



**HAL**  
open science

## Comparison of spatial approaches to assess the effect of residing in a 20-minute neighbourhood on body mass index

Guannan Yang, Lukar E Thornton, Mark Daniel, Basile Chaix, Karen E Lamb

### ► To cite this version:

Guannan Yang, Lukar E Thornton, Mark Daniel, Basile Chaix, Karen E Lamb. Comparison of spatial approaches to assess the effect of residing in a 20-minute neighbourhood on body mass index. *Spatial and Spatio-temporal Epidemiology*, 2022, 43, pp.100546. 10.1016/j.sste.2022.100546 . hal-03973116

**HAL Id: hal-03973116**

**<https://hal.science/hal-03973116v1>**

Submitted on 3 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Comparison of spatial approaches to assess the effect of residing**  
2 **in a 20-minute neighbourhood on body mass index.**

3

4 **Yang G, Thornton LE, Daniel M, Chaix B, Lamb KE.**

5

6 *Spat Spatiotemporal Epidemiol.* 2022;43:100546.

7 doi: 10.1016/j.sste.2022.100546.

8

9

10 **Keywords**

11 Built environment, urban planning, health-environment association, spatial analysis,  
12 spatial autocorrelation

13

14

15 **Abbreviations**

16 BMI, body mass index; PA, physical activity; OLS, ordinary least squares; 20MN, 20-  
17 minute neighbourhood; ProjectPLAN, Places and Locations for Activity and Nutrition  
18 study; km, kilometre; SES, socioeconomic status; ABS, Australian Bureau of Statistics;  
19 IRSAD, Index of Relative Socio-economic Advantage and Disadvantage; SA1,  
20 Statistical Areas Level 1; SA2, Statistical Areas Level 2; kg, kilograms; m, metres; CI,  
21 confidence interval; SD, standard deviation; IQR, interquartile range.

22

23

24

25

26



28 **Abstract**

29 Beliefs that neighbourhood environments influence body mass index (BMI) assume  
30 people residing proximally have similar outcomes. However, spatial relationships are  
31 rarely examined. We considered spatial autocorrelation when estimating associations  
32 between neighbourhood environments and BMI in two Australian cities. Using cross-  
33 sectional data from 1329 participants (Melbourne=637, Adelaide=692), spatial  
34 autocorrelation in BMI was examined for different spatial weights definitions. Spatial  
35 and ordinary least squares regression were compared to assess how accounting for  
36 spatial autocorrelation influenced model findings. Geocoded household addresses were  
37 used to generate matrices based on distances between addresses. We found low positive  
38 spatial autocorrelation in BMI; magnitudes differed by matrix choice, highlighting the  
39 need for careful consideration of appropriate spatial weighting. Results indicated  
40 statistical evidence of spatial autocorrelation in Adelaide but not Melbourne. Model  
41 findings were comparable, with no residual spatial autocorrelation after adjustment for  
42 confounders. Future neighbourhoods and BMI research should examine spatial  
43 autocorrelation, accounting for this where necessary.

44

## 45 **1. Introduction**

46 Many health issues arise from having a high BMI (body mass index) and an enormous  
47 research effort is placed on investigating the determinants of BMI. Research has shown  
48 that BMI is geographically and spatially clustered, suggesting that people residing  
49 proximally generally have a more similar BMI than to those residing further away  
50 (Mills et al., 2020; Qiu et al., 2020; Kim et al., 2018; Guessous et al., 2014; Huang et  
51 al., 2015; King et al., 2006). Explanations for this clustering have included geographical  
52 variations in access to built environmental features considered health promoting (e.g.,  
53 retailers selling fresh food, recreational facilities including open space) or health  
54 harming (e.g., fast food retailers, car-dominant streetscapes) with these variations said  
55 to influence diet and physical activity (PA) behaviours (Lam et al., 2021; Papas et al.,  
56 2007). However, few studies investigating relationships between the neighbourhood  
57 built environment and BMI have adequately accounted for spatial autocorrelation. This  
58 is because commonly used methods like linear regression (Duncan et al., 2014; Hobbs  
59 et al., 2018) assume spatial independence. Ignoring spatial autocorrelation can lead to  
60 incorrect inference by not appropriately estimating standard errors when data are  
61 spatially correlated (Chi and Zhu, 2007). Also, since in any real situation, sample sizes  
62 will be finite, ordinary least squares (OLS) model parameters may be biased when  
63 violations of independence occur (Keitt et al., 2002).

64

65 To handle potential spatial autocorrelation, some studies have accounted for the  
66 clustering of participants living in the same neighbourhoods often defined by an  
67 administrative unit boundary (Cunningham-Myrie et al., 2020; King et al., 2006; Mason  
68 et al., 2020; Tseng et al., 2014; van Lenthe and Mackenbach, 2002; Feng and Wilson,  
69 2015). However, accounting for clustering may be insufficient to explain all of the  
70 spatial relationships in the data. This is because administrative units could be spatially  
71 proximate (even sharing a common boundary), thus sharing common environmental  
72 conditions, yet considered independent. Research has demonstrated that different  
73 conclusions may be reached when comparing spatial regression and multilevel  
74 regression analysis (Takagi et al., 2012; Takagi and Shimada, 2019; Chaix et al., 2005a;  
75 Chaix et al., 2005b), implying that multilevel regression may not fully account for  
76 spatial autocorrelation. As highlighted elsewhere, approaches exclusively based on  
77 accounting for clustering according to geographical areas, rather than individual  
78 neighbours, are insufficient to assess how neighbouring individuals affect one another  
79 (Zangger, 2019). Spatial regression models based on distances between participant  
80 locations, in contrast, enable a better reflection of the effects of neighbouring  
81 individuals on one another (Zangger, 2019).

82

83 Policies and land use planning guidelines that advocate for health promoting built  
84 environments are becoming increasingly popular (McKinnon et al., 2020; Kent and  
85 Thompson, 2014). Originally proposed in Portland (City of Portland Bureau of

86 Planning and Sustainability, 2012), the 20-minute neighbourhood (20MN) concept has  
87 since been adopted in Melbourne (State of Victoria Department of Environment, 2019).  
88 Variations are now seen in European cities, including Paris (Paris En Commun, 2020).  
89 In Melbourne, the written definition of a 20MN has undergone various iterations (State  
90 of Victoria Department of Environment, 2019; State of Victoria Department of  
91 Environment, 2017; State of Victoria Department of Transport, 2014; State of Victoria  
92 Department of Environment, 2015). In brief, a 20MN is a setting that provides residents  
93 with the services and resources to meet most of their daily needs (e.g., access to healthy  
94 food stores, recreational facilities, health and education services) within a 20-minute  
95 non-motorised transport trip from home (noting that in more recent versions of the  
96 policy the mode of transport has been restricted to walking) (Victorian Government  
97 Department of Environment Land Water and Planning, 2020). Urban renewal and  
98 liveability policies that advocate for 20MNs do so under the assumption that these  
99 encourage more physically active lifestyles that in turn lead to better population health.

