American Journal of Epidemiology Advance Access published April 25, 2012

American Journal of Epidemiology

DOI: 10.1093/aje/kwr434

Practice of Epidemiology

Multicollinearity in Associations Between Multiple Environmental Features and Body Weight and Abdominal Fat: Using Matching Techniques to Assess Whether the Associations are Separable

Cinira Leal*, Kathy Bean, Frédérique Thomas, and Basile Chaix

* Correspondence to Dr. Cinira Leal, INSERM U707, Faculté de Médecine Saint-Antoine, 27 rue Chaligny, 75012 Paris, France (e-mail: ciniraleal@hotmail.com).

Initially submitted June 3, 2011; accepted for publication November 2, 2011.

Because of the strong correlations among neighborhoods' characteristics, it is not clear whether the associations of specific environmental exposures (e.g., densities of physical features and services) with obesity can be disentangled. Using data from the RECORD (Residential Environment and Coronary Heart Disease) Cohort Study (Paris, France, 2007–2008), the authors investigated whether neighborhood characteristics related to the sociodemographic, physical, service-related, and social-interactional environments were associated with body mass index and waist circumference. The authors developed an original neighborhood characteristic-matching technique (analyses within pairs of participants similarly exposed to an environmental variable) to assess whether or not these associations could be disentangled. After adjustment for individual/neighborhood socioeconomic variables, body mass index/waist circumference was negatively associated with characteristics of the physical/service environments reflecting higher densities (e.g., proportion of built surface, densities of shops selling fruits/vegetables, and restaurants). Multiple adjustment models and the neighborhood characteristic-matching technique were unable to identify which of these neighborhood variables were driving the associations because of high correlations between the environmental variables. Overall, beyond the socioeconomic environment, the physical and service environments may be associated with weight status, but it is difficult to disentangle the effects of strongly correlated environmental dimensions, even if they imply different causal mechanisms and interventions.

body mass index; environment; epidemiologic methods; matching; residence characteristics; waist circumference

Abbreviations: BMI, body mass index; HDI, Human Development Index; RECORD, Residential Environment and Coronary Heart Disease; TRIRIS, trois Ilots Regroupés pour l'Information Statistique.

Numerous studies have investigated relations between residential environmental characteristics and obesity (1). However, only a few studies have simultaneously explored the effects of various dimensions related to the sociodemographic, physical, service-related, and social-interactional environments (2–6).

To develop efficient public health interventions addressing the obesity epidemic, it is important to identify exactly which aspects of the environment influence obesity risk (7). For example, demonstrating that the density of fast-food restaurants was associated with obesity risk and demonstrating that the density of sports facilities was associated with obesity risk would lead to different interventions. However, many neighborhood characteristics, especially those related to the densities of physical features and services, are strongly correlated with each other. Therefore, it is not clear whether it is even possible to disentangle the effects of these different environmental dimensions, even if they are hypothesized to influence obesity risk through distinct causal pathways. An important methodological concern is that regression models provide estimates of environmental effects adjusted for each other even if these associations are not grounded on sufficient data allowing separation of the effects (8, 9).

Our empirical objective was to study relations between numerous correlated neighborhood characteristics related to the sociodemographic, physical, service-related, and social-interactional environments and body mass index (BMI) or waist circumference after adjustment for individual and neighborhood socioeconomic characteristics. Our methodological objective was to assess whether or not it is possible to disentangle the associations among these neighborhood variables, which are strongly correlated with each other, using both classical multiple adjustment and an original neighborhood characteristic-matching technique based on analysis of each environmental effect within pairs of individuals similarly exposed to another environmental variable.

MATERIALS AND METHODS

Population

Data from the first wave of the RECORD (Residential Environment and Coronary Heart Disease) Cohort Study (www.record-study.org) were used for the analyses. As described elsewhere (9–14), 7,230 participants aged 30–79 years were recruited without a priori sampling in 2007–2008 during free preventive medical checkups conducted by the Centre d'Investigations Préventives et Cliniques in the Paris, France, metropolitan area. As an eligibility criterion, participants were residing in one of 10 (out of 20) administrative districts of Paris or one of 111 other municipalities in the metropolitan area selected a priori. Of the persons selected for participation, 83.6% agreed to participate and completed the data collection protocol. All participants underwent a physical examination, completed questionnaires, and were geocoded on the basis of their residential address in 2007–2008. The study protocol was approved by the French Data Protection Authority. After we excluded persons with missing anthropometric values, 7,230 participants were included in the analyses for BMI and 7,076 were included in the analyses for waist circumference, all of them living in 646 neighborhoods (a TRIRIS (trois Ilots Regroupés pour l'Information Statistique) geographic unit (15) comprising, on average, 11 participants (interdecile range (10th–90th percentiles), 4–19)).

Measures

Outcomes. Data related to height, weight, and waist circumference were obtained from medical examinations. Height (using a wall-mounted stadiometer) and weight (using calibrated scales) were recorded by a nurse, allowing for the calculation of BMI (weight (kg) divided by height (m) squared) (16, 17). Waist circumference was measured in centimeters using an inelastic tape placed midway between the lower ribs and the iliac crests on the midaxillary line (18). BMI and waist circumference were analyzed as continuous variables.

Individual variables. The following sociodemographic characteristics of participants were considered in our analyses: age, sex, education, mother's and father's education, occupation, household income, and Human Development Index (HDI) of the participant's country of birth. Age was divided into 3 classes: 30–44, 45–59, and 60–79 years. Education was divided into 4 classes: no education, primary education or lower secondary education, higher secondary education or lower tertiary education, and upper tertiary education. Mother's and father's educational levels were divided into 3 classes: primary school or less, secondary school, and tertiary school. Household income adjusted for household size was divided into 4 categories. Occupation was coded in 4 categories: high white-collar, intermediate, low white-collar, and blue-collar. We followed the approach of Beckman et al. (19) in attributing to each individual the 2004 HDI of his/her country of birth, as a proxy for the country's social development level. Following the United Nations Development Programme (20), binary variables were used to distinguish among people born in low-development countries (HDI <0.5), medium-development countries (HDI 0.5–0.8), and high-development countries (HDI >0.8). Five individual variables related to physical activity and dietary habits were considered in our analyses: total walking time over the previous 7 days, energy expenditure at work over the previous 7 days, energy expenditure in recreational and sports activities over the previous 7 days, average number of fruits and vegetables consumed daily, and frequency of consumption of fast-food meals.

