eAppendix

The "residential" effect fallacy in neighborhood and health studies: formal definition, empirical identification, and correction

eAppendix 1: Assumptions required for estimating a valid causal residential intervention effect

In this paper, we aimed to estimate the causal effect of a hypothetical intervention raising the residential number of services on the probability that a trip is walked. There are potential concerns about drawing causal inferences based on our estimate, even after the correction of the “residential” effect fallacy. Below, we provide a tentative list of assumptions that would need to be met so that the estimate of the residential intervention effect corrected from the “residential” effect fallacy represents the causal intervention effect.

1) Estimation of the regression model used for the calculation

- We assume no systematic measurement error in the probability that a trip is walked according to the exposure variable (the residential accessibility to services).
- We assume no systematic measurement error in the number of services across the different types of transport modes (e.g., differential misclassification of the count of services according to the type of place, itself associated with a particular mode use).
- We assume no systematic nondifferential measurement error in the number of services (i.e., a uniform decrease or increase in the count of services), which would bias the intervention effect estimate due to the nonlinear relationship between the number of services and walking.
- Also we make the assumption of no random measurement error in the number of services (i.e., the exposure) which, as a regression dilution bias, would result in an attenuation towards the null of the association between the exposure to services and the probability that a trip is walked.
- We assume no random measurement error in the transport outcome, which would increase the standard error of the estimated association.

- The study examined the association between the number of services in a 1 km radius buffer around the residence and walking. This was not considered to be an assumption of the study (that would be related to the expected spatial scale of this causal determinant of walking), but to be a logical consequence of the hypothetical choice of policymakers to intervene within 1 km of the residence of a number of residents. However, when we attempted to estimate the causal effect of the number of services around the trip origins and destinations, we also chose a 1 km radius for the trip origin/destination buffers for the sake of coherence. This was related to the assumption that the causal effect that we attempted to take into account (as a potential confounder of the intervention effect) was best captured with a radius of 1 km. It might be that the optimal radius to capture this effect is shorter. However, this is likely not of critical importance due to the strong autocorrelation in the number of services that should minimize the impact of small differences in the radius to define the buffers.
- We assume that there is no selection bias in the sample used for estimating the association of interest. Specifically, we assume that the participation in the sample is not influenced by both the exposure (or a determinant of the exposure) and the outcome (or a determinant of the outcome).
- We assume that there is no additional unmeasured trip-level, individual-level, or environmental confounders in the relationship between the residential number of services and the probability of walking, after our correction of the "residential" effect fallacy and after accounting for neighborhood preferences at the time of moving in one's current neighborhood (neighborhood selection factors) and the other covariates. We assume that residents of neighborhoods with different levels of spatial accessibility to services are otherwise exchangeable conditional on adjustment factors.

- We assume that the positivity assumption (or experimental treatment assignment assumption) is held. We assume that there are both exposed and unexposed individuals at every combination of levels of the observed confounders in the population of interest. Under this assumption, our regression modeling is not based on excessive extrapolations.
- The random effect linear probability model including linear and quadratic terms for the service accessibility variables is assumed to be not misspecified. The following standard assumptions are hypothesized to apply: homoscedasticity; normal distribution of residuals at the trip level and at the individual level; and absence of correlation between the exposure of interest or other covariates and the level-2 individual-level random intercept.

2) Calculation of the residential intervention effect estimate based on the regression estimates

- The method acknowledges that the residential intervention may also influence mode choice in trips far from the residence through the influence that the residential neighborhood has on the overall choice of mode (e.g., buying a car or a public transport pass). This is not an assumption that we force into the model, but a potential mechanism that is allowed for in the model and that is incorporated into the calculation only if estimated to be at play.
- We assume that there is no relevant time varying environmental variable correlated with the implementation of the intervention of interest. For example, if interventions aiming at increasing the number of services were systematically implemented simultaneously with urban design changes (i.e., if the two were consubstantial), the true, pure intervention effect would not correspond to the one estimated in this study.
- Provided that the assumptions listed above are satisfied, the regression estimates reflect average causal effects of the number of services in the residential, trip level,

and trip origin neighborhoods on walking. However, these causal effects are averages of heterogeneous effects across different population subgroups. Our calculation assumes that the various neighborhoods in our study territory contain a comparable mix of heterogeneous residents leading to a uniform intervention effect on walking.