100

101 Our study aimed to i) examine spatial autocorrelation in BMI using individual  
102 household addresses, and ii) assess differences in BMI between people living in 20MNs  
103 and non-20MNs, accounting for spatial autocorrelation. The statistical approach used  
104 here is arguably an important advance over prior research exploring neighbourhood-  
105 level influences on BMI where assumptions related to the spatial independence of  
106 observations may be totally or partially violated. Advanced statistical modelling is vital

107 to create a more robust evidence base regarding the relationships between place and  
108 health.

109

## 110 **2. Methods**

111 Our study used data from the cross-sectional Places and Locations for Activity and  
112 Nutrition study (ProjectPLAN) which aimed to investigate the benefits of living in a  
113 20MN on PA and dietary behaviors in two Australian cities, Melbourne, Victoria and  
114 Adelaide, South Australia. Melbourne is the capital city of the state of Victoria located  
115 in the south-east edge of Australia; Greater Melbourne covers an area of 9992.5km<sup>2</sup> and  
116 has a population of almost 5 million (City of Melbourne, Last accessed June 2022;  
117 Australian Bureau of Statistics, Last accessed June 2022). Adelaide is the capital city  
118 of the state of South Australia located in the southern central part of Australia; Greater  
119 Adelaide covers an area of 3259.8km<sup>2</sup> (Australian Bureau of Statistics, Last accessed  
120 June 2022) with a population of approximately 1.3 million (Australian Bureau of  
121 Statistics, Last accessed July 2022a). Cities were considered separately in the analysis  
122 as the driving distance from the centre of Melbourne to the centre of Adelaide is over  
123 700km. Data collection took place from October 2018 to May 2019. The study was  
124 approved by the Deakin University Human Research Ethics Committee (HEAG-H  
125 168\_2017).

126



127 2.1. Neighbourhood characteristics

128 20MNs were conceived as address points having high-service and amenity provision,  
129 with the particular services and amenities considered for inclusion based on 20MN  
130 policy documents and a review of existing literature (State of Victoria Department of  
131 Environment, 2019; Thornton et al., 2022). These services and amenities included  
132 multiple attributes across five domains: 1) healthy food outlets (e.g., supermarkets), 2)  
133 recreational resources (e.g., gyms), 3) community resources (e.g., primary schools,  
134 general practitioners, pharmacies), 4) public open spaces, and 5) public transport access.  
135 Attributes of the 20MN are described in Supplementary Table S1. Non-20MNs had low  
136 service and amenity provision, defined as fewer than five individual attributes in  
137 Melbourne and four individual attributes in Adelaide (due to different public transport  
138 infrastructures). Full details of the approach to operationalise the 20MN are described  
139 elsewhere (Thornton et al., 2022). In brief, geocoded services and amenities from a  
140 combination of government and commercial sources were overlaid on a pedestrian  
141 network layer. For healthy food outlets, recreational resources and community  
142 resources, a 1.5-kilometre (km) distance pedestrian network service area, was created  
143 using the service or amenity location as the start point. Different measures were used  
144 for public open space and public transport, with accessibility to these domains guided  
145 by literature or Australian planning guideline recommendations where available. Areas  
146 of overlap between service areas could then be identified to determine 20MNs and non-  
147 20MNs. The level of service and resource provision was determined across the whole

148 of the Melbourne and Adelaide metropolitan areas. This allowed the identification of  
149 address points with access to all services and resources (these address points were  
150 defined as having a 20MN) and addresses with low numbers of services and resources  
151 (non-20MNs). This approach means we were not bound by a clustered study design  
152 whereby participants needed to be within the same administrative units, typical of many  
153 studies of neighbourhood effects. The distribution of 20MNs in both cities are shown  
154 in Figures 4 and 5 of Thornton et al. (2022). In general, residential addresses that had a  
155 20MN were commonly situated within inner and mid-suburban areas of Melbourne,  
156 although they extended to the outer areas along the train lines. In Adelaide, residential  
157 addresses with a 20MN were typically located centrally in the city, although they were  
158 also found in townships extended through the middle and outer edges of the city.

159

160 Areas of low and high socioeconomic status (SES) were identified using the Australian  
161 Bureau of Statistics (ABS) Index of Relative Socio-economic Advantage and  
162 Disadvantage (IRSAD). Areas with low socioeconomic conditions were defined by  
163 extracting deciles 1, 2 or 3 of the IRSAD at Statistical Areas Level 1 (SA1, i.e.,  
164 geographical areas built from whole Mesh Blocks which are the smallest geographical  
165 areas defined by the ABS) which were located within IRSAD deciles 1, 2 or 3 at the  
166 larger Statistical Areas Level 2 (SA2). This approach was adopted to ensure that small  
167 areas of low socioeconomic conditions were located within a larger community that

168 also had low socioeconomic conditions. The process was repeated for SA1s and SA2s  
169 within deciles 8, 9 and 10 to represent areas with high socioeconomic conditions.

170

## 171 2.2. Sampling and recruitment

172 Stratified random sampling was used to identify households within each city  
173 (Melbourne or Adelaide), neighbourhood type (20MN or non-20MN), and level of  
174 neighbourhood SES (low or high) to be sent an invitation letter with a URL and unique  
175 password for accessing either a food or PA survey. The sampling approach allowed the  
176 household location of participants to be captured which was necessary to account for  
177 the spatial autocorrelation between participants. Only address points within residential  
178 Mesh Blocks were included in the sampling frame to reduce the number of invitations  
179 sent to non-residential addresses. Address point data were sourced from routinely  
180 available government data sources (Department of Environment, 2021; Government of  
181 South Australia, 2021). The total number of eligible addresses within each city by  
182 neighbourhood type and SES are shown in Table 1. We aimed to recruit equal numbers  
183 within each stratum. Higher numbers of letters were mailed to address points within  
184 low SES areas in both cities due to lower anticipated response rates in these areas.  
185 Reminder letters were also issued to some address points to increase response rates.

186

187 Given the sampling approach, there were eight combinations of city, neighbourhood  
188 type and status (i.e., Melbourne 20MN low SES; Melbourne 20MN high SES;

189 Melbourne non-20MN low SES; Melbourne non-20MN high SES; Adelaide 20MN low  
190 SES; Adelaide 20MN high SES; Adelaide non-20MN low SES; Adelaide non-20MN  
191 high SES). Within each combination, randomly selected eligible households were  
192 mailed an invitation to participate in either the online food or PA survey. Participants  
193 aged 18 years or older were invited to complete the food survey if they were at least  
194 jointly if not fully responsible for the household food shopping. The person (aged  $\geq 18$   
195 years) in the household who had the most recent birthday was asked to complete the PA  
196 survey. Participants were asked to confirm they lived at the address on file. In total, 782  
197 participants (3.7% response rate) from Melbourne and 830 participants (4.2% response  
198 rate) from Adelaide consented to complete either the food or PA survey (Table 1).

