Journal of Transport & Health 2 (2015) 219–229

Contents lists available at ScienceDirect

Journal of Transport & Health

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

Are GIS-modelled routes a useful proxy for the actual routes followed by commuters?

Alice M. Dalton [a,b,*], Andrew P. Jones [a,b], Jenna Panter [b,c], David Ogilvie [b,c]

[a] Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7 TJ, UK
[b] UKCRC Centre for Diet and Activity Research (CEDAR), University of Cambridge, Cambridge, UK
[c] Medical Research Council Epidemiology Unit, University of Cambridge, Cambridge, UK

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form
29 September 2014
Accepted 1 October 2014
Available online 3 November 2014

Keywords:
Global positioning system
Geographic information system
Commute environment
Active travel

ABSTRACT

Active commuting offers the potential to increase physical activity among adults by being built into daily routines. Characteristics of the route to work may influence propensity to walk or cycle. Geographic information system (GIS) software is often used to explore this by modelling routes between home and work. However, if the validity of modelled routes depends on the mode of travel used, studies of environmental determinants of travel may be biased.

We aimed to understand how well modelled routes reflect those actually taken, and what characteristics explain these differences. We compared modelled GIS shortest path routes with actual routes measured using QStarz BT-Q1000X global positioning system (GPS) devices in a free-living sample of adults working in Cambridge and using varying travel modes. Predictors of differences, according to length and percentage overlap, between the two route sets were assessed using multilevel regression models and concordance coefficients.

The 276 trips, made by 51 participants, were on average 27% further than modelled routes, with an average geographical overlap of 39%. However, predictability of the route depended on travel mode. For route length, there was moderate-to-substantial agreement for journeys made on foot and by bicycle. Route overlap was lowest for trips made by car plus walk (22%). The magnitude of difference depended on other journey characteristics, including travelling via intermediate destinations, distance, and use of busy roads.

In conclusion, GIS routes may be acceptable for distance estimation and to explore potential routes, particularly active commuting. However, GPS should be used to obtain accurate estimates of environmental contexts in which commuting behaviour actually occurs. Public health researchers should bear these considerations in mind when studying the geographical determinants and health implications of commuting behaviour, and when recommending policy changes to encourage active travel.

1. Introduction

Active commuting, walking or cycling some or all of the journey to and from work, offers considerable potential to increase the prevalence of activity amongst adults as it can be built into daily routines (Department of Health, 2009). However, the prevalence of active travel has declined in recent decades coincident with increases in car ownership (Jones et al., 2012). There is evidence that characteristics of the physical environment, such as traffic density (Winters et al., 2010), street lighting (Panter et al., 2010), sidewalk availability (Rodríguez and Joo 2004) and other aspects of neighbourhood design (Bauman et al., 2012) might influence the propensity of individuals to be active travellers. Empirical research has therefore attempted to determine associations between travel mode and the objective environmental characteristics of commuting routes to inform potential interventions. However, the evidence is limited by the fact that studies have often assumed that the characteristics of the shortest or fastest routes between home and work reflect the true routes of study participants (Panter et al., 2010; Rodríguez and Joo, 2004). If these assumed routes do not correspond to those actually taken (Oliver et al., 2010; Winters et al., 2010), the environmental exposures attributed to participants in these studies may not reflect those actually received. For example, the commuter may

* Corresponding author at: Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7 TJ, UK. Tel.: +44 1603 591361.
E-mail address: a.dalton@uea.ac.uk (A.M. Dalton).

http://dx.doi.org/10.1016/j.jth.2014.10.001
2214-1405/

be exposed to a higher number of fast food outlets on their journey than a modelled route would suggest, which may be associated with higher consumption of energy-dense food and increased risk of obesity (Burgoine et al., 2014).

The choice of route and mode of travel may depend on the social context in which household-level decisions are made (Guell et al., 2012), as well as on individual preferences (Panter and Jones, 2010). Indeed, route choices may be conditioned by environmental characteristics, with longer routes potentially being selected to provide more favourable commuting environments such as quieter roads. Such processes have recently been referred to as being manifestations of 'selective daily mobility' (Chaix et al., 2013). Furthermore, trips to and from work are often multi-purpose (Ye et al., 2007), with stops at locations en-route such as shops and schools (Noland and Thomas, 2007). As a result of this 'trip chaining', the overall length and duration of the commute may be much longer (McGuckin et al., 2005) and the environmental exposures very different from those that are modelled. Route choice modelling, used in transport planning disciplines to model a set of potential routes that a person may chose based on individual preferences and route characteristics (Prato, 2009), can be used to explore potential environmental exposures during the commute. However, these methods are often not adopted due to the need for specialist knowledge and software (Aldred, 2014) meaning that the evidence base available to inform interventions is potentially limited by the uncertainties inherent in more basic route modelling, usually performed using geographic information system (GIS) software, to identify shortest or fastest paths between origins and destinations (Badland et al., 2008; Cronberg and Bonsall, 2006; Panter et al., 2010; Rodríguez and Joo, 2004).