- We make the assumption that the intervention (increase in the number of services) is evenly distributed over the 1 km residential buffer; thus the portion of the residential intervention that affects a participant located in a nonresidential place depends on the fraction of the residential buffer that is included in this nonresidential buffer.
- The calculation of the residential intervention effect estimate was based, among other, on the estimated associations between the trip origin/trip destination numbers of services and walking. These were average effects, i.e., the associations between the trip origin or trip destination number of services and walking were estimated accounting for all trip origins and destinations in the database. However, this estimate was used to calculate the intervention effect estimate for trip origins and destinations close to the residence (as their buffer had to overlap the residential buffer). Thus we hypothesize that the estimated associations between the trip origin/destination numbers of services and walking were appropriate to estimate the intervention effect for trip origins/destinations overlapping the residential buffer.
- In the calculation of the overall intervention effect estimate based on the regression estimates, an absence of residential migration between the pre-intervention and post-intervention states is assumed. Also, we assume no change in the places visited for this calculation. Regarding the notion of post-intervention disequilibria, our calculation is based on the stable unit treatment assumption, i.e., the intervention implemented for one person does not affect the level of intervention for another person. While this assumption was made in the calculations in the main article, the last section of

eAppendix 7 provides calculations relaxing this assumption of an absence of interferences between participants.

- It should be noted that the present work took into account the “residential” effect fallacy only in the association of services with the probability that a trip is walked. However, our calculations did not account for the fact that an increase in the residential number of services may also increase the number of trips, and that a similar “residential” effect fallacy may bias this association. Correcting for the “residential” effect fallacy in the number of trips would be more complicated.
- Moreover, an intervention on services in the residential neighborhood may affect not only the number of trips but also the length of trips and how trip origins / destinations are close or not from the residence.[1] Because our regression models were purposely not adjusted for distance, the association estimated between services and walking may operate through a switch of mode in a given trip with fixed start and end points but also through a switch from a longer motorized trip to a shorter walking trip. However, our estimation of the intervention effect did not allow for a change in the extent to which the origin and destination buffers of each trip overlapped the residential neighborhood. This is a limitation since an intervention raising the residential number of services may increase the percentage of trips whose origin and/or destination buffers overlap the residential neighborhood. Thus our illustrative study was to some extent grounded on the simplifying hypothesis that participants would anyway visit all the places visited during the follow-up even if the residential number of services was increased.

eAppendix 2: Review of previous literature

There are very few articles in the whole literature on neighborhood / environmental effects on health that relate at least approximately to the topic of the "residential" effect fallacy bias that we describe.

Inagami et al.

In a well-known article,[2] Inagami and colleagues found that a low socioeconomic status of the residential neighborhood was associated with a worst self-rated health; that exposure to a low socioeconomic status in nonresidential neighborhoods was associated with a worst self-rated health (only in some of the models); and that adding the nonresidential term to the model increased the residential effect estimate. Thus the standard calculation of the residential effect estimate led to an underestimation of the supposedly true residential effect, i.e., the confounding bias was in the form of a suppression effect rather than an amplification effect as in our article.

However, the conceptual and analytical framework of Inagami and colleagues was entirely different from ours. In the work of Inagami, the nonresidential exposure variable was calculated as the difference between the nonresidential exposure and the residential exposure, thus expressed as nonresidential relative disadvantage. First, it should be noted that such a "relative exposure effect" is an original effect in itself, distinct from the effect of the absolute level of nonresidential exposure we are interested in. Indeed, such a relative exposure specification likely captures a different effect operating through distinct mechanisms, such as the influence of cognitive processes of comparison of residential and nonresidential neighborhoods. Second, Inagami and colleagues implicitly recognize that there was a positive correlation between the absolute socioeconomic status in the residential and nonresidential neighborhoods (Data and methods, Measures, Operationalizing non-residential neighborhood exposure, fourth paragraph). However, the suppression effect of confounding documented by Inagami (rather than amplification effect in our case) implies a negative correlation between the residential socioeconomic status and nonresidential socioeconomic exposure variable (as opposed to a positive correlation in our "residential" effect fallacy application). This negative correlation stems from the relative definition of the nonresidential exposure that was used (difference with the residential exposure) and may be attributable to a "regression to the mean", i.e., to the fact that participants with a particularly high residential socioeconomic status will often have nonresidential places with a comparably lower socioeconomic status (thus a negative relative exposure) while participants with a particularly low residential socioeconomic status will often have nonresidential places with a comparably higher socioeconomic status.