199

## 200 2.3. Variables

### 201 2.3.1. *Exposure*

202 The exposure was neighbourhood type (20MN or non-20MNs).

203

### 204 2.3.2. *Outcome*

205 BMI (weight [kilograms, kg]/height [metres, m]<sup>2</sup>) was calculated from self-reported  
206 weight and height.

207

208 2.3.3. *Other covariates*

209 Potential confounders included gender (male; female), age, individual SES based on  
210 highest qualification obtained (less than university; university) and self-rated financial  
211 stress indicated by the ability to manage on household income (very difficult or difficult;  
212 just getting by; comfortable or very comfortable).

213

214 2.4. Statistical analysis

215 Participants who were pregnant or who had missing or implausible BMI ( $>50\text{kg/m}^2$  or  
216  $<15\text{kg/m}^2$ ) were excluded, leaving 1382 (86%) participants from the original sample  
217 eligible for analysis. The statistical analysis was conducted for eligible participants with  
218 complete data for the exposure and potential confounding variables. A high percentage  
219 of the eligible sample had complete data ( $n=1329$ , 96%). Descriptive characteristics for  
220 the eligible, complete case and omitted samples are shown in Supplementary Table S2.

221

222 Analyses were conducted separately for Melbourne and Adelaide because there were  
223 slightly different definitions of 20MN used (due to different public transport systems)  
224 in each city. Food survey and PA survey participants were combined for analyses and a  
225 parameter for survey type included in the models as these surveys had different target  
226 populations; the food survey was completed by the primary or joint primary household  
227 food purchaser and the PA survey was any adult in the household. The resulting sample  
228 sizes were 637 for Melbourne (Melbourne Food Survey=287, Melbourne PA

229 Survey=350) and 692 for Adelaide (Adelaide Food Survey=351, Adelaide PA  
230 Survey=341). Demographic characteristics for the two cities were compared to the 2016  
231 Census data for Greater Melbourne (Australian Bureau of Statistics, Last accessed July  
232 2022b) and Greater Adelaide (Australian Bureau of Statistics, Last accessed July  
233 2022a).

234

#### 235 *2.4.1. Global spatial autocorrelation*

236 Moran's  $I$  was calculated to assess spatial autocorrelation in BMI for both Melbourne  
237 and Adelaide under different spatial weights definitions (Cliff and Ord, 1971; Cliff and  
238 Ord, 1972). Moran's  $I$  is the most commonly used measure of spatial autocorrelation,  
239 providing a measure of how related observations are based on the location at which  
240 they are measured. The correlation ranges from values of -1, indicating negative spatial  
241 autocorrelation or complete spatial dispersion, to 1, indicating high positive spatial  
242 autocorrelation. A value of 0 indicates complete spatial randomness.

243

244 Eight spatial weights matrices were considered based on both the Euclidean and road  
245 network distance between individual household addresses to investigate whether  
246 findings were sensitive to the weight chosen. A 3-km bandwidth weight matrix was  
247 considered as this distance was directly related to the 20MN exposure definition. Under  
248 the assumption that people residing in 20MNs have similar health and behaviour  
249 outcomes, it could be hypothesised that those living proximally (i.e., within the pre-

250 specified threshold of a 20MN) have correlated outcomes so this matrix choice enabled  
251 us to assess this assumption. In this definition, participants were defined to be  
252 neighbours (with a '1' entered in the matrix) if they lived within a 3-km distance of one  
253 another; otherwise, participants were not considered neighbours ('0' in the matrix). It  
254 was assumed that 1.5-km is the distance an adult can walk in 20 minutes, as discussed  
255 previously, so participant pairs whose distances were less than 3-km could share the  
256 same amenity or service, as illustrated in Figure 1. In addition, a 1.5-km bandwidth  
257 weight matrix, corresponding to the assumed distance that can be travelled in a 20-  
258 minute walk, was considered. Matrices were row-normalised, meaning each element in  
259 row  $i$  (where  $i$  represents each study participant) was divided by the sum of the elements  
260 of row  $i$ , meaning all weight elements took a value between 0 and 1. This means that  
261 the spatial weight is shared among the total number of neighbours of participant  $i$  (e.g.,  
262 for the 3-km bandwidth weight matrix, if participant  $i$  had 3 neighbours, each was  
263 weighted by  $1/3$ ).

264

265 Inverse-distance weights using the calculated distance between each pair of  
266 observations were also considered; this matrix allows all observations to have  
267 neighbours, with those living nearer assumed to be more related than those living  
268 further. The analysis using the inverse-distance matrix thus assumes that participants  
269 would be more influenced by those who live closer to them than those who live further  
270 away but enables all participants to have a spatial relationship with all other participants

271 within strata of city (Adelaide or Melbourne). In addition, the inverse-distance-squared  
272 matrix was considered which allows the weights between neighbours to decrease  
273 rapidly with increasing distance.

#### 274 2.4.2. Regression models

275 Unadjusted and confounder-adjusted OLS linear regression models were fitted to  
276 examine associations between having a 20MN (no/yes) and BMI in each of the two  
277 cities. The Moran's *I* test was used to assess spatial autocorrelation in the residuals of  
278 these models using all spatial weights definitions.

279

280 To account for potential residual spatial autocorrelation in the OLS regression models,  
281 spatial error regression models were fitted to compare with results from the OLS  
282 regression models. The spatial lag model was not considered because these models are  
283 most suitable for situations where there is an intuitive or theoretical motivation that the  
284 dependent variable will depend on its neighbours' values (e.g., housing price, local  
285 government actions) (LeSage and Pace, 2009). The purpose of the spatial modelling in  
286 this analysis was to avoid residual spatial autocorrelation where a spatial pattern in the  
287 residuals could exist due to omitted random factors. Hence, spatial error models were  
288 used in this study. The spatial error models do not assume the interdependence between  
289 any two neighbouring observations operate in any particular way, meaning that the  
290 outcome (i.e., BMI) of a given individual will be to some degree dependent on the  
291 behaviours of their neighbours due to some unmeasured variables, thus accounting for



292 any unmeasured environmental exposures or individual characteristics shared by  
293 neighbouring participants (Zangger, 2019). The spatial error model is expressed as  
294 follows:

$$295 \quad y = \mathbf{X}\beta + \varepsilon$$

$$296 \quad \varepsilon = \lambda \mathbf{W}\varepsilon + u$$

297 where  $u \sim N(0, \sigma^2 \mathbf{I})$  is the  $N \times 1$  error vector, assumed to be independent and  
298 identically distributed;  $\mathbf{W}$  is the  $N \times N$  spatial weight matrix (either 3-km bandwidth  
299 or inverse-distance) and  $\lambda \mathbf{W}\varepsilon$  is the vector of the spatially weighted average of errors  
300 (from the observation's neighbours, when the matrix is row-normalised);  $\lambda$  is the  
301 spatial autoregressive coefficient, representing the spatial autocorrelation of the error  
302 term  $\varepsilon$ . Since  $\varepsilon = (\mathbf{I} - \lambda \mathbf{W})^{-1}u$ , the error variance-covariance matrix is written as:

$$303 \quad \sigma^2 (\mathbf{I} - \lambda \mathbf{W})^{-1} (\mathbf{I} - \lambda \mathbf{W}^T)^{-1}$$

304 Once the variance-covariance matrix of error was obtained, after estimating  $\lambda$  by  
305 maximum likelihood, the estimated parameters for the association between  
306 neighbourhood type and BMI in the spatial error model were estimated by OLS method.  
307 The coefficients of neighbourhood type and their 95% confidence intervals (CIs) can  
308 be interpreted as the differences in mean BMI between participants residing in 20MNs  
309 and non-20MNs.