Neighborhood variables. Details on the neighborhood sociodemographic, physical, service-related, and social-interactional variables considered in this study are reported in Table 1. Most variables were defined within 500-m radius circular buffers centered on each participant's residence. Other neighborhood variables were defined using the ecometric method, through which individuals' perceptions were aggregated at the neighborhood level (TRIRIS geographic unit) based on 3-level (perception items, individuals, area units) hierarchical modeling (21). Neighborhood variables were analyzed as standardized continuous variables and as 4 categories.

Statistical analysis

To account for within-neighborhood correlation in BMI/waist circumference, we used multilevel linear regression models with random effects estimated at the TRIRIS level.

Statistical analyses took place in 3 steps:

1. Considering all of the individual sociodemographic variables, only those that were associated with BMI/waist circumference were retained in the models.
2. We estimated associations between the different neighborhood characteristics and BMI/waist circumference, adjusted for individual sociodemographic covariates (each environmental variable in a separate model). Then we examined whether associations persisted after adjustment for neighborhood socioeconomic level. As additional analyses, first we tested whether sex modified the relations between environmental characteristics and BMI/waist circumference. Second, we examined whether variables related to physical activity and dietary behavior mediated some of the relations between environmental characteristics and BMI/waist circumference.
3. Considering the environmental variables that were associated with the outcomes in step 2, we conducted additional analyses to attempt to disentangle the different associations (strong correlations existed among the environmental variables).

First, we fitted models adjusted for individual sociodemographic characteristics that simultaneously included 2 environmental variables beyond neighborhood socioeconomic status, to assess whether the associations could be separated through multiple adjustment. Second, we developed an original neighborhood characteristic-matching technique that was applied to the environmental variables that were most strongly associated with BMI/waist circumference.

The principle of this matching technique is to estimate the effect of a neighborhood variable X_1 within pairs of individuals similarly exposed to another environmental variable X_2 that is correlated with X_1. A possible strategy if the outcome were binary would be to use conditional logistic regression to assess whether a given environmental variable X_1 was associated with the outcome within pairs of individuals having comparable values for X_2 (22). In order to apply a similar strategy to our continuous outcomes, we determined a risk score for BMI/waist circumference based on the individual/neighborhood socioeconomic variables retained in the model. We constructed pairs of participants with comparable values (see details below) for a neighborhood variable X_2. We then fitted nonmultilevel linear models in which the difference in BMI/waist circumference between the 2 individuals of the pair was used as the outcome and in which the explanatory variables were the difference in variable X_1 and the difference in risk score between the individuals of the pair. Statistical units in these analyses were not the individuals themselves but pairs of individuals. We standardized the difference in exposure to X_1 within pairs of participants to obtain comparable results across analyses matched on neighborhood characteristics. The final models were fitted as follows:

$$\Delta\text{BMI in pairs (defined according to factor } X_2) = \alpha + \beta_1 \Delta X_1 + \beta_2 \Delta \text{ risk score.}$$

Regarding assessment of uncertainty in these estimates, we randomly generated 500 different combinations of pairs of participants having comparable values for the environmental variable X_2 (the participants were ranked by blocks of 10 observations according to the value of X_2, and the pairs were generated within each block). Fitting the model in these 500 different samples, we report the median value of coefficients as the effect estimates and the 2.5th and 97.5th percentile ranges as 95% credible intervals.

As a validation of the approach, we expected that an association observed in a classical regression analysis between an environmental variable X_1 and BMI/waist circumference would also be observed within pairs of participants constructed on the basis of a random variable (rather than on the basis of a correlated environmental variable X_2). Moreover, with these neighborhood characteristic-matched analyses, we expected that the stronger the correlation between environmental variables X_1 and X_2, the larger would be the decrease in the difference of exposure to environmental variable X_1 within pairs of participants exposed to a comparable value of X_2 and the less likely we would be to observe any effect of X_1 within such pairs.

All of the analyses were conducted using SAS 9.2 (SAS Institute Inc., Cary, North Carolina).

RESULTS

Descriptive information about the RECORD participants is presented in Web Table 1, which is posted on the *Journal*'s website (http://aje.oxfordjournals.org/). Mean BMI and waist circumference were 25.4 (25th–75th percentile range, 22.6–27.7) and 85.1 cm (25th–75th percentile range, 76–93), respectively.

In a model adjusting for individual/neighborhood covariates, higher BMIs and waist circumferences were observed among older participants, males, people born in low-HDI countries, those with a low educational level, those with less educated mothers, and those living in less educated neighborhoods (Web Table 2).

After we accounted for neighborhood educational level, the other neighborhood socioeconomic variables were not associated with BMI/waist circumference. Accordingly, only neighborhood education was retained as an indication of neighborhood socioeconomic level in the following analyses.

The interaction tests showed that the associations between neighborhood education and BMI/waist circumference were modified by sex and were stronger for women (Web Table 2). Thus, the term for interaction between neighborhood education and sex was introduced in all subsequent analyses (multiple adjustment models and matched analyses).

After adjustment for neighborhood education and its interaction with sex, only a few environmental variables remained associated with BMI/waist circumference (Table 2). Population density was negatively associated with waist circumference but not with BMI. The proportion of built surface and (contrary to expectation) the level of neighborhood physical deterioration were negatively associated with both outcomes.