Overall, the study by Inagami and colleagues investigates how a nonresidential effect defined in a different way than the residential effect (relative exposure) negatively confounds the residential effect of interest. Differently, our study investigates how a nonresidential effect defined in a similar way than the residential effect positively confounds the residential effect. Thus our study is the first to address the "residential" effect fallacy bias described here.

Sharp et al.

Another article by Sharp and colleagues[3] based on longitudinal data from the same cohort than in the Inagami article examined the extent to which controlling for the exposure to nonresidential neighborhood disadvantage affected the relationship between residential neighborhood disadvantage and self-rated health. As opposed to Inagami et al., nonresidential disadvantage was assessed in absolute rather than relative terms. The work by Sharp and colleagues reported that the residential effect estimate was slightly attenuated when the model was adjusted for nonresidential exposures. This reduction of the strength of the association was attributed to a mediation of the residential association by the nonresidential term, although no concrete description of the causal mechanism involved in this supposed mediation and no explanation based on directed acyclic graphs of why this should be mediation rather than confounding were provided. Even if the observed reduction in the residential association when controlling for the nonresidential term is coherent with our own findings, the conceptual framework that we develop is substantially different. Although we acknowledge that nonresidential exposures mediate to some extent the estimated residential neighborhood-walking association (the residential environment influences the transport modes used, which influence the types of places visited), we emphasize that most importantly, nonresidential exposures confound the residential neighborhood-walking association.

eAppendix 3: Services accounted for in our spatial accessibility variable

Consistent with previous literature,[4-6] the spatial accessibility to destinations within a walking distance has been found to be a major determinant of both transport and recreational walking in our RECORD Study.[7-9] Services, especially because a large spectrum of them were accounted for in our study, represent a large share of the potential destinations of participants. Thus, our variable is expected to capture in a relatively reliable way a factor that has a direct causal effect on the likelihood of transport walking.

We report below the list of services accounted for in our variable of spatial accessibility to services. Although neither sport facilities nor parks were taken into account in the list of services that was analyzed, it can be seen that a large fraction of the services available were included in this study.

List of services analyzed:
A101 – Police
A102 – Treasury
A103 – National Employment Agency
A104 – Gendarmerie
A203 – Bank
A206 – Post office
A207 – Package delivery point
A208 – Municipality post office
A301 – Car repair
A302 – Automobile technical inspection service
A303 – Car rental
A401 – Mason
A402 – Plasterer, painter
A403 – Wood worker, carpenter, locksmith
A404 – Plumber, roofer, heating engineer
A405 – Electrician
A406 – Construction company
A501 – Hairdresser
A502 – Veterinarian
A504 – Restaurant
A505 – Real estate agency
A506 – Laundry
A507 – Beauty care

B101 – Hypermarket
B102 – Supermarket
B103 – Large do-it-yourself store
B201 – Minimarket
B202 – Grocery
B203 – Bakery
B204 – Butcher / delicatessen shop
B205 – Frozen food store
B206 – Fish market
B301 – Bookshop, stationery store
B302 – Clothing store
B303 – Home equipment store
B304 – Shoe store
B305 – Home appliance store
B306 – Furniture store
B307 – Sports store
B308 – Wallpaper and wall covering store
B309 – Drugstore, hardware, handiwork
B310 – Perfumery
B311 – Watch and jewellery
B312 – Florist
B313 – Optical store
D108 – Health center
D201 – General practitioner
D202 – Cardiologist
D203 – Dermatology and venereology
D204 – Medical gynecology
D205 – Gynecology obstetrics
D206 – Gastroenterology
D207 – Psychiatry
D208 – Ophthalmology
D209 – Otorhinolaryngology
D210 – Pediatrics
D211 – Pulmonology
D212 – Diagnostic radiology and medical imaging
D213 – Stomatology
D221 – Dental surgeon
D231 – Midwife
D232 – Nurse
D233 – Masseur physiotherapist
D235 – Speech therapist
D236 – Orthoptist
D237 – Chiropodist
D238 – Audioprosthesist
D239 – Occupational therapist
D240 – Psychomotrician
D241 – Medical radiology operator
D301 – Pharmacy
D302 – Medical analysis laboratory
F301 – Cinema

F302 – Theater

eAppendix 4: Regression models estimated to determine the naïve and corrected intervention effect estimates

eAppendix Table 1. Regression models estimated to determine the naïve and corrected residential intervention effect estimates

