A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP5445.

Land-Use Change and Cardiometabolic Risk Factors in an Urbanizing Area of South India: A Population-Based Cohort Study

Carles Milà,[1,2,3] Otavio Ranzani,[1,2,3] Margaux Sanchez,[4] Albert Ambrós,[1,2,3] Santhi Bhogadi,[5] Sanjay Kinra,[6] Manolis Kogevinas,[1,2,3] Payam Dadvand,[1,2,3] and Cathryn Tonne[1,2,3]

[1]Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
[2]Universitat Pompeu Fabra, Barcelona, Spain
[3]CIBER Epidemiología y Salud Pública, Madrid, Spain
[4]Vieillissement et Maladies chroniques, Institut national de la santé et de la recherche médicale, Paris, France
[5]Public Health Foundation of India, New Delhi, India
[6]London School of Hygiene and Tropical Medicine, London, UK

Background: Land-use changes in city fringes due to urbanization can lead to a reduction of greenspace that may reduce its associated health benefits.

Objectives: We evaluated the association between changes in residential surrounding built-up land use and cardiometabolic risk factors in an urbanizing peri-urban area of south India and explored the mediating roles of air pollution, physical activity, and stress in these associations.

Methods: We analyzed data on 6,039 adults from the third follow-up of the Andhra Pradesh Children and Parent Study (APCAPS) cohort (2010–2012). We generated trajectories of change in residential surrounding built-up land use (buffer areas) from 1995–2009 (stable, slow increase, fast increase) using remote sensing data and image classification methods. We estimated associations between built-up land use trajectories and natural log-transformed blood pressure, waist circumference, triglycerides, fasting glucose, and non–high-density lipoprotein (non-HDL) cholesterol using linear mixed models. We accounted for multiple mediators and the multilevel structure of the data in mediation analyses.

Results: We observed positive associations between a fast increase in built-up land use within 300 m of the home and all cardiometabolic risk factors. Compared with participants with stable trajectories, those with the largest increase in built-up land use had 1.5% (95% CI: 0.1, 2.9) higher systolic blood pressure, 2.4% (95% CI: 0.6, 4.3) higher diastolic blood pressure, 2.1% (95% CI: 0.5, 3.8) higher waist circumference, and 1.6% (95% CI: −0.6, 3.8) higher fasting glucose in fully adjusted models. Associations were positive, but not statistically significant, for triglycerides, fasting glucose, and non-HDL cholesterol. Physical activity and ambient particulate matter $\leq$2.5 μm in aerodynamic diameter ($PM_{2.5}$) partially mediated the estimated associations. Associations between fast build-up and all cardiometabolic risk factors except non-HDL cholesterol were stronger in women than men.

Discussion: Increases in built-up land use surrounding residences were consistently associated with higher levels of cardiometabolic risk factors. Our findings support the need for better integration of health considerations in urban planning in rapidly urbanizing settings. https://doi.org/10.1289/EHP5445

Introduction

India is undergoing an epidemiologic transition: The contribution of noncommunicable diseases to the total burden of disease in India (measured in disability-adjusted life years) increased from 31% in 1990 to 55% in 2016, led by cardiovascular diseases (14% of total burden in 2016) (India State-Level Disease Burden Initiative Collaborators 2017). Despite wide geographic heterogeneity within the country (Geldsetzer et al. 2018), prevalence of cardiovascular diseases is generally high in both urban and rural areas and is expected to increase in the future (Prabhakaran et al. 2016). The main risk factors for cardiovascular diseases in India in 2016 were diet and high levels of systolic blood pressure (SBP), total cholesterol, plasma glucose, and body mass index as well as high levels of ambient air pollution (India State-Level Disease Burden Initiative CVD Collaborators 2018).

The proportion of the Indian population living in urban areas is projected to increase from 33% in 2015 to 53% by 2050 (UN DESA 2018). Previous remote sensing studies in India have identified a decrease in natural and agricultural areas with increasing urbanization in large cities and medium-sized settlements (Gibson et al. 2015; Roy et al. 2015) that may lead to poorer air quality (Larkin et al. 2016). Remote sensing data are able to capture land-use changes over time and space (Rogan and Chen 2004), thus fostering opportunities for research on the health impacts of land-use change around the globe.

Greenspace has been linked to numerous health benefits. An accumulating body of evidence has associated this exposure with a lower risk of cardiovascular mortality (Gascon et al. 2016; Twohig-Bennett and Jones 2018), cardiometabolic risk factors such as blood pressure (Twohig-Bennett and Jones 2018), and possibly fasting glucose and obesity (James et al. 2015; Twohig-Bennett and Jones 2018). Possible mechanisms linking greenspace and cardiovascular disease include mitigation of environmental hazards (e.g., air pollution), stress reduction, and promotion of physical activity (Markevych et al. 2017; Twohig-Bennett and Jones 2018).

Despite the growing literature on the health effects of greenspace exposure, or its converse—built-up areas consisting of non-vegetated, human-constructed elements—evidence from low- and middle-income countries (LMICs) remains limited (Corlin et al. 2018; Lane et al. 2017). Generalizability of findings from high-income countries (HICs) to LMICs is challenging due to different definitions of greenspace (Taylor and Hochuli 2017). Indeed, previous studies from HICs focusing on urban greenspace (e.g., parks, urban forests, gardens) may have limited applicability to settings where greenspace largely represents farmland and bare, open areas (Taylor and Hochuli 2017).

We investigated the association between changes in residential surrounding built-up land use and cardiometabolic risk factors in an urbanizing area in the south of Hyderabad, India. Our objectives were to *a*) estimate associations between change (1995–2009) in residential surrounding built-up land use and cardiometabolic risk factors (2010–2012) and *b*) explore the mediating roles of air pollution, physical activity, and stress.

Address correspondence to Cathryn Tonne, ISGlobal, Universitat Pompeu Fabra, CIBER Epidemiología y Salud Pública Doctor Aiguader 88, 08003 Barcelona, Spain. Telephone: 34 93 227 1806. Email: cathryn.tonne@isglobal.org

Supplemental Material is available online (https://doi.org/10.1289/EHP5445).

The authors declare they have no actual or potential competing financial interests.

Received 12 April 2019; Revised 18 February 2020; Accepted 18 March 2020; Published 3 April 2020.

Methods

Study Population and Area

We used data from the third follow-up of the Andhra Pradesh Children and Parent Study (APCAPS) (Kinra et al. 2014). APCAPS was approved by the London School of Hygiene & Tropical Medicine (London, UK) and the National Institute of Nutrition (Hyderabad, India). Signed consent forms were obtained from all participants. APCAPS builds on the Hyderabad Nutrition Trial (1987–1990), which included all newborns (index children) in 28 villages in a peri-urban area south of Hyderabad, India (Figure 1A). The third follow-up of APCAPS (2010–2012) recruited adult index children born during the original trial ($n = 1{,}360$) and their parents and siblings ($n = 5{,}584$), yielding a total sample size of 6,944 participants (see Figure S1). Nearly all (95%) index children lived in the same village in APCAPS follow-up 3 (2010–2012) as during the original trial (1987–1990). Participants were representative of the total village population in terms of sex, age, and education level (Curto et al. 2019). The coordinates of the front door of the participants' residences were recorded with a Global Positioning System (GPS) device (accuracy of ~ 4 m) as part of a built environment survey in 2012–2013; village boundaries were traced from aerial imagery. The study area covered approximately 700 km^2, encompassing part of the Hyderabad ring road, dispersed villages, and small towns, farmland, and sparsely vegetated and barren land (Figure 1A). Aerial images of one of the study area villages in 2003 and 2011 are shown in Figure S2A,B.

