur before the age of 70 compared to 23% for Western populations (Prabhakaran et al., 2016). CVD is the leading cause of death in all parts of India, including rural areas (Prabhakaran et al., 2016). Additionally, rural areas are affected by high and/or increasing prevalence of cardiovascular risk factors like hypertension and diabetes.
mellitus (Prabhakaran et al., 2016; Anchala et al., 2014; Kinra et al., 2010). However, large knowledge gaps remain in the descriptive epidemiology of CVD in India given the lack of current, nationally representative surveillance data on the prevalence or incidence of CVD and secular trends in CVD mortality (Prabhakaran et al., 2016).

**Objectives** The Cardiovascular Health effects of Air pollution in Telangana, India (CHAI) is a European Research Council funded project (Jan 2015 to Dec 2018) that investigates the relationship between particulate air pollution from outdoor and household sources with markers of atherosclerosis, an important cardiovascular pathology leading to permanent changes in the structure and function of arteries and ultimately acute events (Künzli et al., 2011). CHAI focuses on rural and peri-urban populations in Telangana, which serves as an excellent model of the population growth, urbanization, and increasing rates of CVD experienced throughout much of India. It remains unclear whether total exposure to particulate air pollution in such populations is decreasing as they shift from inefficient solid fuel to cleaner fuel types (e.g. liquefied petroleum gas and electricity), or increasing with increased reliance on electricity from diesel generators and increasing traffic emissions. Previous epidemiological studies have focused on either outdoor or household air pollution from solid fuel use. In many rural and peri-urban settings in India, individuals are exposed to pollution from multiple outdoor and indoor sources, and epidemiological studies in these populations require a more integrated approach to assess total exposure (Balakrishnan et al., 2014).

The CHAI project aims:

1. To characterize medium-term exposure of participants of the Andhra Pradesh Children and Parents Study (APCAPS) (Kinra et al., 2014) to particles using an integrated approach based on monitoring of outdoor concentrations, personal monitoring, and previously collected questionnaire data.

2. To identify specific locations and activities associated with high particle exposure using methods for the objective measurement of location and time-activity patterns based on global positioning system (GPS) devices and wearable cameras linked to continuous personal monitoring of particle pollution.

3. To quantify the association between exposure to particles and biomarkers of cardiovascular risk (intima media thickness of the carotid artery and arterial stiffness) in APCAPS.
The published literature focusing on carotid intima media thickness (CIMT) in relation to quantitative exposure to PM$_{2.5}$ is derived from high income countries, overwhelmingly in North America and Europe (Adar et al., 2013; Aguilera et al., 2016; Bauer et al., 2010; Breton et al., 2016a, 2016b, 2012; Diez Roux et al., 2008; Gan et al., 2014; Jones et al., 2015; Kaufman et al., 2017; Kunzli et al., 2005; Künzli et al., 2010; Lenters et al., 2010; Perez et al., 2015; Su, 2015; Sun et al., 2013). This literature covers a small fraction of the global exposure range, leaving large uncertainties about the shape of the association at higher exposures. Figure 1 compares PM$_{2.5}$ exposures represented in epidemiological studies of CIMT (Adar et al., 2013; Aguilera et al., 2016; Bauer et al., 2010; Breton et al., 2016a, 2012; Diez Roux et al., 2008; Gan et al., 2014; Jones et al., 2015; Kunzli et al., 2005; Künzli et al., 2010; Lenters et al., 2010; Perez et al., 2015; Su, 2015; Sun et al., 2013) with average levels in the World Health Organization database compiled from official ground based measurements, which under-represents countries lacking official air pollution monitoring (WHO, 2016). The annual average ambient PM$_{2.5}$ concentration in our study area is approximately 30 µg/m$^3$, and the range of personal exposures is likely to include considerably higher concentrations. Previously reported 24-hr personal PM$_{2.5}$ exposures in India range from 70 to 1500 µg/m$^3$ (Andresen et al., 2005; Balakrishnan et al., 2002). Therefore, CHAI has the potential to shed light on the shape of the exposure-response function of PM$_{2.5}$ with CIMT at higher exposures not previously included in the literature.

Methods

Study area and population CHAI builds on the APCAPS prospective cohort investigating risk factors for cardio-metabolic disease over the life course (Kinra et al., 2014). APCAPS includes approximately 7000 participants (4000 with the health outcomes of interest in CHAI), residing in 29 villages southeast of Hyderabad (Figure 2). Approximately half of the APCAPS participants were originally recruited as part of the Hyderabad Nutrition Trial in 1987-1990 and were followed-up in young adulthood between 2003 and 2010. The other half of the cohort is comprised of family members of the nutrition trial participants recruited as part of the most recent APCAPS (third) follow up conducted between 2010 and 2012.

