Social Science & Medicine 73 (2011) 6–12

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

School-based friendship networks and children's physical activity: A spatial analytical approach

Kyle Macdonald-Wallis*, Russell Jago, Angie S. Page, Rowan Brockman, Janice L. Thompson

Centre for Exercise, Nutrition & Health, School for Policy Studies, University of Bristol, Bristol BS8 1TP, UK

ARTICLE INFO

Article history:
Available online 18 May 2011

Keywords:
UK
Physical activity
Accelerometer
Friendship
Peer
Social networks
Spatial analysis
Autoregression

ABSTRACT

Despite the known health benefits, the majority of children do not meet physical activity guidelines, with past interventions to increase physical activity yielding little success. Social and friendship networks have been shown to influence obesity, smoking and academic achievement, and peer-led interventions have successfully reduced the uptake of adolescent smoking. However, the role of social networks on physical activity is not clear. This paper investigates the extent to which friendship networks influence children's physical activity, and attempts to quantify the association using spatial analytical techniques to account for the social influence.

Physical activity data were collected for 986 children, aged 10–11 years old, from 40 schools in Bristol, UK. Data from 559 children were used for analysis. Mean accelerometer counts per minute (CPM) and mean minutes of moderate to vigorous physical activity per day (MVPA) were calculated as objective measures of physical activity. Children nominated up to 4 school-friends, and school-based friendship networks were constructed from these nominations.

Networks were tested to assess whether physical activity showed spatial dependence (in terms of social proximity in social space) using Moran's I statistic. Spatial autoregressive modelling was then used to assess the extent of spatial dependence, whilst controlling for other known predictors of physical activity. This model was compared with linear regression models for improvement in goodness-of-fit.

Results indicated spatial autocorrelation of both mean MVPA ($I = .346$) and mean CPM ($I = .284$) in the data, indicating that children clustered in friendship groups with similar activity levels. Spatial autoregressive modelling of mean MVPA concurred that spatial dependence was present ($\rho = .26$, $p < .001$), and improved model fit by 31% on the linear regression model. These results demonstrate an association between physical activity levels of children and their school-friends, and indicate that spatial modelling is an informative method for incorporating the influence of school social structure into physical activity analysis.

Background

Physical activity is associated with improved physical and mental health among children and adolescents (Department of Health, 2004). The majority of children are not engaging in the recommended 60 min of moderate to vigorous physical activity (MVPA) per day (Butcher, Sallis, Mayer, & Woodruff, 2008; Deverill, Doyle, Erens, Falaschetti, Hedges, Malbut et al., 2003). Interventions have generally not been successful in increasing children's physical activity levels (van Sluijs, McMinn, & Griffin, 2007). It is therefore important to identify and be able to measure the key determinants of physical activity in childhood.

Past research on physical activity has identified that increased levels of physical activity are associated with social support from family and friends (Duncan, Duncan, & Strycker, 2005; Sallis, Prochaska, & Taylor, 2000; Strauss, Rodzilsky, Burack, & Colin, 2001; Zakarian, Hovell, Hofstetter, Sallis, & Keating, 1994), and that adolescents who engage in high levels of physical activity are more likely to encourage others to be physically active (Gentle, Caves, Armstrong, Balding, & Kirby, 1994). Past research into adolescent friendship networks showed male friends engaged in similar amounts of organised physical activity, whereas female friends engaged in similar levels of screen viewing behaviours (de la Haye, Robins, Mohr, & Wilson, 2009).

Social and friendship networks have been found to be associated with many health related outcomes including obesity (Christakis & Fowler, 2007; Valente, Fujimoto, Chou, & Spruijt-Metz, 2009),

* Corresponding author. Tel.: +44 0 117 33 11100.
E-mail address: K.Macdonald-Wallis@bris.ac.uk (K. Macdonald-Wallis).

0277-9536 © 2011 Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.socscimed.2011.04.018

smoking (Christakis & Fowler, 2008; Urberg, Degirmencioglu, & Pilgrim, 1997) and academic achievement (Steinberg, Dornbusch, & Brown, 1992). Furthermore, previous peer-led interventions (Campbell, Starkey, Holliday, Audrey, Bloor, Parry-Langdon et al., 2008) have been successful in reducing uptake of smoking in adolescents. Past research has also identified social structures in school environments in the form of friendship networks (Urberg, Degirmencioglu, Tolson, & Halliday-Scher, 1995), therefore it is plausible that development of physical activity interventions that account for the influence of friendship networks may lead to greater success in increasing children's activity levels.

Spatial analysis is an area of statistics, popular in a wide range of disciplines such as geography, sociology, epidemiology, biology and economics. Data is not only characterised by a set of independent observations, but also according to the relative position of each observation in a spatial frame. This may be in a relative geographical context, in terms of a measurable distance between observations, or in the context of a position within a system, be it social, ecological or mechanical. In the context of social networks, proximity is not measured by a measurable distance (such as Euclidian distance) but by the existence and strength of social ties (friendship and familial) between individuals in the network.

Spatial dependence in general is defined as the likelihood for objects in close relational proximity to influence or be associated with one another and have similar properties. It was described by Waldo Tobler in his first law of geography when he stated that "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Given the focus in this research on 'proximity' in social space, as described above, spatial dependence in a social network can be interpreted as the extent of similarity between individuals connected socially by friendship and family ties.