Data Collection and Covariates

Participants visited clinics established in the study villages as part of the APCAPS (2010–2012), where they responded to a questionnaire and underwent a standardized physical examination. The questionnaire included questions about demographics, household characteristics, health behaviors (physical activity, diet, smoking), and medical history. Physical activity over the past week was evaluated by an interviewer-administered questionnaire [APCAPS–Physical Activity Questionnaire (APCAPS-PAQ)]. The APCAPS-PAQ evaluates physical activity in six domains (work, travel to and from work, leisure, household, sedentary, and sleep), and within each domain the average amount of time spent on specific activities and the frequency of these activities were collected. We calculated the total physical activity

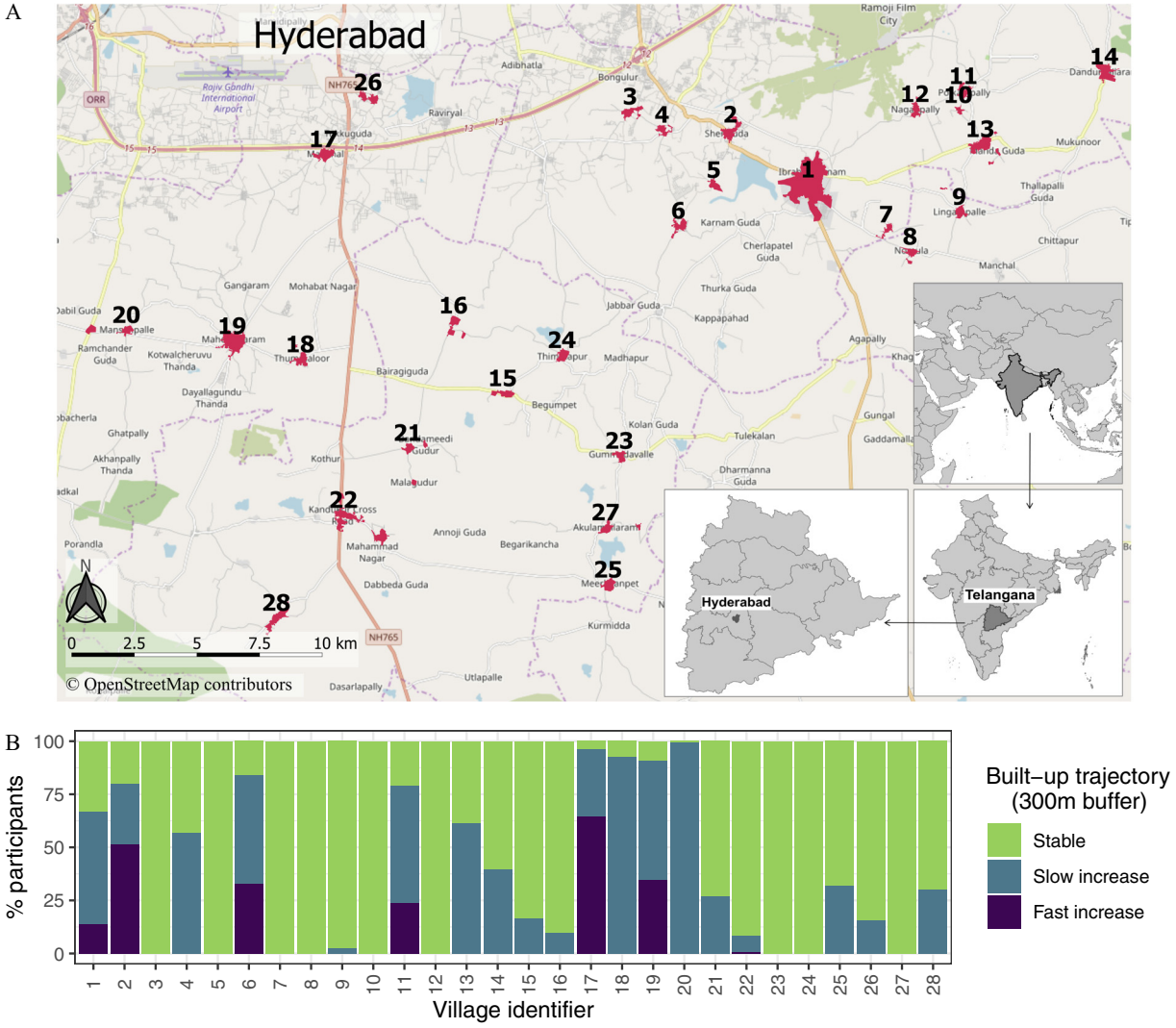

Figure 1. (A) Study area map and (B) distribution of trajectories of residential surrounding built-up land use by village (300-m buffer). Red polygons (in A) represent APCAPS villages. OpenStreetMap was used as background map (in A) under the Open Database License (https://www.openstreetmap.org/copyright). Note: APCAPS, Andhra Pradesh Children and Parent Study; % participants, percentage participants.

for each participant by summing daily metabolic equivalent unit values (METs). METs were attributed to each activity using the approach reported in the APCAPS-PAQ validation, which used the standard guidelines for MET assignment and a correction factor for occupational activities considered "more strenuous than walking" (Matsuzaki et al. 2016). Dietary intake over the past year was evaluated through a semi-quantitative food frequency questionnaire (FFQ). We obtained average daily consumption of sugar and sweets, alcohol, fruits, and carbohydrates from the FFQ. We used an algorithm developed during the validation of the FFQ in the APCAPS area to link information from the FFQ to nutrient databases, retrieving the average nutrient and food group consumption per day (Bowen et al. 2012). Smoking status was evaluated by the following self-reported question: "Have you ever used tobacco on a regular basis (at least weekly)?" Participants were asked about smoked, chewed, and snuffed tobacco. There were three possible answers used to classify the participants smoking status: never, former (stopped >6 months), and current (regular use in the last 6 months). We classified those on current use as active smokers, and we created two dummy indicators: active tobacco smokers and other active tobacco users (chew/snuff). From the participants' medical histories, we used information about current medication use gathered from the following self-reported questions with yes/no answer options: "Have you been diagnosed with any of the following conditions? 1) High blood pressure; 2) Diabetes (High blood sugar)" and "Are you on regular medication for your high blood pressure?; Are you on regular tablets for your diabetes?" A standard of living index (SLI) was calculated from self-reported household assets (Kinra et al. 2014). We derived a stress indicator using nine self-reported questions about physical, personal, and financial worries in the previous 4 weeks using psychosocial stressors items from the Brief Patient Health Questionnaire (Spitzer et al. 2000). Each question had three possible answers: not bothered (1 point), bothered a little (2 points), and bothered a lot (3 points). We summed the points of the nine questions to derive a score (range: 9–27 points), with higher scores indicating more sources and/or intensity of stress.

Outcomes

We collected data on six components of metabolic syndrome (Alberti et al. 2009): SBP and diastolic blood pressure (DBP), waist circumference, triglycerides, fasting glucose, and non–high-density lipoprotein (non-HDL) cholesterol. Participants were asked to fast overnight. SBP and DBP were defined as the average of three consecutive readings using an oscillometric device (HEM 7300; Omron); room temperature was recorded for purposes of adjusting in analyses of blood pressure outcomes. Waist circumference was calculated as the average of two consecutive measures of the natural waist (metallic tape; ADE). We performed assays for triglycerides, HDL, and total cholesterol (C311; Roche), and glucose [using the glucose oxidase/peroxidase-4-aminophenazone-phenol (GOD-PAP) enzymatic method]. Non-HDL was derived as the difference between total and HDL cholesterol. We also derived a composite cardiometabolic outcome using the INTERHEART modifiable risk score (IHMRS, version accounting for cholesterol), a globally validated risk score that includes age, sex, non-HDL and HDL cholesterol, tobacco smoking, secondhand smoking, diabetes, and high blood pressure (McGorrian et al. 2011). Missing individual components of the score were set to 0.

Ambient Fine Particulate Matter

We collected ambient fine particulate matter [PM $\leq$2.5 μm in aerodynamic diameter ($PM_{2.5}$)] concentrations (during 2015–2016) at 23 sites in the APCAPS study area within the Cardiovascular Health effects of Air pollution in Telangana, India (CHAI) project (Tonne et al. 2017) and developed local land-use regression (LUR) models (Sanchez et al. 2018) using geographical predictors collected during a built environment survey (in 2012–2013). The $PM_{2.5}$ model had an adjusted R^2 of 58% and included regional predictors (Sanchez et al. 2018). We used the model to predict $PM_{2.5}$ annual exposure at the residence for all APCAPS participants.