Regarding the food environment, the total number of restaurants, the number of fast-food restaurants, the number of traditional restaurants, and the number of shops selling fruits/vegetables were negatively associated with BMI/waist circumference after adjustment for individual/neighborhood socioeconomic covariates. As is shown in Web Table 3 with environmental variables divided into quartiles, of these 4 environmental variables, only the number of shops selling fruits/vegetables showed a dose-response pattern of association with BMI/waist circumference.

Apart from food services, the availability of basic services in the neighborhood (e.g., banks, post offices) was negatively associated with BMI/waist circumference. Variables of the social-interactional environment were not associated with BMI/waist circumference.

None of the associations between these specific environmental variables and BMI/waist circumference interacted with sex after adjustment for the interaction between neighborhood education and sex (results not reported).

Regarding mediation analyses, as shown in Web Table 2, lower energy expenditure related to sports and recreational activities was associated with higher BMI and greater waist circumference, while more frequent consumption of fast-food meals was associated only with higher BMI. Accounting for these behavioral mediators did not substantially change the associations between environmental exposures and BMI/waist circumference (Web Table 4).

As Table 3 shows, the environmental variables that were associated with BMI/waist circumference were generally strongly correlated with each other: Among the 36 Pearson correlations estimated, the median correlation was 0.72, and the interdecile range was 0.39–0.97.

Table 1. Operational Definitions of Neighborhood Variables Considered in Analyses of Neighborhood Characteristics and Body Mass Index/Waist Circumference, Paris, France, 2007–2008

Environmental Dimension and Neighborhood Variable	Definition	Data Source (Year)
Sociodemographic characteristics		
Educational level[a]	Proportion of residents aged >15 years with an upper tertiary education	French national census (2006)
Household income[a]	Median household income per consumption unit	Tax registry of the General Directorate of Taxation (2006)
Real estate prices[a]	Mean value of dwellings sold in 2003–2007	Paris-Notaries (2007)
Population density[b]	Population density in the neighborhood	French national census (2006)
Physical characteristics		
Proportion of built surface[a]	Proportion of the neighborhood covered with buildings	3-dimensional data from IGN on buildings' ground shape and height (2008)
Mean building height[a]	Mean building height weighted by each building's ground size	3-dimensional data from IGN on buildings' ground shape and height (2008)
Alpha index[a]	Ratio of the number of actual circuits (closed paths) to the maximum number of circuits	Data on road network from IGN (2008)
Gamma index[a]	Ratio of the number of links between nodes in the network to the maximum possible number of links	Data on road network from IGN (2008)
Connected node ratio[a]	No. of street intersections with at least 3 ways (streets) divided by the number of intersections plus cul-de-sacs	Data on road network from IGN (2008)
Density of street intersections[a]	Density of intersections with at least 4 ways	Data on road network from IGN (2008)
Street density[a]	Density of streets	Data on road network from IGN (2008)
No. of monuments[a]	Count of monuments (historical or other)	Data from IAU-IdF (2005)
Proportion of the area covered by water[a]	Proportion of the area covered by water	Linear and polygonal data from IAU-IdF (2008)
Proportion of the area with parks or green spaces[a]	Proportion of the area covered by parks or green spaces	Linear and polygonal data from IAU-IdF (2008)
Level of neighborhood physical deterioration[a]	Ecometric variable	4 items from the RECORD questionnaire (2007–2008)
Active living potential[c]	Ecometric variable	3 items from the RECORD questionnaire (2007–2008)

Table continues

As Web Tables 5 and 6 demonstrate, most of the associations observed between environmental variables and BMI/waist circumference after adjustment for neighborhood education and its interaction with sex did not persist when we included a third environmental variable in the model (neighborhood education, however, remained associated in all models). As the only association that persisted in such models with 3 environmental variables, the number of shops selling fruits/vegetables remained associated with waist circumference after adjustment for population density, but the association with fruit/vegetable shops disappeared after adjustment for other variables, such as the total number of restaurants or the number of basic services.

As Table 4 shows, the associations between environmental variables and waist circumference observed with classical regression models after adjustment for neighborhood education were retrieved with matched analyses using pairs of participants as the statistical units of analysis when pairs of participants were generated on the basis of a random variable (similar results for BMI are reported in Web Table 7). Table 4 shows analyses of these environmental effects of interest successively estimated within pairs generated on the basis of the following environmental variables: the variable with the lowest level of correlation with the environmental exposure of interest that was not independently associated with the outcome and variables associated with the outcome that showed the lowest level, median level, and highest level of correlation with the exposure of interest.

The difference in environmental exposure X_1 within pairs of participants similarly exposed to environmental variable X_2 decreased with increasing degree of correlation between X_1 and X_2 (Table 4). The environmental effects of interest tended to persist when estimated within pairs of participants with similar values of an environmental variable X_2 that was not associated with the outcome and only modestly correlated with

Table 1. Continued

Environmental Dimension and Neighborhood Variable	Definition	Data Source (Year)
Service-related characteristics		
No. of supermarkets[a]	Count of services	Data from INSEE's Permanent Database of Facilities (2008)
No. of shops selling fruits/vegetables (including street markets)[a]	Count of services	Data from INSEE's SIRENE database (2007)
Total no. of restaurants[a]	Count of services	Data from INSEE's SIRENE database (2007)
No. of traditional restaurants[a]	Count of services	Data from INSEE's SIRENE database (2007)
No. of fast-food restaurants[a]	Count of services	Data from INSEE's SIRENE database (2007)
Proportion of fast-food restaurants[a]	No. of fast-food restaurants divided by the total no. of restaurants	Data from INSEE's SIRENE database (2007)
No. of sports facilities[a]	Count of sports facilities	Data from the Census of Sport Facilities from the Paris region of the Regional Directorate of Youth, Sports, and Social Cohesion (2008)
No. of transportation lines[a]	Count of different transit lines (buses, metros, and trains)	Data from the Syndicate of Transports of the Paris Region (2008)
No. of basic services[a]	Count of destinations (administrative agencies, shops, entertainment facilities, etc.)	Data from INSEE's Permanent Database of Facilities (2008)
No. of health-care resources[a]	Count of health-care resources	Data from the Geographic Database for Health-Related Practices (Regional Union of Health Insurance Offices Ile-de-France database)
Social-interactional characteristics		
Deteriorated social interactions[c]	Ecometric variable	5 items from the RECORD questionnaire (2007–2008)
Social cohesion[c]	Ecometric variable	4 items from the RECORD questionnaire (2007–2008)
Collective feeling of insecurity[c]	Ecometric variable	1 item from the RECORD questionnaire (2007–2008)