The availability of global positioning system (GPS) receivers – small, low cost, wearable devices which receive information from global navigation satellites and give a precise point location on the ground – means that it is now possible to track individuals as they use outdoor environments, and the devices are increasingly being employed in physical activity research (Krenn et al., 2011; Maddison and Mhurchu, 2009). For example, recent work has used GPS-derived information to validate self-report methods for assessing the level of physical activity involved in dog walking (Murray et al., 2012), to help correctly identify walking bouts (Kang et al., 2013), and to quantify the contribution that public parks make to physical activity (Evenson et al., 2013). GPS technology offers particular potential to refine our understanding of how the environment shapes commuting behaviours because it allows researchers to identify the actual routes that study participants take to and from work, as well as to validate the use of travel modes reported by participants (Maddison and Mhurchu, 2009). On the other hand, the use of GPS technology increases the burden in studies—both for participants, due to the need to carry receivers and recharge them frequently (Krenn et al., 2011), and for researchers, due to the need to clean and analyse the large volumes of data they produce (Duncan et al., 2013; Stopher et al., 2005). As the computation of estimated routes in GIS software is a rapid procedure, requiring simply the location of the origin and destination of each trip, an important question is whether the refined understanding obtained from using GPS in studies of travel behaviour justifies their considerable resource implications.

Limited empirical research has compared modelled with actual commuting routes measured using GPS technology. Findings have shown that children tend to use quieter streets than GIS models suggest (Duncan and Mummery, 2007) and, unsurprisingly the actual routes of both children and adults are longer than the shortest-route distances that GIS estimate (Duncan and Mummery, 2007; Harrison, 2014; Ramaekers et al., 2013). A small study of 29 journeys made by adults suggested that the mix of land uses suggested by GIS routes had poor concordance with those actually experienced (Badland et al., 2010), and there is also evidence that deviations between actual and modelled routes are associated with individual characteristics including sex and employment status (Ramaekers et al., 2013). However, a limitation of much of the evidence base is that studies have not fully explored the role of travel mode and route characteristics in determining the concordance between actual and modelled routes, limiting our understanding of the utility of modelling the route in studies of the determinants of travel behaviour.

This study aims to assess the implications of employing modelled routes (using GIS) versus those actually taken (captured using GPS), using data from a diverse sample of commuters working in Cambridge, UK. We evaluate whether GIS can produce routes that adequately reflect the characteristics of the actual routes taken using a set of commonly assessed metrics, and ascertain the circumstances under which GIS-derived shortest distance path routes may reflect those actually taken either particularly well or poorly. We identify circumstances in which modelled routes could be useful proxies for actual journey characteristics, and suggest implications of our findings for the assessment of environmental exposures.

2. Material and methods

2.1. Study design, setting and sample

We used cross-sectional data from a sample of commuters participating in the Commuting and Health in Cambridge study in Cambridge, UK. Participants were aged 16 and over, working in Cambridge, and living within 30 km of the city but not in the immediate vicinity of their workplace. They were sampled using a workplace recruitment strategy that targeted a variety of workplaces and employers in a range of geographical locations across Cambridge city centre and urban fringe. Full details of the study protocol are outlined elsewhere (Ogilvie et al., 2010). Data collected during the second and third phases of the study (May to November 2010 and 2011) were used for this analysis. Participants completed a postal questionnaire that provided information on individual and household characteristics including age, gender, type of work, health, physical health problems limiting physical activity, level of education, number of children, home ownership and number of cars, as well as postcodes (zip codes) for their home and workplace locations. The questionnaire included a seven-day retrospective record of travel to and from work, recording the modes of travel used and the time of starting and finishing work each day (Panter et al., 2011). A subsample of participants also completed a seven-day household travel diary, recording all journeys made during that period along with the purpose, start time and location and end time and location for each journey, and the duration and mode of travel of each journey stage. A further subsample of participants wore a QStarz BT-Q1000X GPS data logger, set to record location every 5 s using the American NAVSTAR-GPS network, for seven days. A total of 776 individuals completed the questionnaire, 488 also completed the household travel diary and 194 participants also successfully completed GPS data collection and were eligible for inclusion in this analysis. Ethical approval was obtained from the Hertfordshire Research Ethics Committee (reference numbers 09/H0311/116 and 10/H0311/65) and written informed consent was provided by each participant.

Information from the questionnaire and travel diary was used to identify all journeys to and from work and the modes of travel used. The GPS-tracked data were visually inspected in the GIS software to identify any discrepancies in travel mode with the self-report data. As a result, one trip coded in the diary as 'car plus walk' was coded to the questionnaire response of 'bus' and one trip coded as 'bus plus walk' in the diary was coded to the questionnaire response of 'walk', whilst seven trips (from two participants) were omitted as agreement could not be found.

The analysis reported here required a sample of commuting trips that were broadly representative of the six most commonly reported types of trip: walk only, car plus walk, bicycle only, car plus bicycle, bus, and car or motorcycle only. It therefore used data from 51 participants from the second phase of the study (2010) selected by random quota sampling to achieve at least 50 journeys of each type. Participants reporting bus in combination with walking were combined with those reporting bus only due to small numbers and the fact that any bus journey would likely involve some walking to access the bus stop. An insufficient number of participants had recorded walk-only trips in the second phase of the study to achieve the target sample size, so data for three participants collected during the third phase (2011) were used to augment the sample. The modal number of journeys analysed per individual was five.

Only journeys that began at home or work and ended at work or home were considered. Participants' home and work postcodes were geo-referenced using the Ordnance Survey (OS) Code Point® database (Ordnance Survey, 2012), which contains a precise location on the ground for every postcode in the UK, and the ArcGIS 9.3 geographic information system (GIS) software package (ESRI, 2009). A measure of the urban or rural nature of the home location was computed based on the classification by Bibby and Shepherd (2004), which allocates every census output area of England and Wales a classification of either being urban; town and fringe; village; or hamlet and isolated dwellings, based on population size and the density of residential addresses. For the purposes of this analysis, we define 'urban' as urban or town and fringe, and 'rural' as village or hamlet and isolated dwellings.