CHAI participants were randomly sampled, stratified by village (weighted by proportion of APCAPS participants in each village), from participants in APCAPS third follow up, who were still resident of APCAPS villages and 18 years or older. Target recruitment was 400 participants for 24-hour gravimetric personal exposure measurements. A subset of 40 individuals was
included in a more detailed panel study of predictors of PM$_{2.5}$ exposure, to better understand where and through which activities participants receive the most exposure (Figure 3).

Ethics approval was granted by the Parc de Salut Mar, Public Health Foundation of India, National Institute of Nutrition, Sri Ramachandra University, and the European Research Council. All participants provided informed consent, and information and consent documents were in Telugu, the local language.

Data collection

Area-level exposure data

**Ambient PM$_{2.5}$** To characterize ambient concentrations and their temporal variability, we established three stationary monitoring stations within the study area (Figure 2). Two stationary monitors use a US EPA federal equivalent beta attenuation method (MetOne, eBAM) to measure PM$_{2.5}$, while all three stationary monitoring sites include a TSI DustTrak to provide PM$_{2.5}$ data with higher temporal resolution alongside relative humidity, temperature, and wind speed and direction. These data will fill an important gap in ambient monitoring of PM$_{2.5}$ in rural South India, shed light on the impact of local emissions on ambient air, and be used to adjust for the influence of meteorology in the variation in personal exposures.

**Saturation monitoring** To characterize spatial variation in background air pollution throughout the study area, we conducted 24hr integrated gravimetric PM$_{2.5}$ measurements at 23 locations for 10-11 days in two seasons: post-monsoon and summer. Monitoring was conducted every other day, spanning 20-22 days in each season. Sites were selected to cover the range of variables expected to be important predictors of spatial variation in PM$_{2.5}$ in the area (e.g. distance to road, village-level solid fuel use). PM$_{2.5}$ mass concentration was measured using a pump which drew air through a cyclone attached to a cassette containing a 37mm filter on which PM$_{2.5}$ was deposited. Filters were weighed pre and post exposure in the field. We also measured optical attenuation (880nm) by the mass collected on the filter using a Magee OT21 Sootscan transmissometer, which can be converted into black carbon concentration.

Individual-level exposure data

**Gravimetric personal monitoring** Two repeated measurements per participant of integrated 24-hr gravimetric PM$_{2.5}$ were collected on all CHAI participants (n=400) with one measurement in each period: May to July 2015 and December to March 2016 (Figure 4). We used the same measurement approach as in the saturation monitoring, with the pump worn in a small
backpack by the participant and the inlet placed with a clip near the breathing zone. This approach is considered the gold standard for personal exposure measurements for PM mass concentration. As with the saturation monitoring filters, we used PM$_{2.5}$ optical attenuation to estimate exposure to black carbon. A post-monitoring questionnaire was administered at the end of each personal monitoring period to capture self-reported time activity, and cooking activities including stove and fuel use.

Panel study with repeated measurements In the subset of 40 individuals included in the panel study, we measured continuous personal exposure to PM$_{2.5}$ using the RTI MicroPEM, a miniaturized nephelometer with gravimetric backup that samples every 10 seconds. Measurements were repeated on the same participant up to 6 times to cover all seasons over the course of one year (Figure 4). Two out of the six measurements were designed to coincide with the gravimetric measurements, allowing for comparisons between co-located MicroPEM and gravimetric measurements. The co-located personal monitoring equipment is shown in Figure 5. Participants wore a small backpack with a GPS device to track their locations as well as an accelerometer to detect compliance with the monitoring instructions. Participants also wore a wearable camera that captures time-stamped images of the environment in front of him/her approximately every 30 seconds without user intervention. Compared to time-activity diaries, data derived from wearable cameras is less burdensome, has higher time resolution, and is more objective because it does not rely on subject recall which can be systematically biased (Kelly et al., 2011). These data are coded manually by trained research staff according to a predefined ontology with concepts related to activities likely to influence air pollution exposure, for example, cooking, travel, presence in industrial setting. Approximately 81% of 290,000 photos collected had a sufficiently clear image (e.g. sufficient light, not blocked by clothing) to be coded. Figure 6 presents an example of the continuous PM$_{2.5}$ data overlaid with activities derived from the wearable camera, and location data derived from GPS. Again, post-monitoring questionnaires were administered at the completion of each monitoring period, including self-reported time activity for comparison with the wearable camera.