Exposure Assessment

We characterized changes of land use over our study area for the period 1995–2009 from remote sensing data and derived three trajectories of residential surrounding built-up land use (stable, slow increase, and fast increase) for each participant in 300- and 500-m buffers. Buffers larger than 500 m generally included the entire village, limiting the within-village exposure variability. Previous analyses of GPS data in a subsample of the APCAPS participants found that 80% of the daytime (72% men, 87% women) was spent within 400 m of the residence, whereas daytime spent within 100 m of the residence was <67% for men (Sanchez et al. 2017). The workflow to obtain exposures involved the following steps:

1) We listed all available Landsat 5 and 7 images of the study area (spatial resolution: 30 m) available between 1995 and 2009 and filtered out those with clouds over the study area. We downloaded the selected surface reflectance products (Masek et al. 2006) from the U.S. Geological Survey EarthExplorer platform (https://earthexplorer.usgs.gov/).
2) We defined built-up training areas from the urban core of Hyderabad and village centers in the study area already built up in 1995. Training areas for greenspace were traced from 2009 high-resolution imagery available in Google Earth and included bare areas, crops, vegetation, and water bodies. We used the same training areas to classify all images under the hypothesis of monotonic greenspace to built-up transition.
3) We trained random forest (RF) models (Hastie et al. 2001) to classify each pixel as either built-up or greenspace for each image separately. We used the blue, green, red, near infrared (NIR) and the two shortwave infrared (SWIR1 and SWIR2) spectral bands, as well as the normalized difference vegetation index [$NDVI = (NIR - red)/(NIR + red)$] and the normalized difference built index [$NDBI = (SWIR1 - NIR)/(SWIR1 + NIR)$], as input data.
4) We evaluated the predictive ability of the RF models by applying 10-fold spatial cross-validation (Brenning 2012). Accuracy was calculated as the percentage of pixels correctly classified. Models with accuracy below 85% were discarded; if more than one model for the year same remained, the one with the highest accuracy was retained. Table S1 shows the details of the selected images; most of which corresponded to the post-monsoon season when contrast between built-up and greenspace was stronger.
5) We generated binary maps (1 = built-up, 0 greenspace) for the selected models (see Figure S3) and extracted an average built-up proportion in a buffer around the home (300-m and 500-m buffers).
6) Raw built-up trajectories were generated by grouping the generated exposures by household at different points in time (see Figure S4A). We individually smoothed them using local linear regression models (Hastie et al. 2001) (see Figure S4B) to reduce noise. Because our focus was on change over time in land use, we subtracted the mean of all trajectories in order to center them at 0 (see Figure S4C).
7) We grouped the smoothed trajectories by using a longitudinal implementation of the *k*-means algorithm (Genolini and Falissard 2011). We used Euclidian distances and tried different starting points to avoid local maximums. The number

of clusters was chosen according to the Caliński and Harabasz (1974) criterion, which suggested two or three groups. We chose three categories because we were interested in evaluating the impact of variation in the rate of change in the built environment (see Figure S4D). The resulting groups were labeled as stable, slow increase, and fast increase in built-up land use.

Alternative Exposure Indices

We generated three alternative built-up indices (image dated 27 October 2009, buffer 300 m) for sensitivity analysis to explore cross-sectional associations rather than a dynamic exposure, capturing change over time. Alternative indices included additive inverse NDVI, NDBI (Estoque and Murayama 2015), and urban index (UI) (Estoque and Murayama 2015). We also considered the proportion of built-up land use within 300 m extracted from the RF-classified image (dated 27 October 2009) as an additional cross-sectional exposure.

Analyses

All cardiometabolic risk factors were right-skewed; we applied a natural log-transformation to better meet the assumption of normally distributed residuals from regression models. IHMRSs were kept in the original scale. We log-transformed diet covariates and categorized SLI and alcohol intake variables into tertiles to account for nonlinear effects identified in exploratory analyses.

We restricted analyses to nonpregnant participants >18 years of age with nonmissing exposure, sex, and age [6,039/6,944 (87%)]. We excluded participants with missing or invalid measurements for each outcome (i.e., the number of participants included in each analysis varied by outcome; Table 1). For example, when analyzing the outcome SBP and DBP, we excluded participants who were taking hypertensive medication (3.1%), were missing information on room temperature (0.1%), or had measurements taken on the left arm (0.3%); when analyzing the outcome fasting glucose, we excluded participants who were taking diabetes medication (1.5%), had not fasted for 8 h (5.9%), or were missing information for the fasting period (1.8%).

We imputed missing data in covariates using the method of chained equations ($n = 20$) (van Buuren and Groothuis-Oudshoorn 2011) using the same analysis data set, including village identifiers, as input for the imputation. We assumed certain data to be missing at random (MAR) because missing entries were mostly correlated with the village of the participant. We imputed binary covariates using logistic regression; occupation was imputed using a

Table 1. Characteristics of the study population by residential surrounding built-up land use (300-m buffer) trajectory.

Variable	All ($n = 6,039$)	Stable ($n = 3,615$)	Slow increase ($n = 1,888$)	Fast increase ($n = 536$)	Missing [n (%)]
Sex {male [n (%)]}	3,232 (53.5)	1,936 (53.6)	1,036 (54.9)	260 (48.5)	0 (0)
Age [AM (SD)]	36.2 (13.8)	36.4 (13.9)	36.1 (13.7)	35.1 (13.1)	0 (0)
Occupation [n (%)]					2 (0)
Unskilled manual	2,734 (45.3)	1,663 (46)	889 (47.1)	182 (34)	—
Skilled manual	1,371 (22.7)	854 (23.6)	393 (20.8)	124 (23.1)	—
Nonmanual	347 (5.7)	207 (5.7)	94 (5)	46 (8.6)	—
Unemployed	1,585 (26.3)	891 (24.6)	510 (27)	184 (34.3)	—
Education [n (%)]					2 (0)
Illiterate	2,959 (49)	1,770 (49)	938 (49.7)	251 (46.8)	—
Primary school	793 (13.1)	459 (12.7)	265 (14.1)	69 (12.9)	—
Secondary school	1,813 (30)	1,089 (30.1)	547 (29)	177 (33)	—
Superior studies	472 (7.8)	297 (8.2)	136 (7.2)	39 (7.3)	—
Standard of living index (SLI) {tertiles [n (%)]}					0 (0)
Low (<24.6)	1,781 (29.5)	1,114 (30.8)	549 (29.1)	118 (22)	—
Medium (24.6–31.4)	1,893 (31.3)	1,159 (32.1)	599 (31.7)	135 (25.2)	—
High (>31.4)	2,365 (39.2)	1,342 (37.1)	740 (39.2)	283 (52.8)	—
Primary cooking fuel {gas/electricity [n (%)]}	2,490 (42.4)	1,329 (38.3)	848 (45.6)	313 (58.4)	172 (2.8)
Active smoker (combustion)	921 (15.3)	564 (15.6)	288 (15.3)	69 (12.9)	3 (0)
Active smoker (chew, snuff)	638 (10.6)	381 (10.5)	210 (11.1)	47 (8.8)	3 (0)
Physical activity {METs [AM (SD)]}	1.62 (0.21)	1.63 (0.21)	1.62 (0.22)	1.56 (0.18)	319 (5.3)
Stress {score [AM (SD)]}	10.6 (1.9)	10.6 (1.9)	10.7 (2)	10.6 (2.1)	14 (0.2)
Alcohol intake tertiles {g/d [n (%)]}					11 (0.2)
Low (<9.4)	1,886 (31.3)	1,123 (31.1)	608 (32.3)	155 (28.9)	—
Medium (9.4–45.1)	1,975 (32.8)	1,166 (32.3)	603 (32)	206 (38.4)	—
High (>45.1)	2,167 (35.9)	1,320 (36.6)	672 (35.7)	175 (32.6)	—
Salt intake {g/d [GM (GSD)]}	5.7 (1.6)	5.7 (1.6)	5.6 (1.6)	5.9 (1.6)	11 (0.2)
Fruit and vegetables intake {g/d [GM (GSD)]}	171.1 (2)	170.4 (1.9)	167.1 (2)	191.1 (1.9)	11 (0.2)
Total energy intake {kcal/d [GM (GSD)]}	2,125.8 (1.5)	2,129.6 (1.5)	2,109.5 (1.5)	2,157.7 (1.5)	11 (0.2)
Energy from carbohydrates {% [AM (SD)]}	68.1 (9.5)	68.4 (9.3)	67.9 (9.5)	66.8 (10.2)	11 (0.2)
Energy from fat {% [AM (SD)]}	17.2 (5.5)	17 (5.4)	17.2 (5.5)	17.8 (5.6)	11 (0.2)
$PM_{2.5}$ ambient air pollution {$\mu g/m^3$ [AM (SD)]}	32.9 (2.7)	32.3 (2.8)	33.3 (2.4)	35 (1.7)	11 (0.2)
Cardiometabolic risk factors [GM (GSD)]					
SBP (mmHg)	118.8 (1.1)	118.7 (1.1)	118.8 (1.1)	119.3 (1.1)	213 (3.5)[a]
DBP (mmHg)	77.6 (1.2)	77.3 (1.2)	77.9 (1.2)	78.6 (1.2)	213 (3.5)[a]
Waist circumference (cm)	72 (1.1)	71.7 (1.1)	72 (1.2)	73.9 (1.2)	15 (0.2)
Triglycerides (mg/dL)	108.6 (1.7)	107.4 (1.7)	111 (1.7)	108.1 (1.7)	203 (3.4)
Fasting glucose (mg/dL)	91.8 (1.2)	91.7 (1.2)	91.3 (1.2)	94 (1.2)	610 (10.1)[b]
Non-HDL cholesterol (mg/dL)	117.4 (1.4)	114.9 (1.4)	121.2 (1.3)	121.5 (1.3)	173 (2.9)
IHMRS [AM (SD)]	4.82 (4.32)	4.69 (4.27)	4.9 (4.32)	5.46 (4.62)	0 (0)