2.2. Modelling routes for journeys to and from work

Based on a commonly adopted assumption that people would take the shortest distance route (Panter et al., 2010; Rodríguez and Joo, 2004), modelled routes representing the shortest route between participants' home and work locations were derived using ArcGIS. A road network was represented using the OS MasterMap® Integrated Transport Network™ (ITN) database (Ordnance Survey, 2013), and was used to create routes for participant journeys made by car or bus. A pedestrian and bicycle route network was used to model trips made by bicycle or on foot, constructed by supplementing road data, excluding motorways, with local authority data on rights-of-way (public footpaths, bridleways and byways) (Cambridge County Council, 2010), cycle route information from the charity Sustrans (2012), and other informal pathways recorded on the crowd-sourced OpenStreetMap.com (OpenStreetMap, 2010).

2.3. Deriving actual routes for journeys to and from work

For each participant, all GPS points recorded for 2 h either side of self-reported work start and finish times were extracted, with the aim of capturing the entire route to and from work and allowing for any discrepancies between reported and actual arrival and departure times. Additional points were further extracted where the journey was not entirely captured in these 2 h periods, for example when participants stopped en route. The scattering of data points resulting from signal error or reflection from buildings, for example where the participants had entered or travelled close to buildings, was cleaned by manually removing erroneous points, and all data points collected at home/work before or after the journey began/ended were removed. The data points were converted to linear routes using the Hawth's Tool add-on for ArcGIS (Beyer, 2004). This tool connected temporally consecutive data points into line features using a unique identification field of journey date plus direction (to or from work) to differentiate between journeys. For these routes, journey duration (based on number of GPS-tracked points recorded at 5 s intervals), mean speed (distance divided by duration), and time spent stationary (defined as less than 1 km/h) were calculated.

For each trip, the route was manually inspected to determine whether the participant travelled to or from work via an additional intermediate location, such as a school or a shop. The locations were identified by overlaying the routes with background mapping and aerial photography in the GIS software. A 'via' was determined to be present if a participant remained at such a location for more than 5 min without changing mode of travel. These criteria prevented the erroneous identification of 'via' locations attributable, for example, to waiting at traffic lights or public transport interchanges.

2.4. Comparing characteristics of routes

A range of metrics were calculated for both actual and modelled routes. Aspects of the environment that would potentially determine route choice were selected, guided by the literature on environmental correlates of travel behaviour (Fraser and Lock, 2010; Panter and Jones, 2010; Broach et al., 2012; Winters et al., 2010). These included route distance; route directness (route distance divided by straight line distance); the number of leisure, retail and food destinations within a 100 m buffer of the route (a distance commonly used to indicate an accessible distance from a route (Burgoine and Monsivais, 2013; Panter et al., 2010)) taken from OS Points of Interest data (PointX Ltd, 2010), and the percentage of the route which fell on busy roads ('A' or 'B' roads or motorways). In addition, health-related exposures were included to indicate the impact of using modelled routes for assessing such exposures. These comprised two subsets of the OS Points of Interest data: 'healthy intermediate destinations' (sports places, athletics facilities, bowling facilities, golf courses, gyms and sports centres, swimming pools, tennis facilities) and 'unhealthy intermediate destinations' (fast food establishments), within 100 m of the route. There has been some suggestion that the availability of such 'healthy' facilities may facilitate physical activity (Humpel et al., 2002; Karusisi et al., 2013; Pascual et al., 2013); and the latter were selected as previous research has suggested that greater exposure to fast food outlets may be linked to higher consumption of fast food and elevated obesity risk (Burgoine et al., 2014).

Route lengths obtained from the GPS and GIS data were calculated in ArcGIS. Two outcome measures were selected to assess concordance between modelled and actual routes: the difference in length (as a percentage of the modelled route length) and the percentage spatial overlap between the routes. The differences in length were computed as per previous research for assessing concordance (Duncan and Mummery, 2007), and values were expressed as a percentage of the modelled distance, with a positive value indicating that the actual (GPS-tracked) route was longer than the modelled (GIS shortest path) route. Overlap was chosen as a way of assessing the spatial comparability of routes. The overlap between actual and modelled routes was estimated by assessing the percentage of the GPS-tracked route that fell within a 50 m buffer of the GIS-shortest path route. This buffer distance was chosen to allow for errors in location associated with variations in the quality of GPS signal and the fact that road users do not typically travel along the road centreline which was used for the computation of modelled routes. This value was found to be appropriate in a previous study of children's routes to school (Harrison, 2014).

Fig. 1 provides an example of modelled and actual routes for the journey to work, along with one of the measured metrics: the destinations identified en route. The actual route taken – by bicycle in this case – is clearly longer than the modelled route and overlaps only in parts, with the participant passing fewer destinations en route, taking a less direct route and travelling along quieter roads.

2.5. Statistical analysis

Absolute differences between actual and modelled routes for each route metric were calculated, along with 95% confidence intervals using the ci command in Stata Corp (2013), by mode of travel. The level of agreement between metrics was assessed using Lin's concordance coefficient (Lin, 1989), a method of evaluating the reproducibility or interchangeability of alternative measurements on a continuous scale by evaluating the proximity of pairs of observations to the line of equivalence, and 95% limits of agreement ($\pm$1.96 standard deviations of the mean difference) following Bland–Altman techniques (Bland and Altman, 1999). The p-value was reported and the coefficient was tested using McBrides's criteria for strength of agreement (McBride, 2005).