Baseline questionnaire A questionnaire was administered by the field team at time of enrollment (May 2015). The questionnaire collects extensive data on fuel use, housing characteristics, and other factors likely to influence air pollution exposure. Many of the questions in the CHAI baseline questionnaire are the same as in the APCAPS third follow up, allowing us to identify changes in important predictors of exposure over time.

Health data
Health outcome data were collected through the APCAPS third follow up, including data on CIMT and arterial stiffness: CIMT measured by Ethiroli Tiny-16a, Surabi Biomedical Instrumentation; augmentation index by Sphygmocor, AtCor Medical; and pulse wave velocity by Vicorder, Skidmore Medical. We focus our investigation of cardiovascular risk on atherosclerosis because of the growing literature from human and animal studies of a link with particle exposure (Araujo et al., 2008; Brook and Rajagopalan, 2010; Kaufman et al., 2017; Liu et al., 2015; Sun et al., 2005; Tonne et al., 2012). Specifically, we will use CIMT and arterial stiffness as biomarkers of structural and functional arterial vulnerability. CIMT has been identified as a useful outcome for studies of air pollution and atherosclerosis because it is a well-validated measure and represents the overall atherogenic burden (Künzli et al., 2011). Meta-analyses indicate an overall positive association between particle exposure and CIMT, largely from cross-sectional studies (Liu et al., 2015; Provost et al., 2015). However, a recent longitudinal study in the US did not observe a relationship between PM$_{2.5}$ and progression of CIMT, although there was a positive association with progression of coronary artery calcium (Kaufman et al., 2017). Arterial stiffness, an alternative measure of atherosclerosis based on the functional performance of vessels, has also been rated relatively highly as a useful marker to investigate the atherogenic effects of particles (Künzli et al., 2011).

Data analysis

Exposure modeling Data from the saturation monitoring will be used in a land use regression model, an approach commonly used in air pollution epidemiology to predict outdoor concentrations at each participant’s residence. For example, the model could include functions of distance from a major road, distance from the village centre, population density, and distance from Hyderabad, among other predictors. Similarly, data from the gravimetric personal monitoring (n=400, 2 repeated measures per person) will be used to fit a model to predict personal exposure using individual (e.g. age, sex, occupation), household (e.g. primary cooking fuel) and area (e.g. village-level prevalence of solid fuel use) predictors. We will test whether the variables selected as predictors in the land use regression model are also predictive of measured personal exposure. This model will allow for prediction of personal exposure on APCAPS participants not included in the personal monitoring measurements. These two measures of exposure for the entire cohort can be used in comparative epidemiology analysis (Figure 7).

Panel study Repeated measures of GPS, activities derived from the wearable camera, and continuous PM$_{2.5}$ data will be analyzed using mixed models to identify specific locations and
activities related to high exposure. Independent of their relationship with air pollution exposures, the GPS and wearable camera data provide useful insights into time activity patterns and daily mobility, for which there is little current information available for rural or peri-urban Indian populations.

**Epidemiological analyses** Regression models will be used to estimate the association between particle exposure and atherosclerosis, accounting for clustering of individuals within villages and households. Analyses will be based on predicted PM$_{2.5}$ and black carbon at residence from the land use regression models as well as personal. We estimate 97% power to detect a 1% increase in CIMT for a 1 µg/m$^3$ increase in PM$_{2.5}$, assuming n=4,000, alpha=5%, and the standard deviation in PM$_{2.5}$ exposure is 3 µg/m$^3$. Detailed covariate data are available from APCAPS to adjust for potential confounders.

**Discussion**

The strengths of CHAI include a large range in exposure to particulate air pollution, including high exposures. The study uses a large number of personal monitoring measurements to estimate exposure, with repeated measurements to characterize within person variability, and detailed panel study data to understand exposure misclassification in the epidemiological analyses. We apply predictive exposure modeling approaches that have not been previously applied in a rural setting influenced by residential biomass combustion. An additional strength is combining these exposure data with the rich health outcome and covariate data from APCAPS, a relatively large cohort. The primary limitation of CHAI is that health outcomes of interest have been collected once, and the epidemiological study is cross-sectional. Future follow-up of the cohort will enable longitudinal analyses in relation to air pollution exposure. Another limitation is the gap between the air pollution measurements in CHAI (2015-2016) and the questionnaire and health outcome data collection in the third follow up of APCAPS (2010-2012). A subset of the detailed questionnaire data collected in CHAI is available for the full APCAPS cohort from the third follow-up. These data, which are contemporaneous with the health outcome data, will be used to predict exposure. We will quantify the extent to which the relationship between individual characteristics (e.g. demographics, occupation) and PM$_{2.5}$ exposure has varied over a five-year period, but expect these changes to be relatively small.