Abbreviations: IAU-IdF, Institute of Urban Planning–Ile-de-France; IGN, National Geographic Institute; INSEE, National Institute of Statistics and Economic Studies; RECORD, Residential Environment and Coronary Heart Disease; SIRENE, Database for a Register of Companies and Their Establishments; TRIRIS, trois Ilots Regroupés pour l'Information Statistique.

[a] The variable was measured in a 500-m radius circular area centered on each participant's residence.

[b] The variable was measured at the census block-group level (fixed boundaries not centered on participants' residences).

[c] The variable was measured at the TRIRIS (census tract) neighborhood level (fixed boundaries not centered on participants' residences).

the exposure. However, as is shown in Table 4 and in Web Tables 7 and 8, when pairs of participants were generated with environmental variables that were associated with the outcome, all of the observed associations between environmental variables and BMI/waist circumference disappeared.

DISCUSSION

We observed that specific demographic, physical, and service-related characteristics of the residential neighborhood were associated with BMI/waist circumference, even after adjustment for individual, maternal, and neighborhood socioeconomic variables. Numerous studies have investigated relations between neighborhood characteristics related to the physical environment/service availability and weight status or abdominal fat (27 articles published through December 2009, based on our recent literature review (1)), but only 8 of them examined environmental variables of these 2 dimensions together after also controlling for the socioeconomic characteristics of the environment as we did (6, 23, 24).

Strengths and limitations

Regarding study strengths, BMI and waist circumference were measured by trained nurses, ensuring the quality of the data, while approximately 60% of the studies in the field have relied on self-reported height and weight (1). Furthermore, none of the studies on the effects of the physical, service-related, and social-interactional environments have considered a measure of abdominal fat beyond BMI (1). Additional study strengths

Table 2. Associations Between Neighborhood Characteristics (Standardized Continuous Variables) and Body Mass Index or Waist Circumference (Separate Models), Paris, France, 2007–2008

Environmental Dimension and Neighborhood Variable	Body Mass Index[a]				Waist Circumference, cm			
	Model 1[b]		Model 2[c]		Model 3[b]		Model 4[c]	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI
Sociodemographic characteristics								
Population density	−0.21**	−0.32, −0.11	−0.08	−0.19, 0.02	−0.56**	−0.84, −0.29	−0.28*	−0.56, −0.01
Physical characteristics								
Proportion of built surface	−0.31**	−0.42, −0.21	−0.12*	−0.24, −0.01	−0.79**	−1.07, −0.51	−0.38*	−0.69, −0.08
Mean building height	−0.27**	−0.38, −0.16	−0.07	−0.18, 0.05	−0.58**	−0.86, −0.30	−0.13	−0.44, 0.18
Alpha index	−0.21**	−0.32, −0.11	−0.06	−0.17, 0.04	−0.59**	−0.86, −0.32	−0.27	−0.55, 0.01
Gamma index	−0.21**	−0.32, −0.11	−0.07	−0.17, 0.04	−0.59**	−0.86, −0.32	−0.27	−0.56, 0.01
Connectivity node ratio	−0.16**	−0.27, −0.06	−0.04	−0.14, 0.07	−0.50**	−0.77, −0.23	−0.22	−0.50, 0.05
Density of intersections	−0.23**	−0.33, −0.12	−0.08	−0.18, 0.03	−0.47**	−0.74, −0.20	−0.13	−0.42, 0.15
Street density	−0.08**	−0.18, 0.02	−0.01	−0.10, 0.09	−0.20	−0.47, 0.06	−0.03	−0.30, 0.23
No. of monuments	−0.04	−0.14, 0.07	0.02	−0.08, 0.12	−0.05	−0.32, 0.21	0.08	−0.18, 0.34
Proportion of the area covered by water	0.02	−0.08, 0.12	0.05	−0.05, 0.14	0.07	−0.19, 0.34	0.14	−0.12, 0.40
Proportion of the area with parks	0.01	−0.09, 0.11	0.001	−0.10, 0.10	−0.01	−0.27, 0.26	0.003	−0.26, 0.26
Level of neighborhood physical deterioration	−0.01	−0.12, 0.10	−0.12*	−0.23, −0.01	−0.10	−0.38, 0.19	−0.35*	−0.64, −0.07
Active living potential	−0.05	−0.16, 0.06	0.03	−0.07, 0.14	−0.07	−0.35, 0.21	0.13	−0.15, 0.41
Service-related characteristics								
No. of supermarkets	−0.22**	−0.32, −0.11	−0.07	−0.17, 0.04	−0.58**	−0.85, −0.30	−0.25	−0.53, 0.04
No. of shops selling fruits/vegetables	−0.26**	−0.37, −0.16	−0.15**	−0.25, −0.04	−0.65**	−0.93, −0.38	−0.40**	−0.68, −0.12
Total no. of restaurants	−0.30**	−0.41, −0.20	−0.14**	−0.25, −0.03	−0.70**	−0.98, −0.42	−0.35*	−0.64, −0.05
No. of traditional restaurants	−0.31**	−0.41, −0.20	−0.14*	−0.25, −0.02	−0.71**	−0.99, −0.44	−0.34*	−0.63, −0.04
No. of fast-food restaurants	−0.27**	−0.38, −0.17	−0.14*	−0.24, −0.03	−0.65**	−0.92, −0.37	−0.35*	−0.63, −0.07
Proportion of fast-food restaurants	0.18**	0.08, 0.29	0.03	−0.08, 0.14	0.35*	0.06, 0.63	−0.03	−0.32, 0.27
No. of sports facilities	−0.01	−0.11, 0.10	0.06	−0.04, 0.16	0.001	−0.27, 0.27	0.16	−0.11, 0.43
No. of transportation lines	−0.17**	−0.27, −0.07	−0.07	−0.17, 0.03	−0.38**	−0.64, −0.11	−0.15	−0.42, 0.12
No. of basic services	−0.32**	−0.43, −0.21	−0.13*	−0.25, −0.02	−0.75**	−1.02, −0.47	−0.34*	−0.65, −0.04
No. of health-care resources	−0.30**	−0.41, −0.20	−0.08	−0.20, 0.04	−0.74**	−1.02, −0.46	−0.27	−0.58, 0.05
Social-interactional characteristics								
Deteriorated social interactions	0.09	−0.02, 0.20	−0.09	−0.20, 0.02	0.15	−0.13, 0.44	−0.26	−0.56, 0.04
Social cohesion	0.08	−0.03, 0.18	0.09	−0.01, 0.20	0.06	−0.22, 0.34	0.09	−0.18, 0.36
Collective feeling of insecurity	0.18**	0.07, 0.29	−0.07	−0.19, 0.05	0.37*	0.08, 0.66	−0.20	−0.53, 0.12