This paper utilises a number of spatial analytical techniques, namely spatial autocorrelation and spatial autoregression, to evaluate the level of spatial dependence of physical activity within school-based friendship networks of 10–11 year old children, with the intention of investigating whether children who are socially proximal as friends at school share similar physical activity levels, or whether individual level predictors are predominantly responsible for influencing physical activity.

Methods

Participants

This paper forms part of a larger study, the Bristol 3Ps Project (www.bristol.ac.uk/enhs/research/projects/bristol3ps.html), which investigated parent and peer influences on physical activity in 10–11 year old children. The sampling and recruitment methodology has been described in detail elsewhere (Jago, Fox, Page, Brockman, & Thompson, 2010). In summary, children aged 10–11 years were recruited from 40 primary schools in Bristol (UK), with data collected between April 2008 and March 2009. To be representative of the socioeconomic background of all schools in the city, schools were sampled proportionately from tertiles according to their Index of Multiple Deprivation (IMD) score, a measure of local neighbourhood deprivation, based on school postcode. The IMD score gives an indication of local deprivation, taking into account factors such as employment, income, health, education and crime levels. Higher scores indicate higher levels of deprivation.

The study was approved by the School of Applied Community and Health Studies Ethics committee at the University of Bristol (ref 004/08, Feb 2008) and informed parental consent and childhood assent were obtained for all participants.

Spatial model structure

Conventional (i.e. aspatial) modelling typically uses the linear regression model (Eq. (1)), where $\underline{y}$ is the response variable, $\boldsymbol{X}$ the set of independent variables to be fitted in the model, $\underline{\beta}$ the vector of parameter estimates and $\underline{\epsilon}$ the independent normally distributed error terms for each observation in the model.

$$\underline{y} = \boldsymbol{X}\underline{\beta} + \underline{\epsilon} \quad \underline{\epsilon} \sim N\left(0, \boldsymbol{I}\sigma^2\right) \tag{1}$$

This model does not however, reflect the possible existence of a correlation between individual observations in close social proximity (though not necessarily geographical proximity), (referred to here as spatial dependence). This omission is important because spatial dependence of observations in the social network in the model may affect estimation of the independent parameters, and thus may not reflect the true association between the outcome $\underline{y}$ vand parameters $\underline{\beta}$.

As explained above, spatial modelling in the research reported here adds an additional term to the model equation to account for this possible effect of proximity in the space of the social network. The spatial models discussed in this paper take two forms. The first is the mixed regressive spatial autoregressive model, also known as the 'spatial lag' model (Eq. (2)), combining conventional independent variables as seen in the linear regression model with an additional spatial term. It involves a measure of the dependent variables of neighbouring observations, weighted by a spatial weights matrix $\boldsymbol{W}$, which is scaled by an estimated parameter ρ. An interpretation of the spatial parameter in this model is that it is a weighted measure of the surrounding observations, and by controlling for similarity in neighbouring individuals, allows more accurate independent parameter estimates to be obtained. The spatial lag parameter ρ also gives an indication of the extent of spatial correlation between observations, and thus allows for easy interpretation of the fitted model.

$$\underline{y} = \rho\boldsymbol{W}\underline{y} + \boldsymbol{X}\underline{\beta} + \underline{\epsilon} \qquad \underline{\epsilon} \sim N\left(0, \boldsymbol{I}\sigma^2\right) \tag{2}$$

The second model is the 'spatial error' model (Eq. (3)) which is primarily used when spatial dependence is believed to be present, but when a spatial lag parameter does not appear to significantly improve model fit. The main structure of the model is consistent with the linear regression model, but it is the error term $\underline{\epsilon}$ that is adjusted to account for spatial dependence. Again the model is intended to give more accurate $\underline{\beta}$ parameter estimates and improved model fit, and interpretation of the spatial error parameter $\underline{\lambda}$ gives an indication of the nature of spatial dependence.

$$\underline{y} = \boldsymbol{X}\underline{\beta} + \underline{\epsilon} \qquad \underline{\epsilon} = \lambda\boldsymbol{W}\underline{\epsilon} + \underline{u} \quad \underline{u} \sim N\left(0, \boldsymbol{I}\sigma^2\right) \tag{3}$$

In the context of this analysis, the outcome variable $\underline{y}$ in the spatial model is an objective measure of physical activity, collected by means of accelerometers, with $\boldsymbol{X}$ denoting other potential covariates of physical activity. The data describing the friendship network is represented by the spatial weights matrix $\boldsymbol{W}$. These three components of the model are each described below.

Physical activity measure

Of the 1684 children invited to participate in the study, 986 (58.6%) consented to provide data for analysis. Physical activity was assessed by asking these participants to wear a GT1 M accelerometer (MTI, Florida) for 5 consecutive days. The accelerometers were set to capture a measure of body motion, or counts, at 10 s intervals. A valid day of data was considered to be when at least 500 min worth of data were obtained, after removal of substantial (≥60 min) periods of zero readings, which were interpreted as time when the accelerometer was not worn (Troiano et al., 2008). Participants were included in analysis if 3 days of valid data were obtained, with 747 (75.8%) of the 986 participants providing data which met these criteria. The accelerometer data were used to derive two physical activity variables. The mean number of counts per minute (Mean CPM) was calculated as an overall measure of volume of physical activity. The mean minutes per day of moderate to vigorous physical activity (Mean MVPA) was calculated as the average amount of time per day accelerometer counts exceed 3200 cpm (Puyau, Adolph, Vohra, & Butte, 2002). This threshold however was determined on older 7164 accelerometers, and with GT1 M monitors giving 9% lower counts (Corder et al., 2007) this threshold was corrected by a factor of .91–2912 cpm.