310

311 *2.4.3. Sensitivity analyses*

312 Two additional types of spatial weights matrix were considered in sensitivity analyses  
313 using both Euclidean and road network distances: an exponential decay weight matrix  
314 with decay constant 1, which allows the weights between neighbours to decrease more  
315 slowly with increasing distance, and an inverse-distance with 3-km bandwidth weight  
316 matrix.

317

318 All analyses were conducted using R (version 3.5.3, The R Foundation) and RStudio  
319 (version 1.1.463, RStudio, Inc.), with spatial analysis conducted using the “spdep”  
320 package (Bivand and Wong, 2018).

321

## 322 **3. Results**

### 323 3.1. Descriptive analysis

324 Descriptive characteristics are shown in Table 2. The average age of participants was  
325 50.6 (standard deviation [SD]=16.4) years in Melbourne and 57.3 (SD=15.7) years in  
326 Adelaide. The ProjectPLAN participants reflect a demographic older than the general  
327 populations of Greater Melbourne and Greater Adelaide (Supplementary Table S3).  
328 There were higher percentages of women than men in each sample (57.3% in  
329 Melbourne and 58.4% in Adelaide); both greater than the overall proportion (51%) of  
330 women in the general population for each city. A much higher proportion of  
331 ProjectPLAN participants had university education (46.4% in Adelaide and 60.4% in

332 Melbourne) than observed in the general population of these two cities (21-28%; see  
333 Supplementary Table S3). Mean BMI was slightly higher in the Adelaide sample at  
334 26.8kg/m<sup>2</sup> (SD=5.0) than in the Melbourne sample (25.4kg/m<sup>2</sup>, SD=4.4). Just over half  
335 of respondents in Adelaide (53.0%) resided in a 20MN. The percentage residing in a  
336 20MN was lower in Melbourne (39.9%).

337

338 The median Euclidean distance between participants was lower in Adelaide (16.7-km,  
339 interquartile range [IQR]: 9.5-26.3) than in Melbourne (20.6km, IQR: 12.5-23.4).  
340 Similarly, participants had a lower median network distance in Adelaide (22.1-km,  
341 IQR:11.8-30.8) than in Melbourne (26.8-km, IQR: 15.1-38.1). The median number of  
342 neighbours within 3-km was higher in Adelaide (38.5, IQR: 24-61) than Melbourne (22,  
343 IQR: 12-33). Further details are shown in Supplementary Table S4.

344

### 345 3.2. Spatial autocorrelation in BMI

346 The Moran's *I* values for BMI based on the eight spatial weights definitions are shown  
347 in Table 3.

348

#### 349 3.2.1 Euclidean distance

350 The estimated spatial autocorrelation in BMI was low and positive for all types of  
351 Euclidean spatial matrix, ranging from 0.008 (Melbourne: inverse-distance spatial  
352 matrix) to 0.059 (Adelaide: 1.5-km bandwidth spatial matrix). The estimated spatial

353 autocorrelation was lower in Melbourne than Adelaide for all spatial weight definitions.  
354 Analyses showed no evidence against the null hypothesis of spatial randomness for  
355 almost all spatial weight definitions in Melbourne, although the Moran's  $I$  test based  
356 on the Euclidean 3-km bandwidth spatial weights on the Melbourne sample had a small  
357  $p$ -value (Moran's  $I=0.032$ ,  $p=0.029$ ). In contrast, three of four Euclidean spatial  
358 matrices in the Adelaide sample showed strong evidence against the null hypothesis of  
359 spatial randomness in BMI.

360

361 The Moran's  $I$  values based on the 3-km bandwidth weights matrices and the 1.5-km  
362 bandwidth weights matrices were slightly larger than the inverse-distance or the  
363 inverse-distance-squared weights matrices in both cities.

364

### 365 *3.2.2 Network distance*

366 Findings using network distances were comparable to those using Euclidean distances.  
367 The spatial autocorrelation in BMI ranged from 0.004 (Melbourne: inverse-distance  
368 spatial matrix) to 0.051 (Adelaide: 3-km bandwidth spatial matrix). Similar to the  
369 Euclidean-distance-based matrices, there was no evidence against the null hypothesis  
370 of spatial randomness for all spatial weight definitions in Melbourne, while three of  
371 four network spatial matrices in Adelaide showed strong evidence to suggest the  
372 presence of positive spatial autocorrelation in BMI. Generally, the values of Moran's  $I$   
373 derived from network distances were smaller than those derived from Euclidean

374 distances.

375

### 376 3.2. Comparison of regression models

377 The difference in mean BMI between residents of 20MN and non-20MN and the  
378 corresponding CIs for each model are shown in Figure 2. The results show that  
379 adjustment for covariates affected the estimated differences, indicating potential bias in  
380 the unadjusted estimates due to confounding. The estimated mean difference in BMI  
381 was positive for all models, indicating a higher mean BMI in people living in non-  
382 20MNs than 20MNs. However, after adjustment for potential confounders, differences  
383 in mean BMI attenuated in both Melbourne ( $0.22\text{kg/m}^2$ , 95% CI=-0.48-0.93; Model 2  
384 Supplementary Table S5) and Adelaide ( $0.71\text{kg/m}^2$ , 95% CI=-0.04-1.45; Model 2  
385 Supplementary Table S6).

386

387 Moran's  $I$  values for the residuals of the unadjusted models were smaller than those  
388 observed for BMI in all the samples (Table 3). There was no residual spatial  
389 autocorrelation in Melbourne after adjusting for living in 20MN (unadjusted OLS  
390 model). Findings for Adelaide were mixed and dependent on the spatial matrix  
391 considered. There was strong evidence of residual spatial autocorrelation in Adelaide  
392 for all but one (the inverse-distance-squared) Euclidean distance spatial matrices.  
393 However, there was only evidence of positive residual spatial autocorrelation for the 3-  
394 km bandwidth matrix when considering the network distance. After accounting for

395 potential confounders in the OLS models, the spatial autocorrelation attenuated to close  
396 to zero in both Melbourne and Adelaide for all spatial weights matrices (Table 3),  
397 indicating complete spatial randomness.