Note: —, No data; AM, arithmetic mean; DBP, diastolic blood pressure; GM, geometric mean; GSD, geometric standard deviation; HDL, high-density lipoprotein; IHMRS, INTERHEART modifiable risk score; MET, metabolic equivalent unit value; $PM_{2.5}$, particulate matter ≤ 2.5 μm in aerodynamic diameter; SBP, systolic blood pressure; SD, standard deviation.

[a]When analyzing the outcomes SBP and DBP, we classified data as missing when participants were taking hypertensive medication (3.1%), were missing information on room temperature (0.1%), or had measurement taken on the left arm (0.3%).

[b]When analyzing the outcome fasting glucose, we classified data as missing when participants were taking diabetes medication (1.5%), had not fasted for 8 h (5.9%), or were missing information for the fasting period (1.8%).

multinomial logit model, education and alcohol tertiles were imputed using proportional odds models, and predictive mean matching was used for the rest of the continuous variables. We checked each of the equations to avoid having highly collinear variables in the same model, we assessed the convergence of the chains, and we verified the plausibility of the imputed values by comparing the densities of the observed and imputed values.

We analyzed the cross-sectional associations between built-up trajectories and cardiometabolic risk factors using linear mixed models with nested random intercepts to account for the clustered structure of the data (households within villages). We fitted regression models in each of the imputed data sets, and then pooled them using Rubin's rules (Rubin 2004). We defined potential mediators and confounders of the exposure–outcome associations via a directed acyclic graph (see Figure S5). We considered three sequential sets of confounders: Model 1 was adjusted for baseline (1995) built-up area, sex, age (continuous, linear, and quadratic terms) and room temperature (continuous, for blood pressure outcomes only). Model 2 was further adjusted for health behaviors: smoking (no smoking/chew or snuff/combustion); salt, fruit and vegetables, and total energy intake (natural log-transformed, continuous); alcohol consumption (in tertiles); and percentage of energy coming from fat and carbohydrates (continuous). Model 3 also included socioeconomic confounders: education (illiterate/primary education/secondary education/superior studies), SLI (in tertiles), and primary cooking fuel (gas or electricity/biomass or coal). We fitted models to the whole sample and in sex and age (younger/older than 30 years of age) subgroups. As sensitivity analyses, we fitted the models using *a*) complete cases data sets, *b*) built-up trajectories based on 500-m buffers, *c*) village identifiers as a fixed effect (within-village models), *d*) difference in built-up land use (300 m) between 2009 and 1995 as an alternative continuous exposure, and *e*) alternative exposure indices (RF, NDVI, NDBI, UI) from 2009.

We performed mediation analyses to explore the influence of physical activity and $PM_{2.5}$ ambient air pollution on the association between 300-m–buffer built-up trajectories (as the exposure) and the six cardiometabolic outcomes. We also evaluated stress as a potential mediator *a priori* but did not include it in our mediation analysis because it was not associated with the exposure (see Figure S9) and was only weakly associated with the outcomes (see Figure S10). We computed indirect, direct, and total effects, adapting the multilevel mediation framework (Krull and MacKinnon 2001) to two mediators (MacKinnon 2012). That is, we fitted the following regression equations for each outcome y:

$$y_{vhi} = \theta_0 + \theta_1 \times BUslow_{vh} + \theta_2 \times BUfast_{vh} + conf_{vhi} + u_v + w_{vh} + \varepsilon_{vhi} \quad (1)$$

$$y_{vhi} = \beta_0 + \beta_1 \times BUslow_{vh} + \beta_2 \times BUfast_{vh} + \beta_3 \times physical_{vhi} + \beta_4 \times pm2.5_{vh} + conf_{vhi} + u_v + w_{vh} + \varepsilon_{vhi} \quad (2)$$

$$physical_{vhi} = \gamma_0 + \gamma_1 \times BUslow_{vh} + \gamma_2 \times BUfast_{vh} + conf_{vhi} + u_v + w_{vh} + \varepsilon_{vhi} \quad (3)$$

$$pm2.5_{vh} = \delta_0 + \delta_1 \times BUslow_{vh} + \delta_2 \times BUfast_{vh} + \varepsilon_{vh} \quad (4)$$

where v, h, and i are the indices for village, household, and individual; and u_v and w_{vh} are random intercepts at the village and household level, respectively. *BUslow* and *BUfast* refer to slow and fast built-up trajectories respectively, and *conf* refers to confounders. Equation 1 corresponds to the estimation of the exposure–outcome relationship; Equation 2 estimates the mediator–outcome associations while adjusting for exposure and other mediators; Equations 3 and 4 estimate the exposure–mediator associations. Equations 1–3 were adjusted for all confounders included in Model 3 of the main analysis. We estimated the effects for the fast built-up trajectory (reference: stable; the slow trajectory was not statistically significantly associated with any of the outcomes) as follows: We computed the indirect mediated effects for physical activity and air pollution as $\beta_3\gamma_2$ and $\beta_4\delta_2$, respectively. The total effect was estimated as θ_2 and the direct effect as β_2. We applied this analysis to each imputed data set and averaged point estimates across them to obtain pooled estimates of the effects. Percentile 95% confidence intervals (CIs) were derived by nonparametric bootstrapping (100 replications of each of the 20 imputed data sets, resampling done at the household unit). We tested for exposure–mediator (alpha = 0.05, p range: 0.49–0.9, tested in fully adjusted models) and mediator–mediator (alpha = 0.05, p range: 0.12–0.99, tested in fully adjusted models) interactions for the different outcomes but none was statistically significant (data not shown).

We made the following assumptions when estimating the total, direct, and indirect effects (VanderWeele 2016): *a*) exposure–outcome confounding, *b*) mediator–outcome confounding, and *c*) exposure–mediator confounding were adequately controlled for by covariate adjustment. Moreover, we assumed that *d*) no mediator–outcome confounder was affected by exposure, given that otherwise the confounder would be both a mediator and a confounder. Although there is no way to ensure that Assumptions a–c were met, we included a broad range of participant-level confounders that accounted for all hypothesized confounding factors between exposure and outcomes. Assumption d required further consideration because inclusion of multiple mediators could violate the assumption if mediators were correlated or if there were an interaction between the mediator effects on the outcome (VanderWeele 2016). We tested these conditions empirically: Correlation between physical activity and ambient $PM_{2.5}$ was weak ($r_{Pearson} = -0.08$), and the interactions between mediators in Equation 2 for the different outcomes were not statistically significant. Therefore, we considered Assumption d to be reasonable.

Data cleaning, analyses, and figures were done in R (version 3.4.0; R Development Core Team) using several packages (Bates et al. 2015; Genolini et al. 2015; Hijmans 2016; Liaw and Wiener 2002; Pebesma 2018; van Buuren and Groothuis-Oudshoorn 2011; Wickham 2017; Wilke 2017).

Results

Figure 1 shows the spatial distribution of built-up trajectories. Villages that were closer to the ring road and the two main roads and with a larger area tended to have a faster transition from greenspace to built-up land use. Median baseline built-up proportion in 1995 was similar in the three groups (0.21, 0.26, and 0.22 for the stable, slow, and fast groups, respectively). Between 1995 and 2009, the median proportion of residential surrounding built-up land use increased by 0.14 for the stable, 0.2 for the slow, and 0.35 for the fast groups (Figure 2).