Model covariates

The aim of this paper was to assess whether spatial approaches can improve our understanding of associations between friendship networks and physical activity. Therefore it was important to model baseline associations between key predictors and children's physical activity, to test if understanding is enhanced by using spatial models compared with traditional aspatial approaches. Physical activity self-efficacy (PASE) was included in the model as the key predictor variable because it is a consistent and strong correlate of children's physical activity (Annesi, 2006; Biddle, Gorely, & Stensel, 2004). PASE was assessed using a validated, self reported set of 6 questions and converted to a summary variable on a continuous scale of 1–4, with higher scores indicating a higher level of self-efficacy (Beets, Pitetti, & Forlaw, 2007). Children's physical activity patterns also differ by physical traits such as gender (Jago, Anderson, Baranowski, & Watson, 2005), BMI and pubertal status (Baker, Birch, Trost, & Davison, 2007), and so the models also controlled for these variables. Pubertal status was self reported and summarised in 3 stages similar to Tanner stages. Height and weight were measured and body mass index (kg/m^2) calculated, which was then converted to a nationally representative age- and gender-specific standard deviation score (BMI SDS). Finally to control for possible socioeconomic effects on physical activity (Inchley, Currie, Todd, Akhtar, & Currie, 2005), IMD score was included as a potential model confounder (because friends might share aspects of deprivation important for physical activity), and was calculated based on the child's home postcode, as a measure of home neighbourhood deprivation (Noble et al., 2007).

The spatial weights matrix

Friendship groupings were assessed via a child-completed social network questionnaire, developed and validated by the Trial of Activity in Adolescent Girls (TAAG) (Voorhees, Murray, Welk, Birnbaum, Ribisl, Johnson et al., 2005). Participants were asked to identify up to 4 of their closest friends in their school ('Best friend', 'Friend 2', 'Friend 3', 'Friend 4') and these nominations were matched up with the data of other children in the study, if they were also participating. Of the 3751 friendship nominations made, 2569 (68.5%) related to other participants in the study, providing sufficient data to develop a network of friendship connections for each school.

This network information can be summarised in the form of a square matrix, with each row (and column) corresponding to an individual child in the study. The values contained in this matrix represent the extent to which children are socially connected with one another, with a zero recorded if no social connection exists. There are a number of different ways in which members of a social network may be considered to be connected, and there are a large number of different spatial weight matrices that may be used (Leenders, 2002). A series of different matrices were proposed for analysis, which could then be tested and compared for suitability to represent a school-based social network.

The simplest spatial weight matrix, used commonly in spatial analysis, is known as a contiguity matrix (Eq. (4)), and quantifies when two individuals are immediately connected by friendship in the network, and assumes that any other individuals are not socially connected (Anselin, 1988).

$$\mathrm{W_{i,j}} = \begin{cases} 1 & \text{if i nominates j as a friend (or vice versa)} \\ 0 & \text{if not friends} \end{cases} \tag{4}$$

This matrix may be seen as the most basic summary of friendships within a social network, but more distant friendships, such as friends-of-friends, are not considered in this construction. Since past studies have demonstrated that health and lifestyle influences extend beyond immediate friends to friends-of-friends and further (Christakis & Fowler, 2007, 2008), it was also important to consider matrices that account for the social distance between individuals. Incorporating social distance into a spatial matrix requires the calculation of $d_{i,j}$ denoting the geodesic or shortest social distance along friendship ties between individuals. Immediate friends connected directly have $d_i j = 1$, whilst friends-of-friends, connected through two ties (via another individual) would have $d_{i,j} = 2$, and so on up to the point that for unconnected pairs $d_i j = \infty$. Given that extremely distant friends are unlikely to influence one another, a variable maximum social distance D was used such that $d_{i,j}{}^{(D)} = \infty$ if $d_i j > D$.

It has been previously suggested that the effect of changing social distance decreases as the social distance between individuals decreases (White, 1983), and thus a weight matrix of inverted social distance was selected to use in analysis (Equation (5)). The weight assigned to each individual corresponded to the level of influence they hold within the spatial autoregressive model. This chosen weight matrix was an inverse function of the social (geodesic) distance between individuals. This meant that immediate (i.e. nominated) friends received the highest weighting (1/2), followed by friends-of-friends (1/3) and so on with the weighting decreasing with increasing social distance. The inverse nature of this function also means that as social distance increases, the rate of decrease in weighting also reduces. This reflects the assumption that the difference in extent of influence between immediate friends and friends-of-friends (1/2 vs. 1/3) is likely to be greater than the difference in influence between more socially distant friends, for example 6th and 7th degree friends (1/7 vs. 1/8). It should be noted that there are many different constructions of spatial and social weights (Getis, 1984; Leenders, 2002), however the above weights matrix is a reasonable assumption of the possible social influence in school-based friendship networks, and thus was justified for use in this analysis.

$$W_{i,j}^{(D)} = \begin{cases} \dfrac{1}{1+d_{i,j}} & \mathit{if}\ d_{i,j} \le D \\ 0 & \mathit{if}\ d_{i,j} > D \end{cases} \tag{5}$$

Statistical analysis

Participants were excluded from the analysis if they failed to provide complete physical activity data or were isolated in the social network, that is, they had no friendship ties with any other valid participant in the sample. Descriptive statistics for all variables and friendship nominations were calculated and independent sample *t*-tests and χ^2 tests were used to examine whether there

Table 1
Descriptive statistics for participants with valid data.