	Model for the naïve estimate		Model for the corrected estimate	
Individual factors				
Age (vs. 35–49)				
50–64	+0.02	–0.08, +0.12	+0.02	–0.06, +0.09
65 and over	–0.02	–0.16, +0.13	–0.02	–0.12, +0.09
Male (vs. female)	–0.03	–0.11, +0.05	–0.02	–0.08, +0.04
Living alone (vs. as a couple)	–0.02	–0.11, +0.07	–0.05	–0.11, +0.02
Education (vs. ≤ low secondary)				
Upper secondary, low tertiary	–0.02	–0.12, +0.07	–0.02	–0.09, +0.05
Intermediate tertiary	+0.09	–0.03, +0.21	+0.04	–0.05, +0.12
Upper tertiary	–0.07	–0.17, +0.03	–0.06	–0.13, +0.02
Employment status (vs. stable job)				
Precarious job	–0.07	–0.27, +0.14	–0.02	–0.18, +0.13
Unemployment	+0.16	–0.09, +0.41	+0.14	–0.04, +0.33
Household income per consumption unit (vs. ≤1285 €)				
>1285 – ≤2200 €	+0.00	–0.08, +0.09	+0.01	–0.06, +0.08
>2200 €	–0.02	–0.11, +0.08	+0.00	–0.07, +0.07
Services as a neighborhood selection factor (vs. low)				
Intermediate	+0.05	–0.07, +0.15	+0.06	–0.02, +0.14
High	+0.05	–0.06, +0.16	+0.06	–0.03, +0.14
Public transport as a neighborhood selection factor (vs. low)				
Intermediate	–0.03	–0.15, +0.10	–0.01	–0.10, +0.08
High	–0.01,	–0.11, +0.09	–0.02	–0.10, +0.06
Bike used in previous trips from home	–		–0.51	–0.59, –0.44
Personal vehicle used in previous trips from home (vs. no)				
In some of the trips	–		–0.32	–0.37, –0.27
In all trips	–		–0.56	–0.60, –0.53
Environmental factors				
Number of services, residential neighborhood				
Linear term	+0.21	+0.11, +0.32	+0.02	–0.07, +0.10
Quadratic term	–0.04	–0.07, –0.01	–0.01	–0.03, +0.02
Number of services, trip origin				
Linear term	–		+0.03	–0.02, +0.08
Quadratic term	–		–0.01	–0.02, +0.00
Number of services, trip destination				
Linear term	–		+0.10	+0.05, +0.15
Quadratic term	–		–0.01	–0.02, –0.00

eAppendix 5: A posteriori recalculation of the naïve biased residential intervention effect estimate

Methodology

To reach an analytical understanding of the genesis of the bias, we recalculated the naïve estimate of the intervention effect from a model accounting for the trip-level number of services. In our naïve model, influences on walking of the nonresidential places visited are spuriously incorporated in the residential neighborhood-walking association when the number of services at these visited places is similar to the residential number of services. Our aim was to mimic this process.

To do so, we had to determine whether the number of services in each nonresidential place visited was "similar" or not to the residential number of services. For the nonresidential place of a participant X, we constructed a database comprising this nonresidential place of X and all the nonresidential places of the other participants. For each of these nonresidential places in the database, we calculated the absolute value of the difference between the nonresidential number of services and the residential number of services of participant X. The number of services in the participant X's nonresidential place buffer was considered to be similar to participant X's residential number of services if the corresponding difference was below the first decile, the second decile, the third decile, the fourth decile, or the fifth decile of ranked differences (alternative definitions of "similarity").

In our recalculation of the naïve estimate, the number of services in a nonresidential place (even if not overlapping the residential neighborhood) was spuriously raised to the intervention target (200, 500, or 1000) if this nonresidential place was "similar" to the residential neighborhood in terms of services.

To recalculate the naïve intervention effect estimate, we re-estimated the regression model for walking with sociodemographic variables, neighborhood selection factors, and the residential, trip origin, and trip destination numbers of services as explanatory variables but without modes used in previous trips (as nonresidential effects incorporated in the naïve estimate are confounded by these modes in previous trips). Based on model coefficients, we calculated the predicted probability of entirely walking in each trip from all model covariates (including the residential, trip origin, and trip destination numbers of services, modified as explained above). This calculation was performed for the pre-intervention state and for each of the post-intervention scenarios (residential services raised to 200, 500, or 1000). For each of these four cases, we calculated the average probability of walking in a trip across all individuals and trips. The intervention effect estimate was computed for each intervention level, and for each cutoff to define whether a nonresidential place was similar to the residential neighborhood in terms of services, as the post-intervention average probability of walking minus the pre-intervention probability, only among participants who experimented the hypothetical intervention (i.e., with less than 200, 500, or 1000 services in their residential neighborhood).