Abbreviation: CI, confidence interval.

* $P < 0.05$; ** $P < 0.01$.

[a] Weight (kg)/height (m)^2.

[b] In models 1 and 3, results were adjusted for age, sex, Human Development Index of the participant's country of birth, and individual and maternal educational levels.

[c] In models 2 and 4, results were further adjusted for neighborhood educational level and its interaction with sex.

Table 3. Correlations Between Environmental Variables Associated With Body Mass Index or Waist Circumference ($n = 7{,}234$), Paris, France, 2007–2008

Variable	Population Density	Proportion of Built Surface	Level of Neighborhood Physical Deterioration	No. of Shops Selling Fruits/Vegetables	Total No. of Restaurants	No. of Traditional Restaurants	No. of Fast-Food Restaurants	No. of Basic Services
Population density		0.72	0.43	0.67	0.66	0.66	0.63	0.68
Proportion of built surface			0.42	0.73	0.86	0.87	0.82	0.89
Level of neighborhood physical deterioration				0.44	0.40	0.39	0.32	0.36
No. of shops selling fruits/vegetables					0.81	0.80	0.82	0.82
Total no. of restaurants						1.00	0.98	0.97
No. of traditional restaurants							0.95	0.97
No. of fast-food restaurants								0.94
No. of basic services								

pertain to the large number of environmental variables included in the analysis, to the fact that most of the neighborhood variables were measured within areas centered on participants' residences, to the investigation of modification and mediation effects for the associations between environmental variables and BMI/waist circumference, and to the careful attempt to examine whether or not it is possible to disentangle the relations of the different environmental characteristics with weight status and abdominal fat. However, a major study limitation is that the RECORD population is not strictly representative of the Paris metropolitan region, even if we a priori selected a panel of municipalities from the region to ensure the presence of people from all socioeconomic backgrounds (neighborhood-related selective participation in the sample has been investigated elsewhere (11)). Another limitation is that the results were based on cross-sectional data and were therefore susceptible to reverse causation (weight status' influencing residential migration to specific neighborhoods) or confounding biases related to selective migration processes (a common characteristic's influencing both residential neighborhood choice and obesity risk), which might have distorted our estimates (7, 11). It is important to emphasize, however, that it would be similarly difficult with longitudinal data to separate the associations of different correlated variables of the physical and service environment with changes in weight status over time.

Study findings and methodological implications

The most consistent association at the neighborhood level was the one between neighborhood education and body weight and abdominal fat. Independently of individual education, the educational level and related values of neighborhood residents may influence personal attitudes and behavior related to weight gain and control. More indirect effects of neighborhood education include its causal influence on several aspects of the neighborhood (e.g., maintenance of the neighborhood, provision of health-enhancing services).

Only a few environmental variables were associated with BMI/waist circumference after adjustment for individual and neighborhood socioeconomic level (and interaction of the latter with sex), in coherence with other recent studies (25). Neighborhood socioeconomic characteristics have been suggested as possible confounders of the relations between physical environment characteristics/service availability and obesity (1, 7, 11, 26). Had neighborhood socioeconomic status been removed from our regression models as is commonly done, more environmental factors would have been (spuriously) associated with BMI/waist circumference.

The crucial issue of whether or not it is possible to separate the effects of highly correlated factors pertaining to the physical or service environment has not been addressed in previous literature. To our knowledge, our study was the first in our domain to focus on the ability of statistical analyses of observational data to disentangle the effects of correlated environmental exposures. A concern is that regression models would provide estimates of environmental effects adjusted for each other even if these associations were not grounded on sufficient data to allow the effects to be separated (27, 28).