		All ($n = 559$)		Girls ($n = 312$)		Boys ($n = 247$)		P[1]
		Mean	SD	Mean	SD	Mean	SD	
Mean Minutes MVPA per Day		36.03	17.45	30.61	13.43	42.86	19.45	**<.001**
Physical Activity Self-Efficacy Score		3.19	.53	3.18	.53	3.20	.54	.634
BMI SDS		.45	1.16	.39	1.17	.53	1.15	.183
Index Multiple Deprivation Score		22.07	16.92	21.47	16.60	22.82	17.32	.351
		N	%	N	%	N	%	P[2]
Tanner Stages	1-2	276	49.4	127	40.7	149	60.3	**<.001**
	3	238	42.6	153	49.0	85	34.4	
	4-5	45	8.0	32	10.3	13	5.3	

Bolded values = $P < .05$.
P[1] = Independent sample *t*-tests for difference by gender.
P[2] = Chi-squared test for difference in pubertal development by gender.

was any difference in physical activity or demographic characteristics of participants who were retained or excluded from analysis.

Friendship nominations for each participant were combined to generate a social network for each school. These networks were graphed using the NodeXL add-in within Microsoft Excel (Smith et al., 2009), with each node representing an individual child in the study. The size of each node was scaled to represent the child's Mean MVPA. The extent of clustering of similar sized nodes could then be visually assessed, thus giving a measure of spatial dependence (clustering of behaviour within groups of people who were socially close to each other).

As the network graphs indicated a possible spatial dependence of physical activity Moran's I statistic (Moran, 1950) was used to calculate the spatial autocorrelation for Mean MVPA minutes and Mean CPM. This test used a simple contiguity spatial weights matrix. Moran's I gave an indication of the level of social clustering of individuals with similar physical activity levels, and thus whether spatial analysis was appropriate for this data. To examine the role of the wider social network, Moran's I statistic was also calculated for larger degrees of separation for Mean MVPA and Mean CPM.

To assess the effect of spatial modelling on model fit and parameters, linear regression models were constructed using the confounders detailed in the previous section as baseline models. As discussed in the model development section, these baseline models predicted Mean MVPA and CPM with physical activity self-efficacy as the key predictor and adjusting for gender, BMI SDS, IMD score and pubertal stage. To examine whether spatial modelling was appropriate for predicting Mean MVPA or Mean CPM, spatial diagnostic tests were performed on this model using the 'spatdiag' test in Stata. This test calculates Lagrange multiplier statistics, indicating whether either of the spatial models discussed in the previous section are appropriate for the data given the baseline model and a proposed spatial weights matrix. In this paper, Robust Lagrange multipliers (Anselin, Bera, Florax, & Yoon, 1996) are used, which are useful for comparison of spatial model types, as they are more accurate in model selection than standard Lagrange multipliers.

Diagnostic tests indicated that a spatial lag model (Eq. (2)) with friendships up to the second degree was most appropriate for predicting Mean MVPA (see below). This spatial autoregressive model was then estimated using the 'spatreg' command in Stata. A likelihood ratio test was used to assess whether the more complex spatial lag model had significantly improved model fit from the baseline linear regression model. All analyses were conducted in Stata 10.1 (College Station, Texas).

Results

The sample of valid data used for analysis was shown to be representative of the overall dataset for all variables with no differences in the physical activity, BMI SDS, IMD or pubertal stage between the included and excluded participants. Descriptive statistics for the valid data are shown split by gender in Table 1. Tests showed that boys engaged in more minutes of MVPA per day than girls (42.9 vs. 30.6, $p < .001$).

Social network graphs from three schools are plotted for illustrative purposes in Fig. 1, with each node representing individual participants sized by their Mean MVPA. It was apparent that similar sized nodes were frequently grouped together in friendship clusters, indicating that children with similar amounts of Mean MVPA per day were closer friends within the social network. This may suggest that spatial autocorrelation of Mean MVPA may be present. Similar patterns of varying strengths were evident for the majority of the 40 schools (data not shown). The Moran's I calculations presented in Table 2 confirms the pattern in Fig. 1, by showing that measures of physical activity are positively and significantly spatially autocorrelated, with Mean MVPA being the most autocorrelated (I = .346, $p < .001$) followed by Mean CPM (I = .284, $p < .001$). The results therefore indicated that spatial analysis on the social network data was appropriate.

Moran's I statistics for Mean MVPA and Mean CPM with increasing friendship distance are displayed in Table 3. Spatial autocorrelation is seen to decrease as the maximum number of