Mean age of participants was 36 y (standard deviation 13.8) and 54% were male (Table 1). Most participants had manual occupations in agriculture, construction, or industry. The total time per week spent on physical activity was higher for the work domain and the total physical activity (in METs) was linearly correlated mainly with time spent on work or travel even when not accounting for time spent sitting at work (see Figure S6); recreational physical activity was scant (79% of men and 96% women did not report any time spent on recreational physical activities). A 42.1% of participants had SBP levels >130 mmHg and/or DBP >80 mmHg. Among males, 7.8% had waist circumferences >90 cm; 17.7% of females had waist circumferences

>80 cm. A 23.7% of participants had triglycerides >150 mg/dL and 19.9% had fasting glucose levels >100 mg/dL. The frequency of unskilled manual occupation and smoking was lower in the fast built-up land-use group, whereas in the SLI, use of gas or electric cooking fuels, fruit and vegetable intake, ambient $PM_{2.5}$, and most health outcomes levels were greater in the fast compared with other groups.

We observed positive associations between a fast increase in built-up land use within 300 m of the home and all cardiometabolic risk factors (Table 2). Compared with participants with stable trajectories, those with the largest increase in built-up land use had 1.5% (95% CI: 0.1, 2.9) higher SBP, 2.4% (95% CI: 0.6, 4.3) higher DBP, 2.1% (95% CI: 0.5, 3.8) higher waist circumference, and 1.6% (95% CI: −0.6, 3.8) higher fasting glucose in the fully adjusted models (Model 3). Associations were larger in magnitude, but less precise, for triglycerides [4% (95% CI: −2.4, 10.8)] and non-HDL cholesterol [3.2% (95% CI: −0.9, 7.5)]. Associations for slow increases in built-up land use (compared with stable) were positive, but not statistically significant, for triglycerides and non-HDL; for other cardiometabolic risk factors, associations were small, negative, and nonsignificant.

Models fit using data based on complete cases yielded similar estimates (see Table S2). Results using trajectories based on 500-m buffers around the home rather than 300-m buffers (see Table S3) resulted in larger estimates for SBP, DBP, waist circumference, and non-HDL for both slow and fast increases compared with stable [e.g., 3.2% (95% CI: 1.3, 5.2) waist circumference increase for fast group vs. stable]; however, effects for triglycerides and fasting glucose were mostly null. Within-village analyses results showed positive, though not statistically significant, estimates for all outcomes in the fast vs. the stable trajectory (see Table S4), suggesting that part of the observed associations in the main models occurred within village. Confidence intervals in these models were wider due to the limited exposure contrast within village (see Table S4). Associations using the difference in built-up land use (300-m buffer) between 2009 and 1995 as a continuous exposure were mostly aligned with the main results. We estimated nonlinear associations for SBP, DBP, and waist circumference: Positive associations were found for individuals with the highest increases in residential surrounding built-up land use, whereas more modest increases resulted in null associations (see Figure S7). Associations for triglycerides and fasting glucose were positive and fairly linear, though not statistically significant (see Figure S7). Alternative built-up indices generated for 2009 yielded similar results for all indicators (see Figure S8): Positive, small associations were found per interquartile range (IQR) change in additive inverse NDVI for SBP [0.5% (95% CI: 0, 1)], DBP [0.7% (95% CI: 0.1, 1.4)], waist circumference [0.4% (95% CI: −0.2, 1)], and triglycerides [1.7% (95% CI: −0.7, 4.2)] in fully adjusted models. We estimated a higher IHMRS in areas with a fast increase in built-up residential surrounding land use (300-m buffer) compared with stable [0.75 points (95% CI: 0.29, 1.2) higher IHMRS], whereas the association for the slow trajectory was smaller and not statistically significant (see Table S5).

Exposure–mediator associations revealed that fast built-up trajectories (compared with stable) were associated with higher long-term ambient $PM_{2.5}$ [2.6 $\mu g/m^3$ (95% CI: 2.2, 3)] and less physical activity [−0.05 METs (95% CI: −0.07, −0.03)] (see Figure S9). Furthermore, physical activity was negatively associated with all outcomes after confounder and other mediator adjustment [e.g., −22.8% (95% CI: −22.7, −17.5) effect in triglycerides per MET; see Figure S10]; estimates for ambient air pollution were generally positive, but not statistically significant [e.g., 0.11% (95% CI: −0.11, 0.32) effect in SBP per 1-$\mu g/m^3$ increase]. The exposure was not associated with stress (see Figure S9); stress also showed weak and inconsistent associations with the outcomes (see Figure S10). Stress was therefore not considered further as a potential mediator.

Figure 3 and Table S6 show total, direct, and mediated effects for physical activity and ambient $PM_{2.5}$ on cardiometabolic risk factors. Indirect effects for physical activity ranged from 0.18% (95% CI: 0.08, 0.3) between fast vs. stable land-use change for SBP to 1.31% (95% CI: 0.67, 1.92) for triglycerides. Estimated indirect effects for ambient $PM_{2.5}$ were less precise: 0.27% (95% CI: −0.05, 0.84) between fast vs. stable land-use change for SBP,

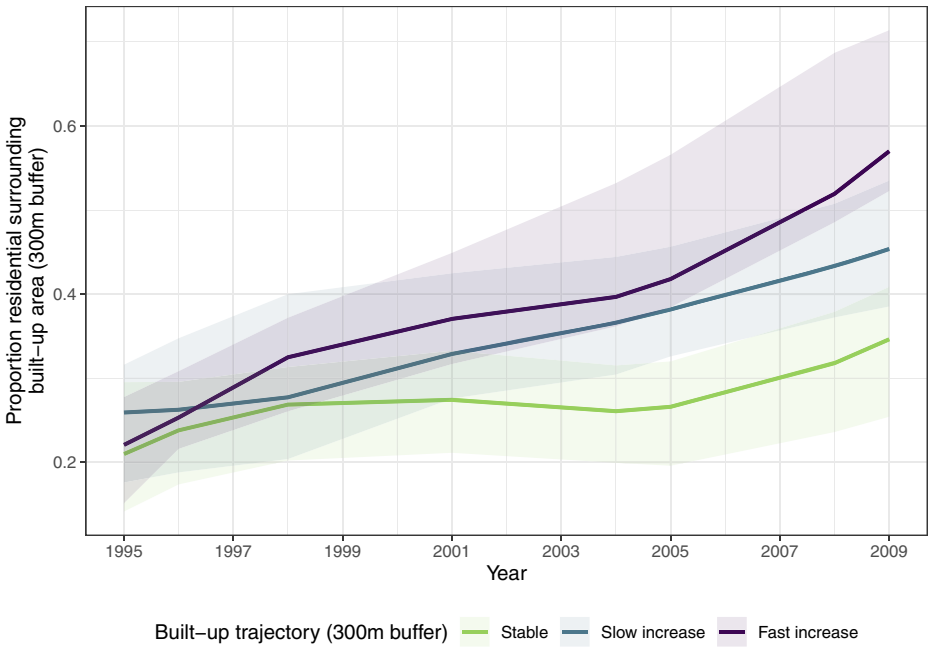

Figure 2. Median [interquartile range (IQR)] residential surrounding built-up land use proportion (300-m buffer) by year and exposure trajectory. Each median is represented by a dark line, with a shaded area represented its IQR.

Table 2. Associations and 95% confidence intervals (CIs) between residential surrounding built-up land use trajectories (300-m buffer) relative to reference (stable) and cardiometabolic risk factors.