The distribution of each outcome variable was tested for normality using the Shapiro–Wilk test, which indicated neither variable was normally distributed. Differences in the percentage overlap by mode between modelled and actual routes were therefore assessed for statistical significance using the non-parametric Spearman's Rho rank correlation coefficient. Forest plots were used to display mean overlap by mode, along with 95% confidence intervals.

Multilevel mixed-effects generalised linear regression models were fitted (using the meglm command in Stata 13 (Stata Corp, 2013)) to examine which characteristics of the modelled route environment and participant independently predicted the degree of concordance in route length, percentage overlap with actual routes and difference in number of unhealthy destinations. A multilevel structure was used to account for the clustering of trips within individuals. Categorical variables with multiple levels (such as travel mode) were included in the regression models only if their overall contribution to the model was significant at $p<0.05$, tested using the log-likelihood test.

Prior to model fitting, multicollinearity was assessed by producing a pair-wise correlation matrix to identify variables that were highly associated, defined as having a Pearson's or Spearman's correlation coefficient of >0.55 based on previous empirical research (Grewal et al., 2004). Only one variable from each correlated pair was selected for modelling, the chosen one being that with the strongest association in the expected direction with the outcome. As a result of this, the measure of route directness was dropped from the model of route environment characteristics as it was significantly correlated with the percentage of busy roads en route at -0.587; route distance, travelling 'via' an intermediate location and the number of intermediate destinations en route were taken forward in model-building. For the model of participant characteristics, no measures were correlated at >0.55. Therefore age, work type, gender, homeownership, health, and whether participants had physical conditions limiting their activity, a degree, a car, or children, were all taken forward. To aid interpretation of the model intercept, all continuous variables were centred on their means. All analysis was conducted using Stata 13 (Stata Corp, 2013).

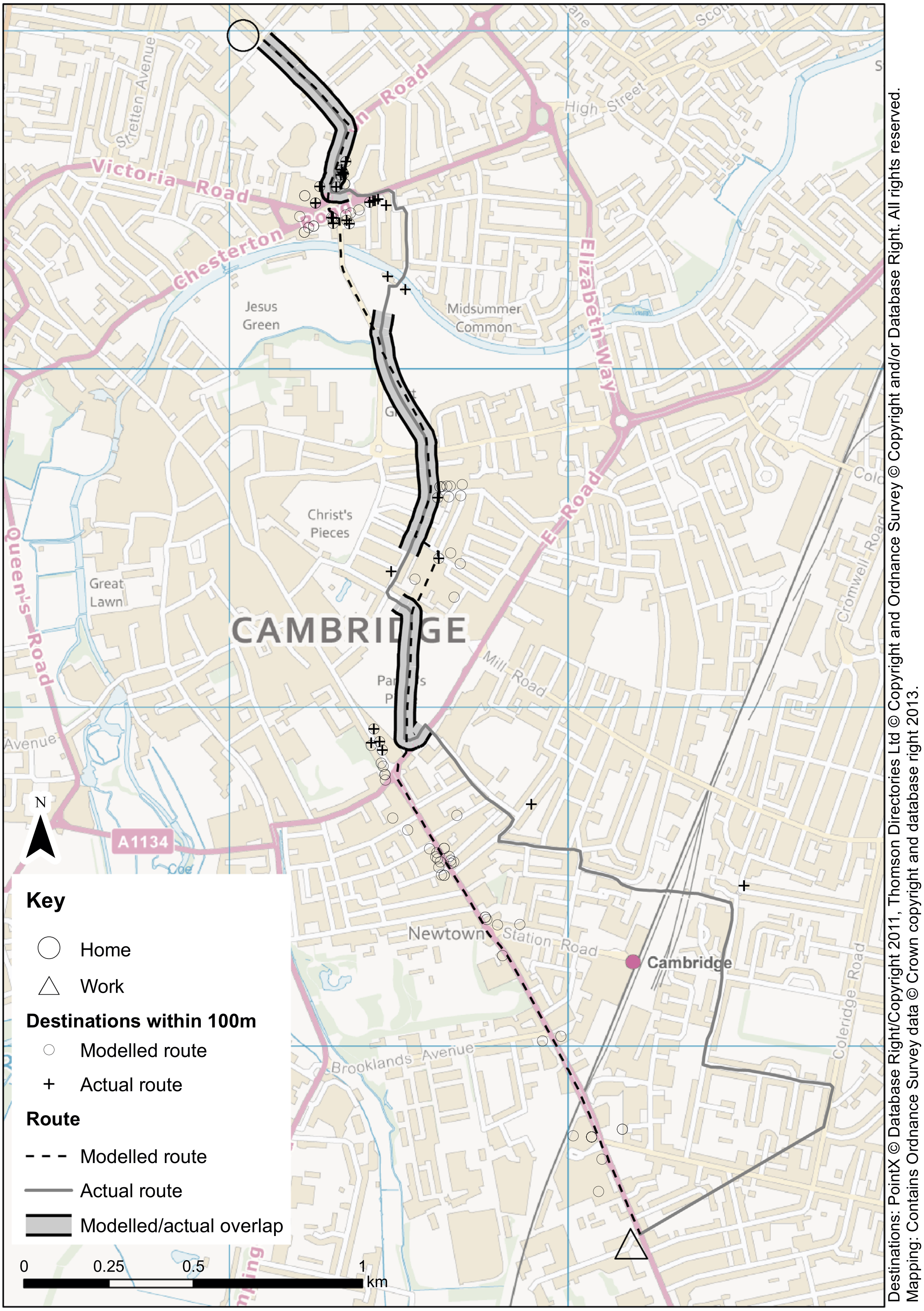

Fig. 1. Example of an actual and a modelled journey between home and work for one participant. Precise home and work locations have been moved in order to preserve participant anonymity.