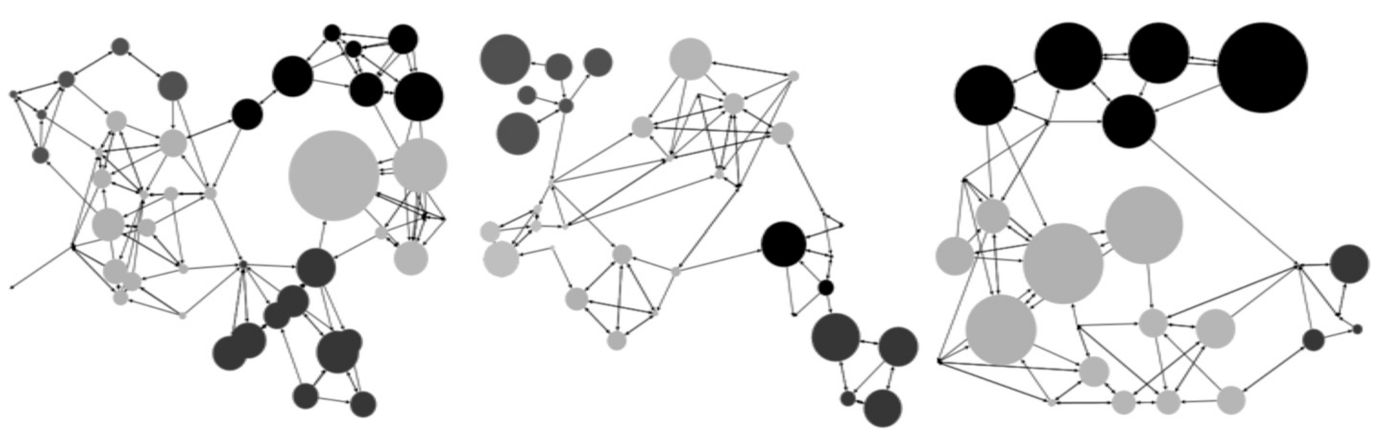

Fig. 1. Social networks graphs of three schools – node size corresponds to child's mean minutes MVPA per day. Key: Nodes shaded by friendship group sub-clusters derived in NodeXL.

Table 2
Moran's I statistic for spatial autocorrelation of immediate friends' physical activity characteristics ($n = 559$).

	Moran's I	Z	P^3
Mean minutes MVPA per day	.346	9.89	**<.001**
Mean Counts per Minute	.284	8.14	**<.001**

Bolded values = $P < .05$.
P^3 = Test for difference in Moran's I statistic from expected mean = −.002.

degrees of separation increases for both MVPA and CPM, with MVPA falling to .218 ($p < .001$) and CPM to .186 ($p < .001$) with up to fifth degree friends included, indicating that similarity in physical activity levels decreases as friendship separation increases.

Table 4 details the results of the spatial diagnostic tests on the series of weight matrices proposed earlier for both Mean MVPA and Mean CPM. It indicates that for a basic linear regression model predicting Mean MVPA by gender, physical activity self-efficacy, BMI, IMD score and pubertal development, the most appropriate spatial model was a spatial lag model, with weight matrix extending to two degrees of separation ($p = .034$). For Mean CPM however, although the same model and weight matrix were indicated to be the most suitable, the Lagrange multipliers were not significant so does not justify spatial modelling for any maximum degree of separation ($p = .185$).

This spatial lag model was then estimated using the 'spatreg' command in Stata, and was statistically compared with the baseline regression model estimated using OLS regression. Results and model comparison are detailed in Table 5. The baseline OLS regression model finds that Mean MVPA is predicted by gender ($p < .001$), self-efficacy ($p = .006$), BMI ($p < .001$), IMD score ($p = .007$) and pubertal status ($p = .005$) in a model that accounts for 17.8% of total variance. The spatial lag model improves on this baseline fit, with 23.4% of variance accounted for by gender ($p < .001$), self-efficacy ($p = .007$), BMI ($p < .001$), pubertal status ($p = .010$ and .065) and the spatial autoregressive parameter ($p < .001$). The spatial autoregressive parameter yielded a coefficient of .26, indicating that an individual's Mean MVPA is positively associated with the Mean MVPA of their first and second degree friends. Spatial modelling was seen to improve model R^2 by 5.6 percentage points (a 31.4% increase in the variance accounted for by the OLS model). A likelihood ratio test conducted to test this improvement indicated that the spatial lag parameter significantly increased model fit ($\lambda = 28.60$, $p < .001$).

Discussion

The results described in this paper indicate that there is spatial autocorrelation of children's physical activity in school-based friendship networks. The data showed strong evidence of clustering of children with similar activity levels when plotted and visually examined. Moran's I statistic supported this by detecting positive and significant spatial autocorrelation for all physical

Table 3
Moran's I statistic for spatial autocorrelation of mean minutes MVPA and mean CPM with increasing degrees of separation included in the weight matrix $W(D)$ ($n = 559$).

	Moran's I	Z	P^3
Mean MVPA minutes per day			
$W^{(1)}$ – Immediate friends only	.346	9.89	**<.001**
$W^{(2)}$ – Second degree friends included	.292	11.47	**<.001**
$W^{(3)}$ – Third degree friends included	.260	11.96	**<.001**
$W^{(4)}$ – Fourth degree friends included	.236	11.78	**<.001**
$W^{(5)}$ – Fifth degree friends included	.218	11.47	**<.001**
Mean Counts per Minute			
$W^{(1)}$ – Immediate friends only	.284	8.14	**<.001**
$W^{(2)}$ – Second degree friends included	.242	9.51	**<.001**
$W^{(3)}$ – Third degree friends included	.216	9.94	**<.001**
$W^{(4)}$ – Fourth degree friends included	.198	9.95	**<.001**
$W^{(5)}$ – Fifth degree friends included	.186	9.81	**<.001**

Bolded values = $P < .05$.
P^3 = Test for difference in Moran's I statistic from expected mean = −.002.

Table 4
Spatial diagnostic test for appropriate spatial model for predicting mean MVPA and mean CPM, given baseline OLS regression model[a] ($n = 559$).