Outcome	Model 1 [percent difference (95% CI)]	Model 2 [percent difference (95% CI)]	Model 3 [percent difference (95% CI)]
SBP			
Slow increase	−0.47 (−1.33, 0.4)	−0.49 (−1.35, 0.38)	−0.51 (−1.36, 0.35)
Fast increase	1.67 (0.28, 3.09)	1.62 (0.23, 3.04)	1.52 (0.13, 2.92)
DBP			
Slow increase	0.03 (−1.11, 1.19)	0 (−1.14, 1.16)	−0.04 (−1.17, 1.1)
Fast increase	2.72 (0.86, 4.61)	2.59 (0.74, 4.48)	2.41 (0.57, 4.28)
Waist circumference			
Slow increase	−0.27 (−1.34, 0.81)	−0.39 (−1.45, 0.68)	−0.45 (−1.48, 0.58)
Fast increase	2.67 (0.93, 4.43)	2.45 (0.74, 4.2)	2.12 (0.46, 3.8)
Triglycerides			
Slow increase	1.33 (−2.66, 5.49)	1.17 (−2.82, 5.32)	1.15 (−2.85, 5.31)
Fast increase	4.59 (−1.84, 11.43)	4.36 (−2.07, 11.2)	3.97 (−2.44, 10.81)
Fasting glucose			
Slow increase	−0.17 (−1.53, 1.21)	−0.23 (−1.59, 1.15)	−0.26 (−1.62, 1.12)
Fast increase	1.74 (−0.45, 3.97)	1.68 (−0.51, 3.91)	1.61 (−0.57, 3.84)
Non-HDL cholesterol			
Slow increase	2.95 (0.28, 5.68)	2.70 (0.05, 5.43)	2.66 (0.01, 5.37)
Fast increase	4.00 (−0.18, 8.37)	3.53 (−0.64, 7.87)	3.20 (−0.94, 7.52)

Note: Mixed effects linear models with nested random intercepts (household within village) were used with the following adjustments: Model 1: baseline built-up land use + sex + age + age^2 + room temperature (for blood pressure outcomes only); Model 2: Model 1 + smoking + salt + alcohol + fruit vegetables + energy intake + % fat + % carbohydrates; Model 3: Model 2 + education + SLI + cooking fuel. Models were fit to multiply imputed data sets and pooled following Rubin's rules. Percent difference in outcome associated with a given predictor calculated as $[\exp(\beta)\text{-}1] \times 100$. DBP, diastolic blood pressure; HDL, high-density lipoprotein; SBP, systolic blood pressure; SLI, standard of living index; % carbohydrates, percentage carbohydrates; % fat, percentage fat.

0.24% (95% CI: −0.2, 0.96) for DBP, 0.47% (95% CI: 0, 0.91) for waist circumference and 0.76% (95% CI: 0.17, 1.51) for fasting glucose. Air pollution indirect effects for triglycerides and non-HDL were mostly null.

Subgroup analysis (Figure 4; Table S7) suggested larger associations in women in all outcomes except non-HDL cholesterol. For example, fasting glucose was 4.7% (95% CI: 1.9, 7.6) higher in fast vs. stable land-use change in women and −1.4% (95% CI: −4.2, 1.5) in men. Results in age subgroups depended on the outcome: Stronger associations were observed in the older group for SBP, DBP, and fasting glucose [e.g., 3.7% (95% CI: 1.2, 6.3) change in DBP in participants >30 years of age vs. 1.5% (95% CI: −0.7, 3.6) for participants ≤30 years of age], whereas for non-HDL and triglycerides, associations were stronger for the younger group. Estimates for waist circumference were similar.

Discussion

Our results provide new evidence for the association between changes in residential surrounding built-up land use and cardiometabolic risk factors in a representative sample of residents in an urbanizing area of a LMIC. We estimated statistically significant increases in SBP, DBP, and waist circumference for participants with fast vs. stable increase in residential surrounding (300 m) built-up land use trajectories. Associations for fast vs. stable trajectories were also positive, but not significant for triglycerides, fasting glucose, and non-HDL cholesterol. Mediation analysis suggested partial mediation by physical activity and ambient $PM_{2.5}$. Subgroup analysis consistently suggested stronger associations in women.

The transition from greenspace to built-up land use between 1995 and 2009 was more marked in larger villages located closer

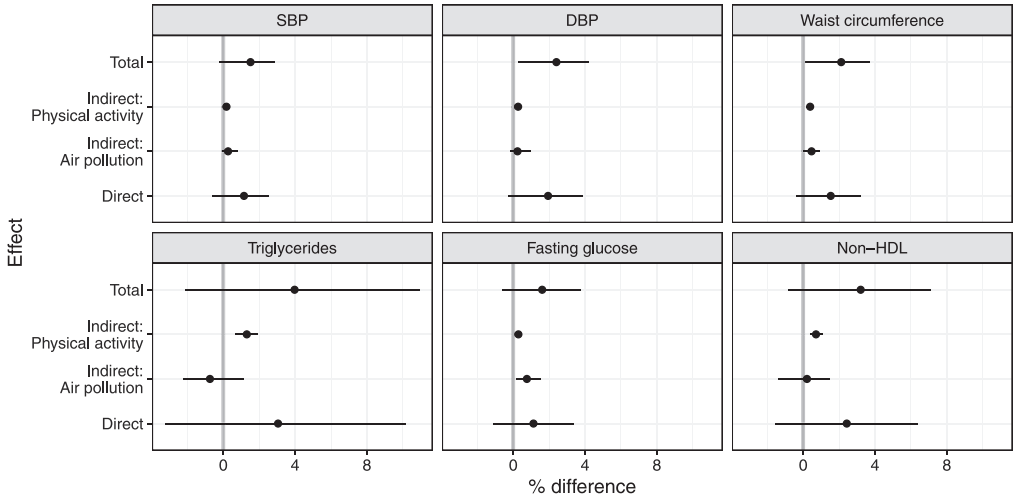

Figure 3. Total, indirect (physical activity), indirect (air pollution), and direct effects and 95% confidence intervals (CIs) between fast increase in residential surrounding built-up land use (300-m buffer) relative to reference (stable) and cardiometabolic risk factors (SBP, DBP, waist circumference, triglycerides, fasting glucose, and non-HDL cholesterol). Calculation of the effects took into account the multilevel nature of the data, the two candidate mediators and the multiply imputed data. CI was derived by bootstrapping. Percent difference in outcome associated with a given effect was calculated as $[\exp(\beta)\text{-}1] \times 100$. A table version of this figure is available in the Supplemental Material (see Table S6). Note: DBP, diastolic blood pressure; HDL, high-density lipoprotein; SBP, systolic blood pressure; % difference, percentage difference.

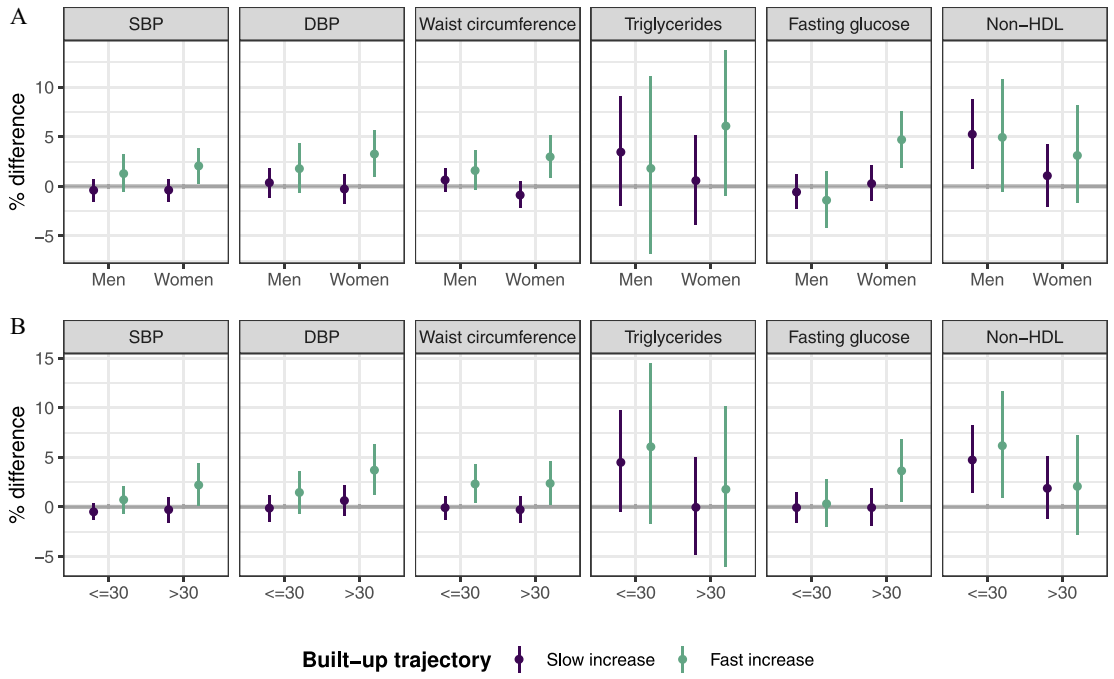

Figure 4. Associations and 95% confidence intervals between change in residential surrounding built-up land use relative to reference (stable) and cardiometabolic risk factors according to (A) sex and (B) age for SBP, DBP, waist circumference, triglycerides, fasting glucose, and non-HDL cholesterol. Mixed effects linear models with nested random intercepts (household within village) adjusted for baseline built-up, sex, age, age^2, room temperature (for blood pressure outcomes only), smoking, salt, alcohol, fruit and vegetables, energy intake, percentage fat, percentage carbohydrates, education, SLI, and cooking fuel. Models fit to multiply imputed data sets and pooled following Rubin's rules. Percent difference in outcome associated with a given predictor was calculated as $[\exp(\beta)\text{-}1]\times 100$. A table version of this figure is available in the Supplemental Material (see Table S7). Note: DBP, diastolic blood pressure; HDL, high-density lipoprotein; SBP, systolic blood pressure; SLI, standard of living index; % difference, percentage difference.

to the ring road and main roads. The loss of agricultural and barren areas and water bodies in the city fringe of Hyderabad has been previously reported in the remote sensing literature (Gibson et al. 2015; Roy et al. 2015; Wakode et al. 2014) and has been mostly due to expansion of residential areas (Wakode et al. 2014).