In addition to the common strategy based on multiple adjustment, our analyses performed within pairs of individuals similarly exposed to a given environmental dimension (neighborhood characteristic-matched analyses) allowed us to examine whether neighborhood characteristics remained associated with BMI/waist circumference after fixing values of another environmental characteristic. This matching technique is a methodological innovation that allowed us to perform analyses among participants who were exchangeable between neighborhood exposure groups on the basis of a specific environmental exposure. In our analyses, matching was not used in itself as an alternative to adjustment. In

Table 4. Associations (With 95% Credible Intervals[a]) Between Environmental Variables and Waist Circumference, Determined Using the Neighborhood Characteristic-Matching Technique[b], Paris, France, 2007–2008

Variable of Interest	No. of Shops Selling Fruits/Vegetables[c]				Level of Neighborhood Physical Deterioration[d]				Proportion of Built Surface[e]			
	Difference in Exposure	2.5th–97.5th Percentiles	β	95% CrI	Difference in Exposure	2.5th–97.5th Percentiles	β	95% CrI	Difference in Exposure	2.5th–97.5th Percentiles	β	95% CrI
Pairs based on a random variable	3.11	3.04, 3.17	−0.46	−0.82, −0.13	1.24	1.22, 1.26	−0.38	−0.77, −0.06	0.15	0.14, 0.15	−0.32	−0.70, −0.01
Pairs based on the least correlated variable unassociated with waist circumference	2.96	2.90, 3.02	−0.43	−0.75, −0.08	1.20	1.17, 1.22	−0.31	−0.68, 0.03	0.11	0.11, 0.11	−0.40	−0.73, −0.03
Pairs based on the least correlated variable associated with waist circumference	2.03	1.97, 2.09	−0.24	−0.55, 0.07	1.03	1.01, 1.05	−0.20	−0.53, 0.14	0.09	0.09, 0.09	−0.26	−0.62, 0.08
Pairs based on the variable with the median correlation level associated with waist circumference	1.73	1.68, 1.77	−0.41	−0.75, −0.06	1.07	1.05, 1.08	−0.33	−0.66, −0.01	0.06	0.06, 0.06	0.07	−0.28, 0.42
Pairs based on the most correlated variable associated with waist circumference	1.63	1.59, 1.67	−0.21	−0.53, 0.13	1.07	1.05, 1.09	−0.19	−0.51, 0.18	0.05	0.05, 0.05	−0.01	−0.35, 0.34

Abbreviation: CrI, credible interval.

[a] The credible intervals for the associations correspond to the 2.5th and 97.5th percentiles of the coefficients estimated in the different samples.

[b] In all models, results were adjusted for a composite score including individual and neighborhood socioeconomic variables.

[c] The variables used for matching were: for the least correlated variable unassociated with waist circumference, the number of monuments ($r = -0.06$); for the least correlated variable associated with waist circumference, the level of neighborhood physical deterioration ($r = 0.44$); for the variable with the median correlation level associated with waist circumference, the number of traditional restaurants ($r = 0.80$); and for the most correlated variable associated with waist circumference, the number of fast-food restaurants ($r = 0.82$).

[d] The variables used for matching were: for the least correlated variable unassociated with waist circumference, the proportion of the area covered by water ($r = -0.10$); for the least correlated variable associated with waist circumference, the number of fast-food restaurants ($r = 0.32$); for the variable with the median correlation level associated with waist circumference, the total number of restaurants ($r = 0.40$); and for the most correlated variable associated with waist circumference, the number of shops selling fruits/vegetables ($r = 0.44$).

[e] The variables used for matching were: for the least correlated variable unassociated with waist circumference, the level of insecurity ($r = 0.04$); for the least correlated variable associated with waist circumference, the level of neighborhood physical deterioration ($r = -0.42$); for the variable with the median correlation level associated with waist circumference, the number of fast-food restaurants ($r = 0.82$); and for the most correlated variable associated with waist circumference, the number of basic services ($r = 0.89$).

the literature, matching is typically employed to reduce model dependence and estimate associations in a more empirical way than would be necessary without matching (29–32). In line with this practice, matching was employed as a diagnostic tool to verify whether the effects of environmental variables that are highly correlated with each other could be disentangled.

The results from matched analyses almost systematically concurred with those from the multiple adjustment approach, with few exceptions. As an exception, the number of shops selling fruits/vegetables remained associated with waist circumference after adjustment for population density (beyond individual/neighborhood socioeconomic covariates) in a multiple adjustment model; however, the effect of the number of shops selling fruits/vegetables on waist circumference among pairs of participants living in neighborhoods with similar population densities disappeared (result not shown). A possible explanation for this divergence is related to the excessive extrapolations made by multiple adjustment models. The absence of association with our neighborhood characteristic-matched design may have been due to the fact that this approach suppresses such excessive extrapolations: Within pairs of participants similarly exposed to population density, the contrast between participants in the availability of shops selling fruits/vegetables was not sufficient to demonstrate any association with BMI/waist circumference.

Regarding the interpretation of our neighborhood characteristic-matched analyses, given 2 environmental variables X_1 and X_2 that are associated with the outcome when examined separately, it is always possible to find a variable X_2 that is sufficiently correlated and redundant with X_1 that no effect of X_1 can be detected within pairs of X_2. Therefore, an inability to document an effect of X_1 within pairs of X_2 should be interpreted not as evidence of the absence of an effect of X_1 but rather as evidence of the impossibility of separating its effects from those of X_2. Overall, our neighborhood characteristic-matched approach is proposed as a diagnostic tool with which to assess the separability of associations with environmental characteristics that epidemiologists are willing to introduce simultaneously in regression models.

Our findings based on matched analyses indicated that among variables related to the density of physical features of the environment and services, we could not identify a specific environmental variable that remained particularly associated with BMI/waist circumference after adjustment for the other environmental variables, even if the different environmental factors examined were hypothesized to operate through different mechanisms.

Our analyses indicated that the number of shops selling fruits/vegetables and the proportion of built surface were perhaps more particularly associated with BMI/waist circumference. On the one hand, the negative associations observed between the number of shops selling fruits/vegetables and weight status/abdominal fat might be explained by the fact that the density of such shops serves as a proxy of the healthiness of the food environment or the healthiness of the environment in general. On the other hand, the negative relations between the proportion of built surface and BMI/waist circumference may be attributable to the demonstrated positive effect of high densities on utilitarian walking (33). However, our analyses, which could not disentangle the effects of these different environmental exposures, do not provide firm support for these hypotheses.

For future research, cluster analyses that integrate several environmental characteristics to create neighborhood typologies (34) might be a promising perspective for dealing with the methodological challenge of investigating the influence of highly correlated environmental variables on weight status/abdominal fat.