3. Results

3.1. Sample characteristics

The sample comprises 51 study participants. The characteristics of the sample at the individual and household level are presented in Table 1. The average age was 45 years and the sample was predominantly female, with most people working in a sedentary occupation

Table 1
Sample characteristics: individual and household ($n=51$).

	All (n)	Male (n)	Female (n)
Individual characteristics			
Mean age (years)	45.0 (51)	48.9 (21)	42.3 (30)
Gender (%)	100.0 (51)	41.2 (21)	58.8 (30)
Work type (%)			
Sedentary occupation	76.5 (39)	90.5 (19)	66.7 (20)
Standing occupation	23.5 (12)	9.5 (2)	33.3 (10)
Health (%)			
Excellent	33.3 (17)	28.6 (6)	36.7 (11)
Very good to good	58.8 (30)	57.1 (12)	60.0 (18)
Fair to poor	7.8 (4)	14.3 (3)	3.3 (1)
Physical conditions limit activity (%)			
Not at all	74.5 (38)	66.7 (14)	80.0 (24)
Very little	19.6 (10)	19.0 (4)	20.0 (6)
Somewhat	5.9 (3)	14.3 (3)	0.0 (0)
Degree (%)			
Lower than degree	15.7 (8)	9.5 (2)	20.0 (6)
Degree	84.3 (43)	90.5 (19)	80.0 (24)
Mode of travel to work (% journeys)			
Bicycle	17.4 (48)	16.5 (17)	17.9 (31)
Bus	18.8 (52)	24.3 (25)	15.6 (27)
Car/motorcycle	19.9 (55)	20.4 (21)	19.7 (34)
Car+bicycle	16.7 (46)	5.8 (6)	23.1 (40)
Car+walk	18.5 (51)	12.6 (13)	22.0 (38)
Walk	8.7 (24)	20.4 (21)	1.7 (3)
Household characteristics			
Children (%)			
None	68.6 (35)	71.4 (15)	66.7 (20)
One or more	31.4 (16)	28.6 (6)	33.3 (10)
Home ownership (%)			
Does not own home	33.3 (17)	33.3 (7)	33.3 (10)
Home owner	67.3 (34)	67.3 (14)	67.3 (20)
Number of cars (%)			
None	11.8 (6)	23.8 (5)	3.3 (1)
One or more	88.2 (45)	76.2 (16)	96.7 (29)
Urban rural (%)			
Rural	19.6 (10)	14.3 (3)	23.3 (7)
Urban	80.4 (41)	85.7 (18)	76.7 (23)

and reporting being in good or very good health. The sample was generally highly educated and most lived in households with no children, were home owners, and had one or more cars. Most of the sample was drawn from urban environments. A total of 276 separate trips between home and work were manually extracted from GPS-tracked data. Among the trips studied, the generally even split between modes reflected the quota sampling strategy employed. The lower prevalence of walk-only and 'car plus bicycle' trips reflects the fact that it was only possible to obtain 24 and 46 of such trips, respectively.

Based on the actual routes derived from the GPS, the mean trip distance was approximately 20 km and the mean duration approximately 36 min. A comparison of trip characteristics between the six modes for the actual routes taken showed that, unsurprisingly, walk-only trips were the shortest (1.3 km, 14 min on average) and those made by car plus walking the longest (30.6 km and 47 min) (Table 2). Walking trips were most direct and car-only trips the least direct. Just over 27% of trips were via an intermediate destination. However, there were statistically significant differences between modes ($p=0.017$), with car/motorcycle trips being most likely to include a stop en route (on 40% of trips) and walk-only trips the least (on 13% of trips). Average speeds ranged from 5.6 km/h for walking trips to 46.9 km/h for car/motorcycle only trips, whilst trips made by car plus bicycle had a slightly lower average speed (38 km/h) than those made by car plus walking (41 km/h), reflecting the fact that those combining the car with walking spent a greater proportion of the journey driving than those who combined the car with cycling. Walking trips involved the least amount of time spent stationary, and bus trips the most.

3.2. Comparison of actual and modelled shortest path trip characteristics

On average, the routes followed between home and work were 4.3 km (27%) longer (95% CI 3.6 to 4.9 km) than the shortest path routes modelled in the GIS (Table 3). However, this result conceals differences between modes of travel: the routes followed on trips made by bicycle, bus and car (whether singly or in combination with other modes) were all longer on average than the modelled routes, whereas those followed on walking trips were actually 0.2 km (13%) shorter (95% CI -0.1 to -0.2 km), indicating that participants used cut throughs and other paths not present in the route networks used for analysis. Lin's concordance coefficient of 0.84 indicates poor agreement between modelled and actual route length overall, but disaggregating the analysis by mode shows substantial agreement for journeys made on foot (0.98) and moderate agreement for journeys made by bicycle (0.93). The lowest level of agreement was for trips made by car or motorcycle only (0.44). Unsurprisingly, route distance was generally longer for participants who travelled via another destination on their way to or from work (Fig. 2).