We estimated moderate increases in cardiometabolic risk factors in participants living in areas rapidly building up compared with areas where land use was stable. Our estimates for blood pressure were comparable to those from a cross-sectional study conducted in Chennai, India, evaluating the effect of residential surrounding greenness and impervious surfaces on markers of vascular aging (Lane et al. 2017). Lane et al. (2017) observed increases of 3.4% in SBP and of 1.5% in DBP for a 1-IQR decrease in NDVI. A recent meta-analysis showed a decrease in DBP of 2 mmHg in participants with the highest greenspace exposure vs. the lowest, estimated from 9,695 participants of 12 studies conducted in HICs (Twohig-Bennett and Jones 2018). Nevertheless, comparison of our results with studies included in the meta-analyses is challenging given that the meaning of greenspace exposure is likely to differ between peri-urban India and urban settings in HICs. The relationship between land use and health is likely to vary considerably across context; however, understanding of this context dependency is hampered by the lack of available evidence from LMICs, particularly in rapidly urbanizing areas.

The estimated effect of the 14-y change in built-up land use and cardiometabolic risk factors was larger than cross-sectional associations based on exposure metrics from a single year. This suggests that dynamic changes in land use may be more relevant for cardiometabolic health in urbanizing settings. Ji et al. (2019) evaluated the association between change in residential surrounding greenness and mortality in oldest-old participants (>80 years of age) in China and found no association. However, there are important differences between our study and that of Ji et al. (2019) in the demographics of the study population (adults vs. elder), as well as in the outcomes investigated (cardiometabolic markers vs. mortality), that limit the comparability of results.

Physical activity appeared to partially mediate associations with all cardiometabolic risk factors by a small extent. Partial mediation by physical activity has been observed in some studies conducted in HICs focusing on the association between greenspace and mortality (James et al. 2016). However, it is unclear whether the hypothesized mechanism of greenspace facilitating recreational physical activity in these studies would also apply to the APCAPS and similar populations. Ambient $PM_{2.5}$ appeared to partially mediate associations for blood pressure, waist circumference, and fasting glucose. Ambient $PM_{2.5}$ has been previously suggested to mediate the association between greenness and mortality (James et al. 2016). We estimated large, although imprecise, direct effects, possibly reflecting the role of other important pathways not included in our analysis, for example, social cohesion and biodiversity, noise, and air pollutants other than $PM_{2.5}$, or heat.

Subgroup analysis by sex consistently suggested larger associations in women. A possible explanation may be the differences in mobility patterns in the population: daytime spent near the residence was substantially longer for females (74%) than males (52%) (Sanchez et al. 2017). Similarly, a study evaluating the association between residential long-term ambient $PM_{2.5}$ and blood pressure in the same population observed stronger associations for women (Curto et al. 2019).

A major strength of our study is the exposure assessment: We derived a spatially resolved dynamic exposure metric that captured both the spatial and temporal components of variation in land use. Other strengths include analysis of a range of cardiometabolic

outcomes in a relatively large sample representative of the general population in a LMIC. We performed a formal mediation analysis considering the multilevel structure of the data and multiple mediators. The main limitation of our study was that cardiometabolic risk factors and mediators were measured once; we were therefore not able to estimate the effects of change in land use over time on change in cardiometabolic risk. We considered a binary built-up/greenspace classification of land use that did not allow differentiating the transition to built-up land use for different types of baseline greenspace. We did not have information about pesticides, a potentially harmful environmental exposure more prevalent in agricultural areas (Markevych et al. 2017). However, possible confounding by pesticide exposure could not explain the positive associations we observed. Our LUR model was fit on air pollution data (2015–2016) collected later than health outcomes (2010–2012) and the last built-up exposure measurement (2009); we therefore assumed the spatial pattern of ambient exposures to remain constant within this period and used geographical predictors in the LUR model collected at most 3 y later (2012–2013) than health outcomes. Our mediation analysis may have underestimated the influence of air pollution and physical activity due to measurement error in these variables. Finally, we cannot rule out the possibility of residual confounding by unmeasured confounders or confounders measured with error.

India is experiencing fast increases of built-up land use in urban fringes (Gibson et al. 2015; Roy et al. 2015). We provide new evidence that these changes are associated with moderate increases in several cardiometabolic risk factors. These relationships appear to operate, in part, through reduced physical activity and increased air pollution. Better integration of health into urban planning in rapidly urbanizing settings is critical to reducing the negative impacts and maximizing the benefits for health due to urban development.

Acknowledgments

We thank all participants of the Andhra Pradesh Children and Parent Study (APCAPS) and the Cardiovascular Health effects of Air pollution in Telangana, India (CHAI) study as well as all study teams involved in those projects. Landsat surface reflectance products were courtesy of the U.S. Geological Survey Earth Resources Observation and Science Center. We acknowledge past high-resolute imagery available in Google Earth. This work was supported by the European Research Council (grant agreement 336167 for the CHAI Project) and the Wellcome Trust (grant 084674/Z for the third follow-up of the APCAPS). C.T. was funded through a Ramón y Cajal fellowship (RYC-2015-17402) awarded by the Spanish Ministry of Economy and Competitiveness. Data related to the APCAPS cohort is available to researchers through a brief application to the cohort's Steering Group (form available from APCAPS website, http://apcaps.lshtm.ac.uk, and submitted by email to apcaps@iiphh.org).

References

Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. 2009. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 120(16):1640–1645, PMID: 19805654, https://doi.org/10.1161/CIRCULATIONAHA.109.192644.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48, https://doi.org/10.18637/jss.v067.i01.

Bowen L, Bharathi AV, Kinra S, Destavola B, Ness A, Ebrahim S. 2012. Development and evaluation of a semi-quantitative food frequency questionnaire for use in urban and rural India. Asia Pac J Clin Nutr 21(3):355–360.

Brenning A. 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package sperrorest. In: *2012 IEEE International Geoscience and Remote Sensing Symposium Proceedings*, 22–27 July, Munich, Germany, 5372–5375.

Caliński T, Harabasz J. 1974. A dendrite method for cluster analysis. Commun Stat Theory Methods 3(1):1–27, https://doi.org/10.1080/03610927408827101.

Corlin L, Lane KJ, Sunderarajan J, Chui KKH, Vijayakumar H, Krakoff L, et al. 2018. Urbanization as a risk factor for aortic stiffness in a cohort in India. PLoS One 13(8):e0201036, PMID: 30067798, https://doi.org/10.1371/journal.pone.0201036.

Curto A, Wellenius GA, Milà C, Sanchez M, Ranzani O, Marshall JD, et al. 2019. Ambient particulate air pollution and blood pressure in peri-urban India. Epidemiology 30(4):492–500, PMID: 31162282, https://doi.org/10.1097/EDE.0000000000001014.

Estoque RC, Murayama Y. 2015. Classification and change detection of built-up lands from Landsat-7 ETM+ and Landsat-8 OLI/TIRS imageries: a comparative assessment of various spectral indices. Ecol Indic 56:205–217, https://doi.org/10.1016/j.ecolind.2015.03.037.

Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Rojas-Rueda D, Plasència A, et al. 2016. Residential green spaces and mortality: a systematic review. Environ Int 86:60–67, PMID: 26540085, https://doi.org/10.1016/j.envint.2015.10.013.

Geldsetzer P, Manne-Goehler J, Theilmann M, Davies JI, Awasthi A, Danaei G, et al. 2018. Geographic and sociodemographic variation of cardiovascular disease risk in India: a cross-sectional study of 797,540 adults. PLoS Med 15(6):e1002581, PMID: 29920517, https://doi.org/10.1371/journal.pmed.1002581.