Preliminary analyses using a data reduction method (principal component analysis) for neighborhood variables provided results comparable to those reported here. For example, applying data reduction to variables related to street connectivity, the 2 axes created (which accounted for more than 70% of the variance) were not associated with BMI/waist circumference, in accordance with the results observed when street connectivity variables were analyzed separately. Applying the same method to food environment variables, the first axis was negatively associated BMI/waist circumference, in agreement with the findings documented here (Web Table 9). It is important to note, however, that data reduction for neighborhood variables does not allow for identification of any specific environmental effects on weight status that would be useful for the elaboration of definite public health intervention strategies.

Regarding neighborhood social interactions, our hypothesis was that they may encourage or discourage individuals to practice leisure-time physical activities or to walk in their neighborhoods. Contrary to this hypothesis, we did not find associations between any neighborhood social-interactional characteristic and BMI/waist circumference. If it is not attributable to measurement error in the neighborhood ecometric variables, the absence of associations may be due to the fact that decreased leisure-time activity in neighborhoods with deteriorated social interactions is counterbalanced by increased transportation-related activity in these neighborhoods.

In conclusion, physical and service-related neighborhood characteristics reflecting higher densities were negatively associated with BMI/waist circumference after adjustment for individual/neighborhood socioeconomic covariates. However, because these environmental characteristics are highly correlated with each other, it is impossible to firmly conclude that one of them has a truly independent causal effect on BMI/waist circumference. On the basis of our data, it is reasonable to conclude only that there is an overall protective effect of high densities. Assessing the possibility of disentangling the associations of different environmental characteristics with health, for which distinct causal mechanisms are hypothesized, is an important methodological challenge for future studies aiming to provide specific guidance in the fight against the obesity epidemic.

ACKNOWLEDGMENTS

Author affiliations: Research Unit in Epidemiology, Information Systems, and Modeling, INSERM U707, Paris, France (Cinira Leal, Basile Chaix); Université Pierre et Marie Curie–Paris 6, UMRS 707, Paris, France (Cinira Leal, Basile Chaix); EHESP School of Public Health, Rennes, France (Cinira

Leal); and Centre d'Investigations Préventives et Cliniques, Paris, France (Kathy Bean, Frédérique Thomas).

All of the authors are members of the RECORD Study Project (Paris, France).

The RECORD Study Project received financial support from the French Institute for Public Health Research (Institut de Recherche en Santé Publique (IReSP)); the French National Institute for Prevention and Health Education (Institut National de Prévention et d'Education pour la Santé (INPES)) (Prevention Program 074/07-DAS in 2007 and financial support in 2010–2011); the French National Institute of Public Health Surveillance (Institut de Veille Sanitaire (InVS)) (Territory and Health Program); the French Ministry of Research and Ministry of Health (an Epidemiologic Cohorts Grant in 2008); the French National Health Insurance Office for Salaried Workers (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAM-TS)); the French National Research Agency (Agence Nationale de la Recherche (ANR)) (Health-Environment Program grant 00153 05 in 2005); the Ile-de-France Regional Health Agency (Agence Régionale de Santé d'Île-de-France (ARS)); the City of Paris (Ville de Paris); and the Ile-de-France Youth, Sports, and Social Cohesion Regional Directorate (Direction Régionale de la Jeunesse et des Sports et de la Cohésion Sociale (DRJSCS)).

The authors thank Dr. Alfred Spira, head of the French Institute for Public Health Research (IReSP), for his advice and support. They are grateful to the French National Institute for Prevention and Health Education (INPES) (and Pierre Arwidson) for its continued support. They also thank Dr. Danièle Mischlich from the Ile-de-France Regional Health Agency (ARS) and Nathalie Catajar and Muriel Hirt from the Ile-de-France Youth, Sports, and Social Cohesion Regional Directorate (DRJSCS) for their support. The authors are grateful to the French National Institute of Statistics and Economic Studies (Institut National de la Statistique et des Études Économiques (INSEE)), which provided support for the geocoding of the RECORD participants and allowed access to relevant geographic data (with special thanks to Pascale Breuil and Jean-Luc Lipatz). The authors thank GeoConcept SA (Bagneux, France) for allowing them access to the Universal Geocoder software and Dr. Alain Weill from the French National Health Insurance Office for Salaried Workers (CNAM-TS) for his support in merging the health-care consumption data from the French National Health Insurance Information System (SNIIR-AM) with the RECORD database. Regarding the data used in the present analysis, the authors are grateful to Paris-Notaires and the National Geographic Institute (Institut Géographique National). They also thank the French National Health Insurance Office for Salaried Workers (CNAM-TS) and the Paris Health Insurance Office for Salaried Workers (Caisse Primaire d'Assurance Maladie, Paris (CPAM-P)) for helping to make this study possible.

These findings were previously presented at the 19th IEA World Congress of Epidemiology, Edinburgh, Scotland, August 7–11, 2011 (35) and the Third North American Congress of Epidemiology, Montreal, Quebec, Canada, June 21–24, 2011 (36).

Conflict of interest: none declared.