398

399 There was little discrepancy in inference assessing the difference in BMI between  
400 20MN and non-20MNs comparing the OLS regression models with the spatial  
401 regression models, irrespective of which spatial matrix or type of distance was used.

402

### 403 3.3 Sensitivity analyses

404 Sensitivity analyses are presented in Supplementary Table S7. Similar to the results  
405 presented in Table 3, there was strong evidence of spatial autocorrelation in BMI in  
406 Adelaide considering the inverse-distance with 3-km bandwidth spatial weight or  
407 Exponential decay weight, with Moran's  $I$  ranging from 0.044 ( $p=0.042$ ; inverse-  
408 distance with 3-km bandwidth network distance matrix) to 0.055 ( $p<0.001$ ; Exponential  
409 decay network distance matrix). Estimated spatial autocorrelation in Melbourne was  
410 lower (0.021 to 0.029) and there was only strong statistical evidence of spatial  
411 autocorrelation using the exponential decay Euclidean distance matrix (Moran's  
412  $I=0.032$ ,  $p=0.037$ ). As with the other spatial matrices considered, the estimated residual  
413 spatial autocorrelation in both Melbourne and Adelaide attenuated to close to zero after  
414 accounting for potential confounders in the OLS regression models.

415

## 416 **4. Discussion**

417 Our study found some evidence of positive global spatial autocorrelation in BMI in  
418 Adelaide for almost all types of spatial weights matrices considered. However, this was  
419 of a low level, with estimated correlations below 0.1. This is accordant with some  
420 previous studies that considered area-level measures to define spatial matrices (Laraia  
421 et al., 2014; Han et al., 2018; Paquet et al., 2016) but in contrast to others that  
422 demonstrate spatial randomness or spatial dispersion (Penney et al., 2014; Schuurman  
423 et al., 2009). Only the inverse-distance-squared weight matrix provided no statistical  
424 evidence of spatial autocorrelation in BMI in Adelaide. Generally, there was no  
425 evidence of spatial autocorrelation in BMI in Melbourne, where estimated spatial  
426 autocorrelation was consistently lower than in Adelaide for all spatial matrices  
427 considered. However, our results do show differences in the magnitude of spatial  
428 autocorrelation depending on the choice of spatial weights matrix, with positive  
429 correlations in BMI ranging from 0.008 to 0.040 in Melbourne and 0.021 to 0.059 in  
430 Adelaide. This highlights the importance of the choice of spatial weights matrix in the  
431 analysis.

432

433 Estimated spatial autocorrelation in BMI was found to be lower using the inverse-  
434 distance weights than the bandwidth weights. The use of the inverse-distance weight is  
435 appealing as it does not enforce an arbitrary threshold on distances when defining

436 neighbours. However, by not imposing any distance limits, the inverse-distance matrix  
437 considers weighting between pairs even if they reside very far from one another. In this  
438 approach, two participants residing 50-km and 51-km from participant  $i$  would have  
439 different weights, even though arguably these participants reside too far for either to  
440 have an impact on participant  $i$ . Therefore, perhaps the inverse-distance weights  
441 become meaningless beyond certain distances. Interestingly, the estimated spatial  
442 autocorrelations in BMI from sensitivity analysis using combined inverse-distance and  
443 threshold weights were more comparable to those of the bandwidth only approach than  
444 the inverse-distance only approach.

445

446 Comparison of adjusted OLS and spatial regression models examining BMI by  
447 neighbourhood type showed little difference in findings when accounting for spatial  
448 autocorrelation using any of the spatial weights matrix definitions. After adjustment for  
449 potential confounders in the OLS models, there was no evidence of residual spatial  
450 autocorrelation in either city. This suggests the spatial autocorrelation in BMI observed  
451 in Adelaide may be due to the shared environment or characteristics of the neighbours.  
452 Of course, these findings are study- and context-specific, and our results in no way  
453 negate a need to appropriately consider spatial relationships in research in this area. Our  
454 results provide some indication that mean BMI was higher in non-20MNs than 20MNs,  
455 with the estimated mean difference higher in Adelaide than Melbourne. All observed  
456 differences were subsequently attenuated, however, after adjustment for confounders.



457 This was so even for Adelaide, where CIs included the null, though with minimal  
458 overlap. Importantly, as our samples were older and more highly educated than the  
459 general population of Greater Melbourne and Greater Adelaide, our findings may not  
460 be generalisable as representative of the broader population. No other studies have  
461 examined the effect of having a 20MN on BMI. However, studies examining composite  
462 neighbourhood environment measures have yielded mixed findings. In Australia, Tseng  
463 et al. (2014) did not find meaningful correlations in the relationship between an index  
464 of neighbourhood access to food resources, PA facilities and transport and BMI among  
465 women living in socioeconomically disadvantaged areas. In the UK, Hobbs et al. (2018)  
466 found that a combination of access to both food and PA environmental facilities may be  
467 important for BMI and that the richness of environmental attributes is not necessarily  
468 linearly associated with BMI. Future research on liveable environments, considering  
469 multiple measures of neighbourhood attributes, is required to understand built  
470 environment relationships with BMI.

471

472 Choosing a spatial weights matrix is challenging but critical for spatial analysis as it  
473 proposes the spatial dependence structure imposed and can affect both the spatial  
474 autocorrelation test and the estimates of spatial regression models (Anselin, 2002;  
475 Fingleton, 2003). In research conducted elsewhere examining the ability to estimate  
476 area-level relative risk of birth defects in Australia, the researchers found that models  
477 incorporating distance-based spatial weights matrices performed better than adjacency-

478 based weights (Earnest et al., 2007). One reason Earnest et al. cited for this is that  
479 adjacency matrices are likely to perform better with more regular shaped areas than  
480 obtained when dealing with administrative data and they argue the need for careful  
481 consideration of appropriate spatial matrices. In addition to the methods used in the  
482 present study, a spatial weight matrix could be chosen through an exploratory analysis  
483 to assess the degree of spatial autocorrelation (Earnest et al., 2007). One approach to  
484 reduce the arbitrariness of the choice is to select from a set of predefined matrices by  
485 criterion, such as comparison of values of Moran's  $I$ , AIC or model boosting method  
486 (Kostov, 2010).