Environmental exposures along the route were significantly associated with the difference between actual and modelled routes. However, these differences also varied according to mode of travel used (Table 2). For example, route directness showed extremely poor

Table 2
Environmental characteristics of trips: GPS tracked versus GIS shortest path (averages, $n=276$).

Characteristic	Mode	Actual route (GPS-tracked)	Modelled route (GIS shortest path)	Absolute difference (mean)	CI (95%) for absolute difference		Limits of agreement (95%)		Difference (%)	Lin's concordance coeff.	p[a]
					Lower	Upper	Lower	Upper			
Distance (mean, km)	Bicycle	7.6	6.7	0.9	0.5	1.3	−1.9	3.6	13.4	0.93	< 0.001
	Bus	21.0	16.5	4.4	3.2	5.7	−4.6	13.4	26.7	0.86	< 0.001
	Car/MC[b]	25.3	18.7	6.5	4.1	9.0	−11.5	24.5	34.8	0.44	< 0.001
	Car+bicycle	23.7	18.7	5.0	3.9	6.1	−2.3	12.2	26.7	0.64	< 0.001
	Car+walk	30.6	24.4	6.2	5.3	7.1	0.1	12.3	25.4	0.78	< 0.001
	Walk	1.3	1.5	−0.2	−0.2	−0.1	−0.5	0.2	−13.3	0.98	< 0.001
Route directness (mean ratio, 1=most direct)	Bicycle	1.40	1.26	0.14	0.1	0.2	−0.3	0.6	11.1	0.03	0.696
	Bus	1.47	1.17	0.29	0.2	0.4	−0.2	0.8	25.6	0.06	0.133
	Car/MC	1.79	1.20	0.60	0.3	0.9	−1.8	2.9	50.0	0.07	0.002
	Car+bicycle	1.41	1.12	0.29	0.2	0.3	−0.1	0.7	25.9	0.05	0.011
	Car+walk	1.50	1.17	0.32	0.3	0.4	−0.0	0.7	27.4	0.07	0.002
	Walk	1.17	1.56	−0.39	−0.6	−0.2	−1.2	0.4	−25.0	−0.18	0.054
Busy roads: A/B roads and motorways on route (mean, %)	Bicycle	8.4	15.2	−6.8	−11.2	−2.4	−36.4	22.8	−44.7	0.07	0.523
	Bus	63.6	68.1	−4.5	−8.6	−0.3	−33.9	25.0	−6.6	0.71	< 0.001
	Car/MC	67.9	69.0	−1.0	−6.1	4.0	−37.3	35.2	−1.4	0.71	< 0.001
	Car+bicycle	75.3	84.2	−8.9	−15.3	−2.6	−50.9	33.1	−10.6	−0.16	0.155
	Car+walk	75.7	70.2	5.5	−3.3	14.4	−56.3	67.4	7.8	−0.23	0.071
	Walk	3.3	6.8	−3.5	−8.4	1.4	−26.2	19.2	−51.5	0.31	0.044
Intermediate destinations on route[c] (mean, number)	Bicycle	96.5	105.2	−8.7	−27.0	9.6	−132.2	114.8	−8.3	0.80	< 0.001
	Bus	159.5	213.6	−54.1	−96.2	−12.0	−350.5	242.3	−25.3	0.51	< 0.001
	Car/MC	47.0	132.0	−85.0	−122.0	−48.0	−353.0	183.0	−64.4	0.08	0.264
	Car+bicycle	74.2	142.5	−68.3	−106.5	−30.1	−320.6	184.0	−47.9	0.37	< 0.001
	Car+walk	65.9	193.7	−127.7	−166.6	−88.9	−398.8	143.2	−65.9	0.02	0.635
	Walk	11.4	15.8	−4.4	−6.7	−2.2	−14.8	6.0	−27.8	0.95	< 0.001
'Healthy' intermediate destinations (sports) on route (mean, number)	Bicycle	2.1	1.5	0.6	−0.1	1.3	−4.0	5.2	40.0	0.18	0.145
	Bus	5.9	4.9	1.0	0.0	2.0	−6.3	8.3	20.4	0.42	< 0.001
	Car/MC	5.9	3.1	2.8	2.0	3.6	−2.9	8.5	90.3	0.26	< 0.001
	Car+bicycle	6.1	6.1	0.0	−1.8	1.8	−12.1	12.0	0.0	0.09	0.493
	Car+walk	7.3	6.9	0.4	−0.6	1.4	−6.7	7.5	5.8	0.41	< 0.001
	Walk	1.6	1.4	0.3	−0.4	0.9	−2.6	3.1	21.4	0.61	< 0.001
'Unhealthy' intermediate destinations (fast food) on route (mean, number)	Bicycle	2.8	3.1	−0.2	−1.4	0.9	−7.8	7.4	−6.5	0.26	0.043
	Bus	8.8	8.9	−0.1	−1.9	1.7	−12.9	12.7	−1.1	0.68	< 0.001
	Car/MC	2.3	6.0	−3.7	−5.5	−2.0	−16.3	8.9	−61.7	0.07	0.425
	Car+bicycle	3.4	6.7	−3.2	−5.4	−1.1	−17.3	10.8	−47.8	0.28	0.002
	Car+walk	2.9	7.9	−5.0	−6.7	−3.3	−17.1	7.1	−63.3	0.02	0.784
	Walk	0.8	0.8	0.0	−0.3	0.3	−1.4	1.3	0.0	0.96	< 0.001

[a] p=difference between GIS shortest path and actual routes according to each characteristic, for each mode, from Lin's concordance coefficient.
[b] MC=motorcycle.
[c] Retail, food, leisure and education establishments.