Results

The model that was estimated is reported in eAppendix Table 2. When no adjustment was made for the modes in previous trips, both the trip origin and trip destination numbers of services were positively associated with the probability that a trip is walked (with quadratic effects).

The recalculation of the naïve biased estimate implied to spuriously integrate in the residential neighborhood-walking association the influence on walking of the nonresidential places visited if their number of services was “similar” to the residential number of services.

eAppendix Table 3 reports such recalculated naïve estimates for a low threshold (1st decile) to define such similarity (corresponding to a low level of bias) to a higher threshold (5th decile) to define such similarity (corresponding to a higher level of bias). The recalculated estimate was most comparable to the naïve estimate when the effect of nonresidential places was spuriously integrated in the residential neighborhood-walking association if the difference in services between the residential and nonresidential places was below the third or fourth decile of differences for all participants.

eAppendix Table 2. Regression model estimated to determine the recalculated residential intervention effect estimate

Individual factors		
Age (vs. 35–49)		
50–64	+0.03	–0.07, +0.13
65 and over	–0.01	–0.15, +0.13
Male (vs. female)	–0.03	–0.11, +0.05
Living alone (vs. as a couple)	–0.03	–0.11, +0.06
Education (vs. ≤ low secondary)		
Upper secondary, low tertiary	–0.03	–0.12, +0.06
Intermediate tertiary	+0.07	–0.04, +0.19
Upper tertiary	–0.07	–0.17, +0.02
Employment status (vs. stable job)		
Precarious job	–0.06	–0.27, +0.14
Unemployment	+0.20	–0.05, +0.44
Household income per consumption unit (vs. ≤1285 €)		
>1285 – ≤2200 €	+0.00	–0.09, +0.09
>2200 €	–0.01	–0.11, +0.08
Services as a neighborhood selection factor (vs. low)		
Intermediate	+0.03	–0.08, +0.14
High	+0.04	–0.07, +0.15
Public transport as a neighborhood selection factor (vs. low)		
Intermediate	–0.03	–0.15, +0.09
High	–0.02	–0.12, +0.08
Environmental factors		
Number of services, residential neighborhood		
Linear term	+0.04	–0.07, +0.15
Quadratic term	–0.01	–0.04, +0.02

Number of services, trip origin		
Linear term	+0.10	+0.05, +0.16
Quadratic term	–0.02	–0.03, –0.01
Number of services, trip destination		
Linear term	+0.12	+0.07, +0.17
Quadratic term	–0.02	–0.03, –0.01

eAppendix Table 3. Recalculation* of the naïve estimate of the hypothetical intervention effect of raising the number of services in the residential neighborhood to 200, 500, or 1000 for participants in the intervention groups (i.e., with a residential number of services below 200, 500, or 1000), according to the degree of similarity between the nonresidential and residential places needed to spuriously transfer the effect†

Transfer of nonresidential effect to the residential association if the two places are similar	Intervention: residential number of services raised to 200	Intervention: residential number of services raised to 500	Intervention: residential number of services raised to 1000
Similar if below the 1st decile†	0.016 (0.007, 0.025)	0.044 (0.020, 0.069)	0.089 (0.043, 0.135)
Similar if below the 2nd decile†	0.017 (0.008, 0.027)	0.050 (0.025, 0.074)	0.099 (0.054, 0.145)
Similar if below the 3rd decile†	0.018 (0.009, 0.027)	0.053 (0.029, 0.078)	0.107 (0.062, 0.153)
Similar if below the 4th decile†	0.018 (0.009, 0.028)	0.056 (0.031, 0.081)	0.114 (0.068, 0.159)
Similar if below the 5th decile†	0.018 (0.009, 0.028)	0.057 (0.032, 0.082)	0.117 (0.071, 0.162)

* The exact set of predictors included in the model to recalculate the naïve estimate is reported in eAppendix Table 2. The intervention effect estimate is expressed on the probability scale.