487

488 Our use of the 3-km bandwidth is appealing as it is linked to the distance deemed  
489 important for our 20MN exposure definition. However, this approach places an  
490 arbitrary threshold on the definition of 'neighbours' through this dichotomisation (e.g.,  
491 participants located within 2.99-km of one another are classified as neighbours while  
492 those 3.01-km apart are not). Defining participants as neighbours (an entry of 1 in the  
493 matrix) or not (0) using the 3-km bandwidth is analogous to the common approach used  
494 when considering administrative boundaries where areas (e.g., postal areas or ZIP codes)  
495 are defined as neighbourhoods if they share a common boundary. Although attractive  
496 due to simplicity, the common approach using administrative boundaries does not allow  
497 for vastly varying sizes typical of administrative units. For example, in Australia inner  
498 city administrative units are much smaller than those in regional or rural areas. This

499 means the distance between participants deemed ‘neighbours’ in remote areas can be  
500 much greater than the distances considered between neighbours in more urban areas.  
501 While understandable as administrative units are created based on the number of  
502 households and population density, the implications of this are that within the same  
503 study, the distance used to represent neighbours can vary extensively.

504

505 The use of distance-based spatial weights in built environment and health research  
506 avoids this issue. The inverse-distance spatial matrix, or variations of this matrix (e.g.,  
507 inverse-distance-squared), allows the relative distance between *all* participants to be  
508 considered, avoiding the specification of arbitrary boundaries and administrative  
509 neighbouring areas of varying size, whether distance-based (e.g., the 3-km bandwidth)  
510 or not (e.g., defining administrative areas as neighbours if they share common  
511 boundaries). However, it should be noted that the inverse-distance spatial matrix is  
512 typically calculated based on Euclidean distances. Although Euclidean distances often  
513 work well in the context of highly urbanised areas where the underlying level of  
514 connectivity is high (Thornton et al., 2012), this means that barriers (e.g., highways or  
515 bodies of water) between participants are not accounted for when defining neighbours.  
516 This has important implications when considering built environment and health  
517 research where increasingly exposure measures, such as access to retail or public open  
518 space, are defined based on pedestrian street network distances. Therefore, to better  
519 reflect neighbours in this context, spatial matrices should also be based on a network

520 distance measure, as considered in this study, to represent the walking or driving  
521 distance between study participants as this may better reflect accessibility between  
522 participants. Our findings show that the magnitude of spatial autocorrelation can differ  
523 when using Euclidean or network distances, although our inference regarding the  
524 relationship between 20MNs and BMI was not affected by which type of distance was  
525 used.

526

527 Importantly, it should be noted that our study was able to consider distances based on  
528 participant location as our sample sizes were modest at less than 700 in each sample.  
529 Larger studies may encounter difficulties in fitting spatial models considering between  
530 participant distances. In Stata, for example, the maximum possible matrix size is  
531 currently 11,000. Therefore, large studies could either incorporate weighted averages  
532 of the response as model covariates as shown in spatial modelling elsewhere (Lee and  
533 Neocleous, 2010) or define neighbours based on areal units.

534

535 Strengths of our study include our approach to create spatial matrices considering the  
536 precise spatial location of participants through geocoded home addresses. This enabled  
537 greater specification of spatial relationships than the use of arbitrary boundaries. While  
538 research in this area is often hampered by a lack of access to sensitive spatial location  
539 data, particularly if using secondary data sources to examine neighbourhood effects on  
540 health, our project greatly reduced the risk of participant identification by analysts by

541 separately creating spatial matrices before merging only the spatial matrices, not the  
542 address data, to anonymised participant characteristic data. This demonstrates the  
543 possibility of spatial matrices being created externally without the need to release  
544 participant data with geocoded addresses. This is an approach that may be considered  
545 by other researchers wishing to take spatial relationships into account who are unable  
546 to access participant locations. It may be possible for data custodians to create these  
547 matrices using open-source software to be linked to participant characteristic data  
548 without the geocoded locations.

549

550 Although we considered a number of different spatial matrices in this analysis, future  
551 research could consider other types of spatial matrix, including different distance decay  
552 functions or exploring different thresholds in a mixed approach combining distance  
553 decay with different distance cut-points. Our choice of a 3-km cut-point was based on  
554 accessible distances related to the 20MN definition. However, across different contexts,  
555 there will be questions about what this threshold distance should be. The choice of the  
556 spatial matrices used in the primary analysis for this study was based on common matrix  
557 types likely to be adopted by behavioural researchers in neighbourhoods and health  
558 wishing to undertake spatial analysis and is by no means exhaustive. However, our  
559 findings suggest that spatial autocorrelation values are relatively low and have minimal  
560 impact on modelling findings.

561

562 The lack of consideration of spatial models in most studies of the environment and BMI  
563 could be attributable to several reasons. Researchers may be unfamiliar with these  
564 approaches, or the diverse and complex nature of spatial models may be off-putting.  
565 However, the analysis conducted in this study was undertaken using open source  
566 statistical and geographical tools, demonstrating that these methods are becoming more  
567 accessible to researchers. Although the findings from ProjectPLAN suggest little need  
568 for spatial regression models when examining differences in BMI by neighbourhood  
569 type after adjustment for individual characteristics, it is important to acknowledge that  
570 these findings are context specific. Therefore, future studies of the built environment  
571 on BMI, or other health outcomes, should take care to examine and appropriately  
572 account for spatial autocorrelation.  
573

## **Authors' contributions**

GY and KL led the writing of the article with input from all authors (LT, MD, BC). GY conducted the statistical analysis. KL and LT led the conceptualisation of the study. All authors provided critical feedback on the article drafts and approved the final manuscript.

## **Acknowledgements**

The authors thank Neil Coffee, Kylie Ball, and Yan Kestens for their contributions to the ProjectPLAN study conceptualisation and data collection tools. We further acknowledge Keren Best, Simon Best, Ralf Dieter-Schroers, and Steffi Renehan for assistance with data collection.

## **Funding**

This work was supported by an Australian Research Council Discovery Project Grant [DP 170100751].

## References

- Anselin L. (2002) Under the hood Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27: 247-267.
- Australian Bureau of Statistics. (Last accessed July 2022a) *Greater Adelaide: 2016 Census All persons QuickStats* <https://www.abs.gov.au/census/find-census-data/quickstats/2016/4GADE>.
- Australian Bureau of Statistics. (Last accessed July 2022b) *Greater Melbourne: 2016 Census All persons QuickStats* <https://www.abs.gov.au/census/find-census-data/quickstats/2016/2GMEL>.
- Australian Bureau of Statistics. (Last accessed June 2022) *Greater Capital City Statistical Area (GCCSA)* [https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Greater%20Capital%20City%20Statistical%20Areas%20\(GCCSA\)~10003](https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Greater%20Capital%20City%20Statistical%20Areas%20(GCCSA)~10003).
- Bivand RS and Wong DWS. (2018) Comparing implementations of global and local indicators of spatial association. *Test* 27: 716-748.
- Chaix B, Merlo J and Chauvin P. (2005a) Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *Journal of Epidemiology and Community Health* 59: 517-526.
- Chaix B, Merlo J, Subramanian S, et al. (2005b) Comparison of a spatial perspective with the multilevel analytical approach in neighborhood studies: the case of mental and behavioral disorders due to psychoactive substance use in Malmö, Sweden, 2001. *American Journal of Epidemiology* 162: 171-182.
- Chi G and Zhu J. (2007) Spatial regression models for demographic analysis. *Population Research and Policy Review* 27: 17-42.
- City of Melbourne. (Last accessed June 2022) *Melbourne facts and figures* <https://www.melbourne.vic.gov.au/about-melbourne/melbourne-profile/Pages/facts-about-melbourne.aspx>.
- City of Portland Bureau of Planning and Sustainability. (2012) Portland Plan. 20-minute neighborhoods analysis: background report and analysis area summaries. In: Bureau of Planning and Sustainability (BPS) (ed).
- Cliff A and Ord K. (1972) Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* 4: 267-284.
- Cliff AD and Ord K. (1971) Evaluating the percentage points of a spatial autocorrelation coefficient. *Geographical Analysis* 3: 51-62.
- Cunningham-Myrie CA, Younger NO, Theall KP, et al. (2020) Understanding neighbourhood retail food environmental mechanisms influencing BMI in the Caribbean: a multilevel analysis from the Jamaica Health and Lifestyle Survey: a cross-sectional study. *BMJ Open* 10: e033839.