Table 3
Best fit multilevel linear regression model of difference in distance between actual and modelled routes to work (as a percentage of the modelled distance), and environmental characteristics of the predicted routes ($n=276$, indicative adjusted $R^2=0.234$).

	Robust Coefficient	p	CI (95%)	
			Lower	Upper
(Intercept)	38.813	< 0.001	25.877	51.748
Distance (GIS, km)	−0.957	0.017	−1.746	−0.169
Via intermediate location	26.497	< 0.001	15.791	37.202
Mode used (Car/MC[a]=ref):	−44.056	< 0.001	−63.741	−24.371
Bicycle				
Bus	−18.474	0.041	−36.216	−0.732
Car+bicycle	−18.481	0.049	−36.887	−0.074
Car+walk	−10.066	0.283	−28.455	8.324
Walk	−74.530	< 0.001	−100.489	−48.571

[a] MC=motorcycle.

levels of agreement for all types of trip involving cars, which consistently involved longer than modelled routes. All journeys, except those made by car plus walking, were along a route with fewer busy roads than the modelled route suggested, with walkers tending to choose the quietest routes. Trips made by all modes also passed fewer destinations than a modelled route would predict, although routes followed by walkers did show substantial agreement on this measure (0.95). The lowest agreement was seen for trips involving cars, which

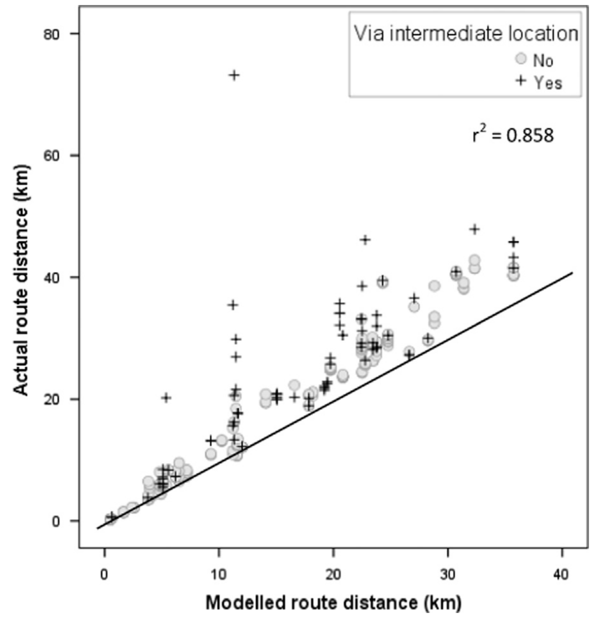

Fig. 2. Comparison of modelled and actual route distance of trips, separated into those that did or did not go via an intermediate location on route, and showing line of equivalence (modelled distance = actual distance), $n=276$.

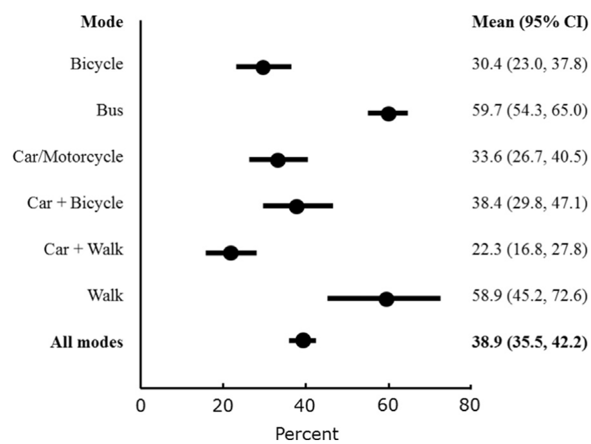

Fig. 3. Mean percentage of route overlap: GPS-tracked versus GIS shortest path ($n=276$).

involved exposure to a significantly lower number of 'unhealthy' food outlet intermediate destinations, and a significantly higher number of 'healthy' intermediate destinations, than the modelled routes suggested.

The mean amount of overlap between actual and modelled routes was relatively low at 39%. However, this also showed much variation according to mode of travel ($p<0.001$). On average, overlap was highest for bus (60%) and walking (59%) trips and lowest for car plus walk trips (22%) (Fig. 3). The correlation between the measures of difference in distance and percentage overlap was statistically significant ($p<0.001$), but fairly low at $r=-0.309$.

3.3. Predictors of concordance between shortest path and actual routes

In the multivariable regression model, the difference between GIS modelled shortest path and actual route length was independently associated with modelled route length, the inclusion of an intermediate 'via' location and the mode of travel used (Table 3). The difference declined by just under 1% (coefficient −0.957) for every 1 km increase in modelled route length, but was increased by 26% (coefficient 26.497) by the inclusion of an intermediate 'via' location. Regression analysis suggested that trips most likely to include a 'via' were ones made by people with children, and during commutes to rather than from work (results not shown). The difference was also significantly lower for trips made by bicycle, car plus bicycle, bus or walking than for trips made by car or motorcycle only. No participant characteristics were significantly associated with the difference in route length. By way of illustration, the model results suggest that a route with a modelled shortest distance in the GIS of 15.78 km (the mean), did not go via a location en route, and was undertaken solely by car or motorcycle would have an actual length that was 38.8% greater than that modelled, as shown by the coefficient for the intercept value. The indicative R^2 value of 0.24 is relatively modest, suggesting the presence of considerable unexplained variance in the difference in route distance.