REFERENCES

1. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. *Obes Rev*. 2011;12(3):217–230.
2. Black JL, Macinko J, Dixon LB, et al. Neighborhoods and obesity in New York City. *Health Place*. 2010;16(3):489–499.
3. Giles-Corti B, Macintyre S, Clarkson JP, et al. Environmental and lifestyle factors associated with overweight and obesity in Perth, Australia. *Am J Health Promot*. 2003;18(1):93–102.
4. Rutt CD, Coleman KJ. Examining the relationships among built environment, physical activity, and body mass index in El Paso, TX. *Prev Med*. 2005;40(6):831–841.
5. Nelson MC, Gordon-Larsen P, Song Y, et al. Built and social environments associations with adolescent overweight and activity. *Am J Prev Med*. 2006;31(2):109–117.
6. Grafova IB. Overweight children: assessing the contribution of the built environment. *Prev Med*. 2008;47(3):304–308.
7. Chaix B. Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. *Annu Rev Public Health*. 2009;30:81–105.
8. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med*. 2004;58(10):1929–1952.
9. Leal C, Bean K, Thomas F, et al. Are associations between neighborhood socioeconomic characteristics and body mass index or waist circumference based on model extrapolations? *Epidemiology*. 2011;22(5):694–703.
10. Chaix B, Bean K, Leal C, et al. Individual/neighborhood social factors and blood pressure in the RECORD Cohort Study: which risk factors explain the associations? *Hypertension*. 2010;55(3):769–775.
11. Chaix B, Billaudeau N, Thomas F, et al. Neighborhood effects on health: correcting bias from neighborhood effects on participation. *Epidemiology*. 2011;22(1):18–26.
12. Chaix B, Jouven X, Thomas F, et al. Why socially deprived populations have a faster resting heart rate: impact of behaviour, life course anthropometry, and biology—the RECORD Cohort Study. *Soc Sci Med*. 2011;73(10):1543–1550.
13. Chaix B, Kestens Y, Bean K, et al. Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases—the RECORD Cohort Study [published online ahead of print July 7, 2011]. *Int J Epidemiol*. (doi:10.1093/ije/dyr107).
14. Havard S, Reich BJ, Bean K, et al. Social inequalities in residential exposure to road traffic noise: an environmental justice analysis based on the RECORD Cohort Study. *Occup Environ Med*. 2011;68(5):366–374.
15. French National Institute of Statistics and Economic Studies. *IRIS—Definition*. Paris, France: National Institute of Statistics and Economic Studies; 2011. (http://www.insee.fr/en/methodes/default.asp?page=definitions/iris.htm). (Accessed April 4, 2012).
16. Thomas F, Bean K, Pannier B, et al. Cardiovascular mortality in overweight subjects: the key role of associated risk factors. *Hypertension*. 2005;46(4):654–659.
17. World Health Organization. *Obesity: Preventing and Managing the Global Epidemic: Report of a WHO Consultation*. Geneva, Switzerland: World Health Organization; 2000.
18. Benetos A, Thomas F, Pannier B, et al. All-cause and cardiovascular mortality using the different definitions of metabolic syndrome. *Am J Cardiol*. 2008;102(2):188–191.
19. Beckman A, Hakansson A, Rastam L, et al. The role country of birth plays in receiving disability pensions in relation to patterns of health care utilisation and socioeconomic differences:

a multilevel analysis of Malmo, Sweden. *BMC Public Health*. 2006;6:71. (doi:10.1186/1471-2458-6-71).
20. United Nations Development Programme. *Human Development Report 2007/2008*. New York, NY: United Nations; 2008.
21. Mujahid MS, Diez Roux AV, Morenoff JD, et al. Neighborhood characteristics and hypertension. *Epidemiology*. 2008;19(4): 590–598.
22. Breslow NE, Day NE, eds. *Statistical Methods in Cancer Research. Vol 1. The Analysis of Case-Control Studies.* (IARC Scientific Publication no. 32). Lyon, France: International Agency for Research on Cancer; 1980.
23. Scott MM, Dubowitz T, Cohen DA. Regional differences in walking frequency and BMI: what role does the built environment play for blacks and whites? *Health Place*. 2009;15(3):882–887.
24. Grafova IB, Freedman VA, Kumar R, et al. Neighborhoods and obesity in later life. *Am J Public Health*. 2008;98(11):2065–2071.
25. Timperio A, Jeffery RW, Crawford D, et al. Neighbourhood physical activity environments and adiposity in children and mothers: a three-year longitudinal study. *Int J Behav Nutr Phys Act*. 2010;7:18. (doi:10.1186/1479-5868-7-18).
26. Chaix B, Leal C, Evans D. Neighborhood-level confounding in epidemiologic studies: unavoidable challenges, uncertain solutions. *Epidemiology*. 2010;21(1):124–127.
27. Oakes JM. Commentary: advancing neighbourhood-effects research—selection, inferential support, and structural confounding. *Int J Epidemiol*. 2006;35(3):643–647.
28. Messer LC, Oakes JM, Mason S. Effects of socioeconomic and racial residential segregation on preterm birth: a cautionary tale of structural confounding. *Am J Epidemiol*. 2010;171(6):664–673.
29. Costanza M. Matching. *Prev Med*. 1995;24(5):425–433.
30. Greenland S, Morgenstern H. Matching and efficiency in cohort studies. *Am J Epidemiol*. 1990;131(1):151–159.
31. Ho DE, Imai K, King G, et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal*. 2007;15(3):199–236.
32. Karon JM, Kupper LL. In defense of matching. *Am J Epidemiol*. 1982;116(5):852–866.
33. Owen N, Humpel N, Leslie E, et al. Understanding environmental influences on walking; review and research agenda. *Am J Prev Med*. 2004;27(1):67–76.
34. Gershoff ET, Pedersen S, Lawrence Aber J. Creating neighborhood typologies of GIS-based data in the absence of neighborhood-based sampling: a factor and cluster analytic strategy. *J Prev Interv Community*. 2009;37(1): 35–47.
35. Leal C, Bean K, Chaix B. Neighborhood sociodemographic, physical, service-related, and social-interactional characteristics and BMI or waist circumference in the RECORD Study: evaluation of the separability of associations with a neighborhood characteristic matching technique [abstract]. *J Epidemiol Community Health*. 2011;65(suppl 1):P1-211.
36. Leal C, Bean K. Chaix, B. Neighborhood sociodemographic, physical, service-related, and social-interactional characteristics and BMI or waist circumference in the RECORD Study: evaluation of the separability of associations with a neighborhood characteristic matching technique [abstract]. *Am J Epidemiol*. 2011;173(suppl 11): 188-S.