† In the recalculation of the naïve estimate, the number of services in a nonresidential place was spuriously raised to 200, 500, or 1000 if the difference in the number of services between the nonresidential place and the residential neighborhood was below the 1st decile, 2nd decile, 3rd decile, 4th decile, and 5th decile of the differences in the number of services between nonresidential and residential places for all participants.

eAppendix 6: Advanced interpretation of the "residential" effect fallacy bias

Several aspects warrant additional comments.

Source of the correlation between exposures at the residential and nonresidential places visited

The correlation between exposures at the residential and nonresidential places visited may be attributable to a large extent to the spatial autocorrelation in the density of services, but also among other mechanisms to the fact that the residential environment determines mode choice and that mode choice influences which nonresidential environments are visited during daily activities (public transport bringing people from high density to high density areas, and car allowing to travel from remote areas to remote areas). The influence of the latter mechanism on the probability of walking in the nonresidential places visited was captured by the residential association in the naïve model, but was picked up by the trip origin / destination associations in the corrected model and therefore removed from the intervention effect estimate. This is a desirable consequence of our correction because a residential intervention would likely not change the modes used in the nonresidential environments visited far from the residence.

Overall behavioral outcome vs. residential neighborhood-specific behavioral outcome

The "residential" effect fallacy described in this article applies to studies correlating residential environment characteristics with overall behavioral outcomes (cumulating behavior conducted inside and outside the residential neighborhood). This bias would not apply to the few studies that investigated associations between residential characteristics and a location-specific outcome (e.g., behavior only in the residential neighborhood).[8,10-12]

However, it should be noted that studies that only consider a location-specific behavioral outcome are of limited interest, since what matters in a Public health perspective is not where people practice a behavior but whether their overall behavior is concordant with health recommendations. Thus relying on an outcome only for the portion of the behavior that takes place in the residential neighborhood neither implies an awareness of the "residential" effect fallacy nor is a proper way to solve it.

Can the "residential" effect fallacy lead to an underestimation of residential effects?

It is indicated in this article that the "residential" effect fallacy leads to an overestimation of the effect of residential characteristics. This is attributable to the fact that environmental characteristics in residential and nonresidential places are positively correlated. It is difficult to think to a case where a residential characteristic would be negatively correlated with the corresponding characteristic in nonresidential environments. However, if that were happening, the "residential" effect fallacy bias would lead to an underestimation of the residential environment effect.

Interferences between participants

We took into account the extent to which an intervention in the residential neighborhood of participant P1 influenced the nonresidential neighborhoods of participant P1, but our calculations made the assumption that interventions in the residential neighborhoods of participants P2, P3, P4, etc. could not influence the nonresidential neighborhoods of participant P1, even if there was some overlap between them. Thus, the present study was based on the assumption that a nonresidential neighborhood of participant P1 could be affected by the intervention only if this nonresidential neighborhood overlapped the residential neighborhood of participant P1, but not by overlapping the residential neighborhood of other participants (P2, P3, P4, etc.). This assumption would be violated if a significant fraction of each participant's nonresidential neighborhoods overlapped the residential neighborhoods of other participants where the intervention was implemented. Our first motivation to follow this approach ignoring interferences between participants was that, typically, an intervention to raise the number of services would not be conducted in all places lacking services over an entire region or country, but in very definite neighborhoods to target some of the populations with the greatest needs. Thus, in case of an intervention to develop services, it is relatively unlikely for a participant affected by the intervention in her/his residential neighborhood to be also affected in some of her/his nonresidential places located relatively far from the residence. Our second motivation to not take into account these interferences between participants is that the induced effect would vary from one study to the other, according to the number of intervention areas disseminated over the study territory and size of this territory, but also according to the patterns of mobility of participants.

Despite this theoretical preference for an intervention effect estimate that does not take into account interferences, we have recalculated the intervention effect estimate taking into account these interferences, as a sensitivity analysis. When interferences between participants were disregarded, the percentage of nonresidential buffers visited by a participant X that overlapped the residential area of the same participant X had a median value of 41.1% across the participants (interquartile range: 20.0%, 61.9%). As a comparison, when interferences between participants were taken into account, the percentage of nonresidential buffers visited by a participant X that overlapped the residential area of any participant had a median value of 88.9 (interquartile range: 71.9%, 96.3%). Thus, even with only 227 participants spread over a large study territory, accounting for interferences between individuals had a substantial impact on the percentage of nonresidential buffers that were affected by the intervention (due to the large 1 km radius selected for the residential intervention areas and nonresidential areas).