- Department of Environment, Land, Water & Planning. (2021) Vicmap Address Melbourne [last accessed December 2021]. Available from: <https://www.land.vic.gov.au/maps-and-spatial/spatial-data/vicmap-catalogue/vicmap-address>.
- Duncan DT, Sharifi M, Melly SJ, et al. (2014) Characteristics of walkable built environments and BMI z-scores in children: evidence from a large electronic health record database. *Environmental Health Perspectives* 122: 1359-1365.
- Earnest A, Morgan G, Mengersen K, et al. (2007) Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *International Journal of Health Geographics* 6: 54.
- Feng X and Wilson A. (2015) Do neighbourhood socioeconomic circumstances not matter for weight status among Australian men? Multilevel evidence from a household survey of 14 691 adults. *BMJ Open* 5: e007052.
- Fingleton B. (2003) Externalities, economic geography, and spatial econometrics: conceptual and modeling developments. *International Regional Science Review* 26: 197-207.
- Government of South Australia, Department for Infrastructure and Transport. (2021) Roads: Government of South Australia [last accessed December 2021]. Available from: <https://data.sa.gov.au/data/dataset/roads>.
- Guessous I, Joost S, Jeannot E, et al. (2014) A comparison of the spatial dependence of body mass index among adults and children in a Swiss general population. *Nutrition Diabetes* 4: e111.
- Han EJ, Kang K and Sohn SY. (2018) Spatial association of public sports facilities with body mass index in Korea. *Geospatial Health* 13: 542.
- Hobbs M, Griffiths C, Green MA, et al. (2018) Neighbourhood typologies and associations with body mass index and obesity: A cross-sectional study. *Preventive Medicine* 111: 351-357.
- Huang R, Moudon AV, Cook AJ, et al. (2015) The spatial clustering of obesity: does the built environment matter? *Journal of Human Nutrition and Dietetics* 28: 604-612.
- Keitt TH, Bjørnstad ON, Dixon PM, et al. (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25: 616-625.
- Kent JL and Thompson S. (2014) The three domains of urban planning for health and well-being. *Journal of Planning Literature* 29: 239-256.
- Kim D, Wang F and Arcan C. (2018) Geographic association between income inequality and obesity among adults in New York State. *Preventing Chronic Disease* 15: E123.
- King T, Kavanagh AM, Jolley D, et al. (2006) Weight and place: a multilevel cross-sectional survey of area-level social disadvantage and overweight/obesity in Australia. *International Journal of Obesity* 30: 281-287.
- Kostov P. (2010) Model boosting for spatial weighting matrix selection in spatial lag models. *Environment and Planning B: Planning and Design* 37: 533-549.

- Lam TM, Vaartjes I, Grobbee DE, et al. (2021) Associations between the built environment and obesity: an umbrella review. *International Journal of Health Geographics* 20: 7.
- Laraia BA, Blanchard SD, Karter AJ, et al. (2014) Spatial pattern of body mass index among adults in the diabetes study of Northern California (DISTANCE). *International Journal of Health Geographics* 13: 48.
- Lee D and Neocleous T. (2010) Bayesian quantile regression for count data with application to environmental epidemiology. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 59: 905-920.
- LeSage J and Pace RK. (2009) *Introduction to spatial econometrics*, New York: Chapman and Hall/CRC.
- Mason KE, Pearce N and Cummins S. (2020) Do neighbourhood characteristics act together to influence BMI? A cross-sectional study of urban parks and takeaway/fast-food stores as modifiers of the effect of physical activity facilities. *Social Science and Medicine* 261: 113242.
- McKinnon G, Pineo H, Chang M, et al. (2020) Strengthening the links between planning and health in England. *BMJ* 369: m795.
- Mills CW, Johnson G, Huang TTK, et al. (2020) Use of small-area estimates to describe county-level geographic variation in prevalence of extreme obesity among US adults. *JAMA Network Open* 3: e204289.
- Papas MA, Alberg AJ, Ewing R, et al. (2007) The built environment and obesity. *Epidemiologic Reviews* 29: 129-143.
- Paquet C, Chaix B, Howard NJ, et al. (2016) Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia. *International Journal of Environmental Research and Public Health* 13: 519.
- Paris En Commun. (2020) *Ville du 1/4h*. Available at: <https://annehidalgo2020.com/thematique/ville-du-1-4h/>.
- Penney TL, Rainham DG, Dummer TJ, et al. (2014) A spatial analysis of community level overweight and obesity. *Journal of Human Nutrition and Dietetics* 27 Suppl 2: 65-74.
- Qiu G, Liu X, Amiranti AY, et al. (2020) Geographic clustering and region-specific determinants of obesity in the Netherlands. *Geospatial Health* 15: 1.
- Schuurman N, Peters PA and Oliver LN. (2009) Are obesity and physical activity clustered? A spatial analysis linked to residential density. *Obesity* 17: 2202-2209.
- State of Victoria Department of Environment, Land, Water and Planning. (2015) Plan Melbourne refresh - Discussion paper.
- State of Victoria Department of Environment, Land, Water and Planning. (2017) Plan Melbourne 2017-2050.
- State of Victoria Department of Environment, Land, Water and Planning. (2019) 20-minute neighbourhoods. Creating a more liveable Melbourne.
- State of Victoria Department of Transport, Planning and Local Infrastructure. (2014) Plan Melbourne: Metropolitan Planning Strategy. Victorian Government.