	Spatial Error Model		Spatial Lag Model	
	Robust Lagrange Multiplier	P^4	Robust Lagrange Multiplier	P^4
Mean MVPA Minutes per Day				
$W^{(1)}$ – Immediate friends only	.047	.827	2.669	.102
$W^{(2)}$ – Second degree friends included	.098	.755	4.445	**.035**
$W^{(3)}$ – Third degree friends included	.053	.817	3.430	.064
$W^{(4)}$ – Fourth degree friends included	.677	.411	1.856	.173
$W^{(5)}$ – Fifth degree friends included	.596	.440	2.063	.151
Mean Counts per Minute				
$W^{(1)}$ – Immediate friends only	.049	.826	.866	.352
$W^{(2)}$ – Second degree friends included	.022	.883	1.754	.185
$W^{(3)}$ – Third degree friends included	.331	.565	1.203	.273
$W^{(4)}$ – Fourth degree friends included	1.033	.309	.536	.464
$W^{(5)}$ – Fifth degree friends included	.845	.358	.741	.389

Bolded values = $P < .05$.
P^4 = Test for significance of Lagrange Multiplier against χ^2_1 distribution.
[a] Baseline regression model predicts mean MVPA by gender, physical activity self-efficacy, BMI, IMD score and pubertal status.

Table 5
Comparison of baseline regression model with spatial lag model using weight matrix $W^{(2)}$ for predicting Mean MVPA.

Baseline OLS regression model ($n = 559$)	Coefficient	95% CI	t	P
Gender (Ref: Female)				
Male	13.16	10.43 to 15.89	9.48	**<.001**
PA Self-Efficacy	3.48	1.00 to 5.97	2.76	**.006**
BMI SDS	−2.08	−3.23 to −.93	−3.56	**<.001**
Index Multiple Deprivation Score	.11	.03 to .19	2.69	**.007**
Tanner Stage (Ref: Stage 1-2)				
Stage 3	4.09	1.28 to 6.91	2.86	**.004**
Stage 4-5	5.16	.05 to 10.28	1.98	**.048**
Constant	15.52	7.18 to 23.86	3.65	**<.001**
Model $R^2 = 0.178$				
Model Log Likelihood (L_0) = −2336.39				
Spatial Autoregressive Lag Model with matrix $W^{(2)}$ ($n = 559$)	Coefficient	95% CI	z	P
Gender (Ref: Female)				
Male	10.74	8.00 to 13.49	7.67	**<.001**
PA Self-Efficacy	3.29	.91 to 5.67	2.71	**.007**
BMI SDS	−2.00	−3.10 to −.90	−3.56	**<.001**
Index Multiple Deprivation Score	.06	−.02 to.14	1.53	.125
Tanner Stage (Ref: Stage 1–2)				
Stage 3	3.53	.83 to 6.23	2.56	**.010**
Stage 4-5	4.62	−.28 to 9.52	1.85	.065
Constant	8.93	.61 to 17.24	2.10	**.035**
ρ (spatial autoregressive parameter)	.26	.17 to .36	5.60	**<.001**
Model $R^2 = 0.234$				
Model Log Likelihood (L_1) = −2322.10				
Likelihood Ratio Test for spatial autoregressive parameter ρ			λ	P^5
			28.60	**<.001**

Bolded values = $P < .05$.
P^5 = Test of $\lambda = -2(L_0 - L_1)$ against χ^2_1 distribution.

activity measures, indicating that children report being friends with children who have similar activity profiles. Spatial autocorrelation of Mean MVPA being greater than that of Mean CPM indicates that social networks are more likely to influence amounts of higher intensity physical activity than overall physical activity levels. This may be because friends are more likely to take part in more structured activities such as team based activities together or that being physically active with friends stimulates more intense play. These results concur with findings from the TAAG study (Voorhees et al., 2005), which found that girls who have more physically active friends reported being more active themselves, and also supports findings that close adolescent friends reported engaging in similar amounts of organised physical activity (de la Haye et al., 2009). The results also provide evidence that overall physical activity levels, including informal activity and physical activity outside of school, are similar among close friends. The fact that spatial autocorrelation of both Mean MVPA and Mean CPM decreased as the scope friendship separation increased indicates that closer friends share more similar physical activity characteristics than more distant friends.

Having been justified by diagnostic tests, spatial autoregression on the data demonstrated that after controlling for known predictors of child physical activity, Mean MVPA is positively and significantly associated with the Mean MVPA of their immediate and second degree friends. Furthermore, the spatial lag model was highly effective in improving model fit from the baseline OLS model, increasing it by 33%. The results of the autoregression models indicate that even after controlling for known predictors of physical activity in children, spatial dependence of physical activity still existed within the social network. As a result it can be concluded that the structure of friendship networks itself has an important association with physical activity, beyond that of individual level predictors such as gender, self-efficacy, BMI and socioeconomic status.

It is important to note that this spatial analysis does not give explicit indication as to why children who are proximal in social network are similar in their activity levels. Similarities may be due to the influence of individuals on their friends, or the fact that children are more likely to develop friendships with others that share similar traits, known as ‘homophily’ (McPherson, Smith-Lovin, & Cook, 2001). It could be that individual level factors known to influence physical activity may also influence the formation of friendship networks (such as children with similar BMI or self-efficacy being more likely to form friendships). Nevertheless, this analysis has identified friendship group influences of physical activity and these effects are shown to be more than simply a function of an individual’s characteristics. Thus children’s social circles are likely to affect their quantities of physical activity and sedentary time.