We report in eAppendix Table 4 the information already provided in Table 2 in the main article. We have added to this Table the estimate of the intervention effect corrected from the "residential" effect fallacy and taking into account the interferences between the participants.

eAppendix Table 4. Naïve and corrected estimates of the hypothetical intervention effect of raising the number of services in the residential neighborhood to 200, 500, or 1000 for participants in the intervention groups (i.e., with a residential number of services below 200, 500, or 1000)[a]

	Intervention: residential number of services raised to 200	Intervention: residential number of services raised to 500	Intervention: residential number of services raised to 1000
Naïve estimate	0.020 (0.010, 0.029)	0.055 (0.030, 0.079)	0.109 (0.063, 0.154)
Corrected estimate	0.007 (0.001, 0.014)	0.019 (0.000, 0.038)	0.039 (0.004, 0.073)
Corrected estimates taking into account interferences	0.007 (0.000, 0.014)	0.021 (0.003, 0.040)	0.045 (0.010, 0.079)

[a]The exact sets of predictors included in the naïve estimate model and in the corrected estimate models are reported in Supplementary Table 1. The intervention effect estimate is expressed on the probability scale.

There was no difference between the two corrected estimates – without and with the interferences between participants – for the intervention raising the number of services to 200, while differences appeared for the intervention raising the number of services to 500, and became larger for the one raising them to 1000. As expected, the intervention effect estimate became higher when taking into account the interferences between participants. However, this increase in the size of the corrected effect estimate when taking into account the interferences between participants was of much lower magnitude than the "residential" effect fallacy itself.

This novel estimate is provided for quantifying interferences between participants but it is not more valid than our corrected estimate that does not account for interferences. This estimate is related to a different intervention where not only the residential neighborhood but also other portions of the activity space of the participants would receive the intervention, which is different than our original estimation target. Whereas the intervention of interest in this article has a univocal definition (affecting a given area around the residence), the estimate accounting for interferences relates to an intervention with an equivocal definition, as it depends on the magnitude of these interferences.

References

1. Khattak AJ, Rodriguez D. Travel behavior in neo-traditional neighborhood developments: A case study in USA. *Transp Res Part A Policy Pract*. 2005;39:481-500.
2. Inagami S, Cohen DA, Finch BK. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Soc Sci Med*. 2007;65:1779-1791.
3. Sharp G, Denney JT, Kimbro RT. Multiple contexts of exposure: Activity spaces, residential neighborhoods, and self-rated health. *Soc Sci Med*. 2015;146:204-213.
4. Owen N, Humpel N, Leslie E, Bauman A, Sallis JF. Understanding environmental influences on walking; Review and research agenda. *Am J Prev Med*. 2004;27:67-76.
5. Saelens BE, Handy SL. Built environment correlates of walking: a review. *Med Sci Sports Exerc*. 2008;40:S550-566.
6. Sugiyama T, Neuhaus M, Cole R, Giles-Corti B, Owen N. Destination and route attributes associated with adults' walking: a review. *Med Sci Sports Exerc*. 2012;44:1275-1286.
7. Chaix B, Kestens Y, Duncan DT, et al. A GPS-Based Methodology to Analyze Environment-Health Associations at the Trip Level: Case-Crossover Analyses of Built Environments and Walking. *Am J Epidemiol*. 2016;184:570-578.
8. Chaix B, Simon C, Charreire H, et al. The environmental correlates of overall and neighborhood based recreational walking (a cross-sectional analysis of the RECORD Study). *Int J Behav Nutr Phys Act*. 2014;11:20.
9. Karusisi N, Thomas F, Meline J, Brondeel R, Chaix B. Environmental conditions around itineraries to destinations as correlates of walking for transportation among adults: the RECORD cohort study. *PLoS One*. 2014;9:e88929.

10. Troped PJ, Wilson JS, Matthews CE, Cromley EK, Melly SJ. The built environment and location-based physical activity. *Am J Prev Med*. 2010;38:429-438.

11. Giles-Corti B, Timperio A, Bull F, Pikora T. Understanding physical activity environmental correlates: increased specificity for ecological models. *Exerc Sport Sci Rev*. 2005;33:175-181.

12. Li F, Fisher KJ, Brownson RC, Bosworth M. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *J Epidemiol Community Health*. 2005;59:558-564.