- Takagi D, Ikeda K and Kawachi I. (2012) Neighborhood social capital and crime victimization: comparison of spatial regression analysis and hierarchical regression analysis. *Social Science and Medicine* 75: 1895-1902.
- Takagi D and Shimada T. (2019) A spatial regression analysis on the effect of neighborhood-level trust on cooperative behaviors: comparison with a multilevel regression analysis. *Frontiers in Psychology* 10: 2799.
- Thornton LE, Dieter-Schroers R, Lamb KE, et al. (2022) Operationalising the 20-minute neighbourhood. *International Journal of Behavioral Nutrition and Physical Activity* 19.
- Thornton LE, Pearce JR, Macdonald L, et al. (2012) Does the choice of neighbourhood supermarket access measure influence associations with individual-level fruit and vegetable consumption? A case study from Glasgow. *International Journal of Health Geographics* 11: 29.
- Tseng M, Thornton LE, Lamb KE, et al. (2014) Is neighbourhood obesogenicity associated with body mass index in women? Application of an obesogenicity index in socioeconomically disadvantaged neighbourhoods. *Health and Place* 30: 20-27.
- van Lenthe FJ and Mackenbach JP. (2002) Neighbourhood deprivation and overweight: the GLOBE study. *International Journal of Obesity* 26: 234-240.
- Victorian Government Department of Environment Land Water and Planning. (2020) *20-minute neighbourhoods: Create more inclusive, vibrant and healthy neighbourhoods*. Available at: <https://www.planning.vic.gov.au/policy-and-strategy/planning-for-melbourne/plan-melbourne/20-minute-neighbourhoods>.
- Zangger C. (2019) Making a place for space: Using spatial econometrics to model neighborhood effects. *Journal of Urban Affairs* 41: 1055-1080.

## TABLES

**Table 1.** Number of eligible addresses by city and neighbourhood type.

<b>Melbourne</b>	<b>Number of eligible addresses</b>	<b>Total number of mailed letters*</b>	<b>Total number of reminder letters issued</b>	<b>Total number of letters sent to unique address points**</b>	<b>Number of consenting participants (% response**)</b>
20MN Low SES	4365	7973	3860	4113	112 (2.7%)
20MN High SES	61187	4039	500	3539	204 (5.8%)
Non-20MN Low SES	30064	9786	0	9786	239 (2.4%)
Non-20MN High SES	24653	4292	500	3792	227 (6.0%)
Overall	120269	26090	4860	21230	782 (3.7%)
<b>Adelaide</b>					
20MN Low SES	4537	5701	1600	4101	147 (3.6%)
20MN High SES	24799	4200	0	4200	290 (6.9%)
Non-20MN Low SES	18118	7606	0	7606	216 (2.8%)
Non-20MN High SES	32333	2403	0	2403	177 (7.4%)
Overall	79787	19910	1600	18310	830 (4.5%)

\*Including reminder letters. \*\*From total number of mailed letters subtracting the number of reminder letters.

**Table 2.** Descriptive characteristics of ProjectPLAN participants by city.

	<b>Melbourne (N=637)</b>	<b>Adelaide (N=692)</b>	<b>Overall (N=1329)</b>
<b>Neighbourhood Status</b>			
20MN	254 (39.9%)	367 (53.0%)	621 (46.7%)
non-20MN	383 (60.1%)	325 (47.0%)	708 (53.3%)
<b>Neighbourhood SES</b>			
High	367 (57.6%)	401 (57.9%)	768 (57.8%)
Low	270 (42.4%)	291 (42.1%)	561 (42.2%)
<b>Gender</b>			
Female	365 (57.3%)	404 (58.4%)	769 (57.9%)
Male	272 (42.7%)	288 (41.6%)	560 (42.1%)
<b>Education</b>			
Less than university education	252 (39.6%)	371 (53.6%)	623 (46.9%)
University education	385 (60.4%)	321 (46.4%)	706 (53.1%)
<b>Ability to manage on income</b>			
Comfortable/Very Comfortable	416 (65.3%)	462 (66.8%)	878 (66.1%)
Just getting by	160 (25.1%)	162 (23.4%)	322 (24.2%)
Very difficult/Difficult	61 (9.6%)	68 (9.8%)	129 (9.7%)
<b>Age (years)</b>			
Mean (SD)	50.6 (16.4)	57.3 (15.7)	54.1 (16.4)
<b>Survey Type</b>			
Food	287 (45.1%)	351 (50.7%)	638 (48.0%)
Physical Activity	350 (54.9%)	341 (49.3%)	691 (52.0%)
<b>BMI (kg/m<sup>2</sup>)</b>			
Mean (SD)	25.4 (4.4)	26.8 (5.0)	26.2 (4.8)

**Table 3.** Moran's *I* test of spatial autocorrelation in BMI and the OLS model residuals.

		Melbourne (N = 637)		Adelaide (N =692)		
	Distance type	Weight type	Moran's <i>I</i>	<i>p</i> -value	Moran's <i>I</i>	<i>p</i> -value
BMI	Euclidean	Inverse-distance	0.008	0.357	0.026	0.001
		3-km bandwidth	0.032	0.029	0.051	<0.001
		Inverse-distance-squared	0.019	0.504	0.037	0.170
		1.5-km bandwidth	0.031	0.183	0.059	0.001
	Network	Inverse-distance	0.004	0.691	0.021	0.022
		3-km bandwidth	0.026	0.163	0.051	0.001
		Inverse-distance-squared	0.022	0.420	0.026	0.329
		1.5-km bandwidth	0.040	0.164	0.049	0.040
<b>Residual of unadjusted OLS regression model</b>	Euclidean	Inverse-distance	0.004	0.558	0.015	0.027
		3-km bandwidth	0.024	0.078	0.034	0.001
		Inverse-distance-squared	0.014	0.586	0.027	0.290
		1.5-km bandwidth	0.025	0.248	0.045	0.009
	Network	Inverse-distance	0.000	0.877	0.012	0.153
		3-km bandwidth	0.018	0.304	0.032	0.033
		Inverse-distance-squared	0.018	0.473	0.016	0.504
		1.5-km bandwidth	0.034	0.210	0.033	0.147
<b>Residual of adjusted OLS regression model*</b>	Euclidean	Inverse-distance	0.000	0.748	0.003	0.481
		3-km bandwidth	0.000	0.888	0.000	0.707
		Inverse-distance-squared	0.006	0.496	-0.002	0.939
		1.5-km bandwidth	-0.008	0.876	0.013	0.355
	Network	Inverse-distance	-0.003	0.993	0.004	0.511
		3-km bandwidth	-0.007	0.749	0.002	0.721
		Inverse-distance-squared	0.005	0.896	-0.007	0.922
		1.5-km bandwidth	-0.004	0.970	0.005	0.688

\*Models adjusted for neighbourhood SES, gender, age, education, ability to manage on income and survey type.

## FIGURES

**Figure 1.** Concept of the 3-km bandwidth spatial weight matrix.

It is assumed that 1.5-km is the distance an adult can walk in 20 minutes, so participant pairs whose distances were less than 3-km may share the same facility.

**Figure 2.** Estimated difference in mean BMI with 95% confidence intervals by neighbourhood status from OLS regression and spatial error regression models by city.

OLS adjusted and all spatial weights models presented adjusted for neighbourhood SES, gender, age, education, ability to manage on income and survey type.