As the data reported here are from a cross-sectional study, it is difficult to infer whether spatial dependence resulted from selection into friendship groups of those whose established behaviour is similar, or influence of friends upon individual behaviour. Past research on health behaviours such as smoking (Mercken, Snijders, Steglich, Vertiainen, & de Vries, 2010) and dieting behaviours (Woelders, Larsen, Scholte, Cillessen, & Engels, 2010) have identified that the effect is likely to be a combination of both. However an extension of this analysis would be to incorporate longitudinal data, to assess the change in physical activity whilst evaluating the change in social network structure of participants, giving insight into the relative contributions of friendship influence and selection.

The findings of this analysis have implications for future analysis of physical activity, and the design of future interventions in this field, which must consider the influence of friendship networks in order to be most effective. As previously discussed, targeting peer networks to change behaviour has shown to be effective in adolescent interventions to delay onset of smoking (Campbell et al., 2008). Physical activity interventions may also benefit from this social network approach by, for example using peer members to increase physical activity within their friendship groups, or by targeting low physically active people at the friendship group level rather than individuals at risk.

Finally these results demonstrate the benefits of using spatial techniques in the analysis of friendship groups in all areas of adolescent health research, limiting any potential confounding from friendship influences on positive and negative health behaviours. Our findings indicate that datasets should be tested for the presence of spatial dependence and, where appropriate spatial dependence should be controlled for in analyses. Results demonstrated spatial analysis is a viable option for modelling and investigating this dependence, as it improved model fit. Moreover with the spatial lag model, the parameter ρ may also be interpreted as a parameter of social influence. Given the relatively low cost of incorporating social network questions in data collection and ease of use in analysis, it may be beneficial in other areas of adolescent health research.

Limitations

Whilst findings from this analysis demonstrate the association of school-based friendship groups with objectively measured physical activity, as discussed previously, the cross-sectional nature of the survey data collected means that it is not possible to say whether friends influence each other’s physical activity or whether children befriend others with similar activity levels. In environments such as schools, where the formations of social networks is largely voluntary (de Klepper, Sleebos, van de Bunt, & Agneessens, 2010), homophily is present in the network structure. As a result determining the path of causation is very difficult without the collection of longitudinal data.

As samples were taken from a school setting, only school-based friendships were included in analysis. We have previously reported that many children are friends both with children at school and children from their neighbourhoods who may not attend the same school, and due to the absence of data we are unable to make comparisons on influence of school and neighbourhood friends (Jago et al., 2009).

Due to requirements for symmetrical spatial weight matrices for modelling, analysis in this paper assumed that if one child nominated another as a friend, then this perception of friendship would be reciprocated by the other person. This assumption of undirected friendship ties allowed spatial autoregression to be performed in Stata, but is not necessarily correct, as friendships may only be perceived by one child, but not reciprocated by the other. It would be of interest to expand spatial analysis to consider the influence of directed friendship ties to investigate whether association is stronger in the direction of nomination, as has been found with the obesity risk (Christakis & Fowler, 2007).

Conclusions

The results of this analysis further add to the evidence that school-based friendship groups are associated with physical activity. Spatial lag models showed that a child’s mean minutes of MVPA per day was positively and significantly correlated with that of their immediate and secondary school friends, after controlling for other known predictors of physical activity. Spatial modelling also significantly improved model fit in comparison to conventional linear regression models. These data would support the inclusion of social network questionnaires in other studies investigating the correlates of physical activity in children of this age.

Acknowledgements

This project was funded by a project grant from the British Heart Foundation (ref PG/06/142). This report is also arising from a Career Development Fellowship (to Dr. Jago) supported by the National Institute for Health Research. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health. The authors thank all of the children and schools that participated in the study.

References

Annesi, J. J. (2006). Relations of physical self-concept and self-efficacy with frequency of voluntary physical activity in preadolescents: implications for after-school care programming. *Journal of Psychosomatic Research, 61*(4), 515–520.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Norwell, MA: Kluwer Academic Publishers.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77–104.

Baker, B. L., Birch, L. L., Trost, S. G., & Davison, K. K. (2007). Advanced pubertal status at age 11 and lower physical activity in adolescent girls. *Journal of Pediatrics, 151*(5), 488–493.

Beets, M. W., Pitetti, K. H., & Forlaw, L. (2007). The role of self-efficacy and referent specific social support in promoting rural adolescent girls' physical activity. *American Journal of Health Behavior, 31*(3), 227–237.

Biddle, S. J. H., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of Sports Sciences, 22*(8), 679–701.

Butcher, K., Sallis, J. F., Mayer, J. A., & Woodruff, S. (2008). Correlates of physical activity guideline compliance for adolescents in 100 US cities. *Journal of Adolescent Health, 42*(4), 360–368.

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., et al. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial. *Lancet, 371*(9624), 1595–1602.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370–379.

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine, 358*(21), 2249–2258.

Corder, K., Brage, S., Ramachandran, A., Snehalatha, C., Wareham, N., & Ekelund, U. (2007). Comparison of two Actigraph models for assessing free-living physical activity in Indian adolescents. *Journal of Sports Sciences, 25*(14), 1607–1611.

de Klepper, M., Sleebos, E., van de Bunt, G., & Agneessens, F. (2010). Similarity in friendship networks: selection or influence? The effect of constraining contexts and non-visible individual attributes. *Social Networks, 32*, 82–90.

de la Haye, K., Robins, G., Mohr, P., & Wilson, C. (2009). Obesity-related behaviors in adolescent friendship networks. *Social Networks*.