Genolini C, Alacoque X, Sentenac M, Arnaud C. 2015. kml and kml3d: R packages to cluster longitudinal data. J Stat Softw 65(4):1–34, https://doi.org/10.18637/jss.v065.i04.

Genolini C, Falissard B. 2011. Kml: a package to cluster longitudinal data. Comput Methods Programs Biomed 104(3):e112–e121, PMID: 21708413, https://doi.org/10.1016/j.cmpb.2011.05.008.

Gibson J, Boe-Gibson G, Stichbury G. 2015. Urban land expansion in India 1992–2012. Food Policy 56:100–113, https://doi.org/10.1016/j.foodpol.2015.08.002.

Hastie T, Tibshirani R, Friedman J. 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.

Hijmans RJ. 2016. Raster geographic data analysis and modeling. https://cran.r-project.org/web/packages/raster/index.html [accessed 28 March 2020].

India State-Level Disease Burden Initiative Collaborators. 2017. Nations within a nation: variations in epidemiological transition across the states of India, 1990–2016 in the Global Burden of Disease Study. Lancet 390(10111):2437–2460, PMID: 29150201, https://doi.org/10.1016/S0140-6736(17)32804-0.

India State-Level Disease Burden Initiative CVD Collaborators. 2018. The changing patterns of cardiovascular diseases and their risk factors in the states of India: the Global Burden of Disease Study 1990–2016. Lancet Glob Health, PMID: 30219317, https://doi.org/10.1016/S2214-109X(18)30407-8.

James P, Banay RF, Hart JE, Laden F. 2015. A review of the health benefits of greenness. Curr Epidemiol Rep 2(2):131–142, PMID: 26185745, https://doi.org/10.1007/s40471-015-0043-7.

James P, Hart JE, Banay RF, Laden F. 2016. Exposure to greenness and mortality in a nationwide prospective cohort study of women. Environ Health Perspect 124(9):1344–1352, PMID: 27074702, https://doi.org/10.1289/ehp.1510363.

Ji JS, Zhu A, Bai C, Wu C-D, Yan L, Tang S, et al. 2019. Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study. Lancet Planet Health 3(1):e17–e25, PMID: 30654864, https://doi.org/10.1016/S2542-5196(18)30264-X.

Kinra S, Radha Krishna KV, Kuper H, Rameshwar Sarma KV, Prabhakaran P, Gupta V, et al. 2014. Cohort profile: Andhra Pradesh Children and Parents Study (APCAPS). Int J Epidemiol 43(5):1417–1424, PMID: 24019421, https://doi.org/10.1093/ije/dyt128.

Krull JL, MacKinnon DP. 2001. Multilevel modeling of individual and group level mediated effects. Multivariate Behav Res 36(2):249–277, PMID: 26822111, https://doi.org/10.1207/S15327906MBR3602_06.

Lane KJ, Stokes EC, Seto KC, Thanikachalam S, Thanikachalam M, Bell ML. 2017. Associations between greenness, impervious surface area, and nighttime lights on biomarkers of vascular aging in Chennai, India. Environ Health Perspect 125(8):087003, PMID: 28886599, https://doi.org/10.1289/EHP541.

Larkin A, van Donkelaar A, Geddes JA, Martin RV, Hystad P. 2016. Relationships between changes in urban characteristics and air quality in East Asia from 2000 to 2010. Environ Sci Technol 50(17):9142–9149, PMID: 27442110, https://doi.org/10.1021/acs.est.6b02549.

Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News 2:18–22. http://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf [accessed 28 March 2020].

MacKinnon D. 2012. *Introduction to Statistical Mediation Analysis*. New York, NY: Routledge.

Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, et al. 2017. Exploring pathways linking greenspace to health: theoretical and methodological

guidance. Environ Res 158:301–317, PMID: 28672128, https://doi.org/10.1016/j.envres.2017.06.028.
Masek JG, Vermote EF, Saleous NE, Wolfe R, Hall FG, Huemmrich KF, et al. 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geoscience Remote Sensing Lett 3(1):68–72, https://doi.org/10.1109/LGRS.2005.857030.
Matsuzaki M, Sullivan R, Ekelund U, Krishna KVR, Kulkarni B, Collier T, et al. 2016. Development and evaluation of the Andhra Pradesh Children and Parent Study Physical Activity Questionnaire (APCAPS-PAQ): a cross-sectional study. BMC Public Health 16(1):48, PMID: 26787268, https://doi.org/10.1186/s12889-016-2706-9.
McGorrian C, Yusuf S, Islam S, Jung H, Rangarajan S, Avezum A, et al. 2011. Estimating modifiable coronary heart disease risk in multiple regions of the world: the INTERHEART Modifiable Risk Score. Eur Heart J 32(5):581–589, PMID: 21177699, https://doi.org/10.1093/eurheartj/ehq448.
Pebesma E. 2018. sf: simple features for R. R J 10(1):439–446, https://doi.org/10.32614/RJ-2018-009.
Prabhakaran D, Jeemon P, Roy A. 2016. Cardiovascular diseases in India: current epidemiology and future directions. Circulation 133(16):1605–1620, PMID: 27142605, https://doi.org/10.1161/CIRCULATIONAHA.114.008729.
Rogan J, Chen D. 2004. Remote sensing technology for mapping and monitoring land-cover and land-use change. Prog Plann 61(4):301–325, https://doi.org/10.1016/S0305-9006(03)00066-7.
Roy PS, Roy A, Joshi PK, Kale MP, Srivastava VK, Srivastava SK, et al. 2015. Development of decadal (1985–1995–2005) land use and land cover database for India. Remote Sens (Basel) 7(3):2401–2430, https://doi.org/10.3390/rs70302401.
Rubin DB. 2004. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons.
Sanchez M, Ambros A, Milà C, Salmon M, Balakrishnan K, Sambandam S, et al. 2018. Development of land-use regression models for fine particles and black carbon in peri-urban South India. Sci Total Environ 634:77–86, PMID: 29626773, https://doi.org/10.1016/j.scitotenv.2018.03.308.
Sanchez M, Ambros A, Salmon M, Bhogadi S, Wilson RT, Kinra S, et al. 2017. Predictors of daily mobility of adults in peri-urban South India. Int J Environ Res Public Health 14(7):783, PMID: 28708095, https://doi.org/10.3390/ijerph14070783.
Spitzer RL, Williams JBW, Kroenke K, Hornyak R, McMurray J. 2000. Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. Am J Obstet Gynecol, PMID: 10992206, https://doi.org/10.1067/mob.2000.106580.
Taylor L, Hochuli DF. 2017. Defining greenspace: multiple uses across multiple disciplines. Landsc Urban Plan 158:25–38, https://doi.org/10.1016/j.landurbplan.2016.09.024.
Tonne C, Salmon M, Sanchez M, Sreekanth V, Bhogadi S, Sambandam S, et al. 2017. Integrated assessment of exposure to $PM_{2.5}$ in South India and its relation with cardiovascular risk: design of the CHAI observational cohort study. Int J Hyg Environ Health 220(6):1081–1088, PMID: 28606699, https://doi.org/10.1016/j.ijheh.2017.05.005.
Twohig-Bennett C, Jones A. 2018. The health benefits of the great outdoors: a systematic review and meta-analysis of greenspace exposure and health outcomes. Environ Res 166:628–637, PMID: 29982151, https://doi.org/10.1016/j.envres.2018.06.030.
UN DESA (Population Division of the United Nations Department of Economic and Social Affairs). 2018. *World Urbanization Prospects: The 2018 Revision*, online ed. https://population.un.org/wup/ [accessed 28 March 2020].
van Buuren S, Groothuis-Oudshoorn K. 2011. mice: multivariate imputation by chained equations in R. J Stat Softw 45(3):1–67, https://doi.org/10.18637/jss.v045.i03.
VanderWeele TJ. 2016. Mediation analysis: a practitioner's guide. Annu Rev Public Health 37(1):17–32, PMID: 26653405, https://doi.org/10.1146/annurev-publhealth-032315-021402.
Wakode HB, Baier K, Jha R, Azzam R. 2014. Analysis of urban growth using Landsat TM/ETM data and GIS—a case study of Hyderabad, India. Arab J Geosci 7(1):109–121, https://doi.org/10.1007/s12517-013-0843-3.
Wickham H. 2017. tidyverse: Easily Install and Load the 'Tidyverse'. http://tidyverse.tidyverse.org [accessed 28 March 2020].
Wilke CO. 2017. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. https://wilkelab.org/cowplot/ [accessed 28 March 2020].