**Generalisability** We expect the epidemiological findings to be generalisable to populations in South Asia and elsewhere at a similar stage of economic development and energy transition as the APCAPS villages. The findings from methodological developments in exposure assessment
using wearable cameras regarding feasibility and extent of value added above standard exposure assessment approaches will be more broadly applicable.

Ethical issues There are important ethical considerations in collecting, storing, analyzing, and disseminating results from wearable camera technology (Mok et al., 2015). Most ethical concerns relate to agency, accountability, third party trust, and the delegation of responsibility from researchers to participants (Shipp et al., 2014). CHAI relies on the functional guidelines for ethically acceptable use of first person, point of view, passive wearable cameras in health behavior research provided by Kelly et al. (Kelly et al., 2013). Key precautions include allowing participants to view and delete images before passing on to researchers and maintaining strict separation between data collection, curation, and analysis. For example, in CHAI, only specially trained research staff view the images and code activities, and only activity labels are used in data analysis. Image data at no point leave a tightly controlled “research environment”, meaning that participants are not allowed to save or share their own images. Only aggregated results based on activity codes will be used in research dissemination.

Dissemination Information on exposure levels will be fed back to participants in an easy to understand format. Information at more aggregated levels about locations and activities associated with high exposures will be fed back through community meetings. Meta data and a data sharing policy will be published on the project website (chaiproject.org). Outdoor air pollution data will be made available for sharing with other researchers. Other data, in fully anonymized form, will be made available for sharing with researchers on a restricted basis.
Figure headings

Figure 1. Comparison of the PM$_{2.5}$ average concentrations reported in published literature on the association between PM$_{2.5}$ and carotid intima media thickness (top), and reported in the WHO database (bottom) according to country income level from the World Bank.

[Systematic review from 1950 to Jan 2017 resulting in 19 mean PM$_{2.5}$ estimates from 14 publications. Search key words: “Air pollutants”, “Particulate matter”, “PM2.5”, “PM10”, “Carotid intima-media thickness”, “Carotid IMT”, “CIMT”, “subclinical atherosclerosis”]

Figure 2. Map of study area. Dots indicate study village locations, stars indicate fixed monitor locations.

Figure 3. Schematic indicating sample size of nested subsets of study population.

Figure 4. Monitoring timeline for gravimetric personal monitoring on all CHAI participants and the subset included in the panel study.

Figure 5. Personal monitoring equipment used in panel study including (from left to right) MicroPEM, wearable camera, and gravimetric sampler.

Figure 6. Example of continuous PM$_{2.5}$ data measured with the MicroPEM for one individual with simultaneously measured activities derived from wearable camera data (A). Distance from home derived from GPS data (B).

Figure 7. Interlinkages between exposure data collected or modeled through CHAI and APCAPS health outcome and covariate data used in epidemiological analyses.
References


Figure 1. Comparison of the PM$_{2.5}$ average concentrations reported in published literature on the association between PM$_{2.5}$ and carotid intima media thickness (top), and reported in the WHO database (bottom) according to country income level from the World Bank.

[Systematic review from 1950 to Jan 2017 resulting in 19 mean PM$_{2.5}$ estimates from 14 publications. Search key words: “Air pollutants”, “Particulate matter”, “PM2.5”, “PM10”, “Carotid intima-media thickness”, “Carotid IMT”, “CIMT”, “subclinical atherosclerosis”]
Figure 2. Map of study area. Dots indicate study village locations, stars indicate fixed monitor locations.

Figure 3. Schematic indicating sample size of nested subsets of study population.
Figure 4. Monitoring timeline for gravimetric personal monitoring on all CHAI participants and the subset included in the panel study.

Figure 5. Personal monitoring equipment used in panel study including (from left to right) MicroPEM, wearable camera, and gravimetric sampler.
**Figure 6.** Example of continuous PM$_{2.5}$ data measured with the MicroPEM for one individual with simultaneously measured activities derived from wearable camera data (A). Distance from home derived from GPS data (B).

**Figure 7.** Interlinkages between exposure data collected or modeled through CHAI and APCAPS health outcome and covariate data used in epidemiological analyses.