Department of Health. (2004). *At least five a week: Physical activity and health outcomes: A review of the Chief Medical Officer*. London: Department of Health.

Deverill, C., Doyle, M., Erens, B., Falaschetti, E., Hedges, B., Malbut, K., et al. (2003). *The health of children and Young people, the health survey for England 2002*. London: The Stationary Office.

Duncan, S. C., Duncan, T. E., & Strycker, L. A. (2005). Sources and types of social support in youth physical activity. *Health Psychology, 24*(1), 3–10.

Gentle, P., Caves, R., Armstrong, N., Balding, J., & Kirby, B. (1994). High and low Exercisers among 14-Year-Old and 15-Year-Old children. *Journal of Public Health Medicine, 16*(2), 186–194.

Getis, A. (1984). Interaction modeling using 2nd-Order analysis. *Environment and Planning A, 16*(2), 173–183.

Inchley, J. C., Currie, D. B., Todd, J. M., Akhtar, P. C., & Currie, C. E. (2005). Persistent socio-demographic differences in physical activity among Scottish schoolchildren 1990-2002. *European Journal of Public Health, 15*(4), 386–388.

Jago, R., Anderson, C. B., Baranowski, T., & Watson, K. (2005). Adolescent patterns of physical activity – Differences by gender, day, and time of day. *American Journal of Preventive Medicine, 28*(5), 447–452.

Jago, R., Brockman, R., Fox, K. R., Cartwright, K., Page, A. S., & Thompson, J. L. (2009). Friendship groups and physical activity: qualitative findings on how physical activity is initiated and maintained among 10-11 year old children. *International Journal of Behavioural Nutrition and Physical Activity, 6*, 4.

Jago, R., Fox, K. R., Page, A. S., Brockman, R., & Thompson, J. L. (2010). Parent and child physical activity and sedentary time: do active parents foster active children? *BMC Public Health, 10*, 194.

Leenders, R. T. A. J. (2002). Modeling social influence through network autocorrelation: constructing the weight matrix. *Social Networks, 24*(1), 21–47.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology, 27*, 415–444.

Mercken, L., Snijders, T. A. B., Steglich, C., Vertiainen, E., & de Vries, H. (2010). Smoking-based selection and influence in gender-segregated friendship networks: a social network analysis of adolescent smoking. *Addiction, 105*(7), 1280–1289.

Moran, P. A. P. (1950). Notes on continuous Stochastic Phenomena. *Biometrika, 37*(1–2), 17–23.

Noble, M., McLennan, D., Wilkinson, K., Whitworth, A., Barnes, H., & Dibben, C. (2007). *The English Indices of deprivation*. London: Communities and Local Government.

Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity Research, 10*(3), 150–157.

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medicine and Science in Sports and Exercise, 32*(5), 963–975.

Smith, M. A., Schneiderman, B., Milic-Frayling, N., Mendes Rodrigues, E., Barash, V., Dunne, C., et al. (2009). *Analyzing (social media) networks with NodeXL*In *Proceedings of the fourth international conference on communities and technologies*, 255–264.

Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic-Differences in adolescent achievement - an ecological Perspective. *American Psychologist, 47*(6), 723–729.

Strauss, R. S., Rodzilsky, D., Burack, G., & Colin, M. (2001). Psychosocial correlates of physical activity in healthy children. *Archives of Pediatrics & Adolescent Medicine, 155*(8), 897–902.

Tobler, W. R. (1970). Computer Movie Simulating Urban Growth in Detroit Region. *Economic Geography, 46*(2), 234–240.

Troiano, R. P., Berrigan, D., Dodd, K. W., Masse, L. C., Tilert, T., & Mcdowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise, 40*(1), 181–188.

Urberg, K. A., Degirmencioglu, S. M., & Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology, 33*(5), 834–844.

Urberg, K. A., Degirmencioglu, S. M., Tolson, J. M., & Halliday-Scher, K. (1995). The structure of adolescent peer networks. *Developmental Psychology, 31*(4), 540–547.

Valente, T. W., Fujimoto, K., Chou, C. P., & Spruijt-Metz, D. (2009). Adolescent Affiliations and Adiposity: a social network analysis of friendships and obesity. *Journal of Adolescent Health, 45*(2), 202–204.

van Sluijs, E. M. F., McMinn, A. M., & Griffin, S. (2007). Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *British Medical Journal, 335*(7622), 703–707.

Voorhees, C. C., Murray, D., Welk, G., Birnbaum, A., Ribisl, K. M., Johnson, C. C., et al. (2005). The role of peer social network factors and physical activity in adolescent girls. *American Journal of Health Behavior, 29*(2), 183–190.

White, M. J. (1983). The Measurement of spatial Segregation. *American Journal of Sociology, 88*(5), 1008–1018.

Woelders, L. C. S., Larsen, J. K., Scholte, R. H. J., Cillessen, A. H. N., & Engels, R. C. M. E. (2010). Friendship group influences on body Dissatisfaction and dieting among adolescent girls: a Prospective study. *Journal of Adolescent Health, 47*(5), 456–462.

Zakarian, J. M., Hovell, M. F., Hofstetter, C. R., Sallis, J. F., & Keating, K. J. (1994). Correlates of vigorous Exercise in a predominantly low Ses and Minority high-school Population. *Preventive Medicine, 23*(3), 314–321.