The best fit model for predicting the degree of overlap between actual and shortest path routes is shown in Table 4. The only statistically significant environmental characteristic that predicted overlap was the percentage of busy roads, with the coefficient (0.248)

Table 4
Best fit multilevel linear regression model of percentage overlap between actual and modelled routes to work, and environmental characteristics of the predicted routes (n=276, indicative adjusted R^2=0.218 (using linear regression).

	Robust coefficient	p	CI (95%)	
			Lower	Upper
(Intercept)	21.239	0.029	2.205	40.273
Busy roads (%)	0.248	0.047	0.003	0.493
Mode used (Car/MC=ref)	−8.222	0.268	−22.762	6.319
Bicycle				
Bus	20.280	0.005	6.257	34.302
Car+bicycle	−1.361	0.844	−14.893	12.170
Car+walk	−11.023	0.096	−24.010	1.963
Walk	36.182	0.001	14.707	57.657

[a]MC=motorcycle.

Table 5
Best fit multilevel linear regression model of percentage difference in the number of unhealthy intermediate destinations between actual and modelled routes to work, participant characteristics and environmental characteristics of the predicted routes (n=276, indicative adjusted R^2=0.129 (using linear regression).

	Robust coefficient	p	CI (95%)	
			Lower	Upper
(Intercept)	43.337	0.061	−2.007	88.681
Excellent health (ref=not excellent)	−57.390	0.042	−112.698	−2.082
Children (ref=no)	64.611	0.025	8.030	121.191
Intermediate destinations (10 s)	−2.325	0.007	−4.012	−0.638

indicating that the overlap between these routes increased marginally with every percentage increase in the number of busy roads en route. No participant characteristics were significantly associated with route overlap. The shortest path routes that showed the most overlap with actual routes were those for journeys made by bus (20%) or on foot (36%).

The best fit model for predicting the difference in the number of unhealthy intermediate destinations between actual and shortest path routes is shown in Table 5. A coefficient below zero indicates an over-prediction of destinations when using the modelled routes as opposed to actual routes. The results suggest that a modelled route overestimates the number of unhealthy intermediate destinations when compared to the actual route taken by an average of 57% for people reporting excellent health, yet it underestimates the number of destinations that commuters with children were exposed to by 65%. Unsurprisingly modelled routes with many destinations on them were associated with overestimates of exposure compared to actual routes; for every ten intermediate destinations along a modelled route, there was a 2% over-prediction of exposure to unhealthy intermediate destinations compared to actual routes. Mode of travel was not significantly associated with differences in exposure to unhealthy food outlets between modelled and actual routes.

4. Discussion

This research has found that using GIS software to model the shortest path between people's homes and workplaces may be a useful method to estimate the distance travelled for non-car journeys, with modelled walk and cycle trip distances showing moderate to substantial agreement with actual trips recorded using GPS technology. Much recent work has examined the predictors of active travel (Winters et al. 2010; Bauman et al. 2012; Panter et al. 2010; Rodríguez and Joo 2004). The findings from this study suggest that using GIS for this purpose is adequate, and this may be useful for researchers wishing to model physical activity levels or energy expenditure during the commute to work. However, modelling performed less well when predicting route location and associated environmental exposures, a limitation apparent across all modes. This may be problematic for assessing the risk of obesity or for understanding how to design environments to promote healthy behaviours, such as active travel which has been found to be associated with lower BMI (Flint et al. 2014). In terms of obesogenic environmental exposure, using modelled routes may underestimate exposure to unhealthy food environments for some people, such as those with children, whereas it may overestimate exposure for others, such as those in better health. This reflects the fact that different population groups will make varied choice sets when deciding which route to take, which may introduce bias into studies if health-related behaviours, such as seeking food or physical activity opportunities, are associated with disparities between modelled routes and those actually taken.

We illustrate the magnitude of observed errors in GIS modelling with some examples. Taking one study participant who travelled 20 km home from work by car on one trip, the shortest route modelled using GIS software underestimated their actual distance by 79% and overlapped the actual route by only 15%, with the proportion of the actual route that involved busy roads being half of that predicted. For a second participant who walked a distance of 2.2 km from work to home, their route was more similar in length to that predicted (just 16% shorter) but only 19.2% of the route overlapped the modelled route. Further, their actual route passed no destinations and involved no busy roads, whereas the model indicated that they would pass 14 leisure, food, or retail attractions and spend 41% of the journey on 'A' or 'B' roads or motorways. Researchers from public health backgrounds using geographical analysis to estimate the health implications of exposure and route choice during the commute must be aware of the potential impact of using these methods.

Our findings build on previous research describing the modelled versus actual commuting routes by Badland et al. (2010), by testing associations in a larger sample of participants recruited from multiple workplaces, and commuting to work using different travel modes.

This has allowed us to examine a wider variety of journeys, revealing strong associations between routes taken, propensity to stop at locations en route and mode of travel, providing an insight into the types of journey for which modelled routes may be more or less appropriate. It is noteworthy that no participant characteristics, either at the individual or household level, were associated with either of the outcome variables, suggesting that, at least in our data, there is no particular type of commuter who will choose the most direct route. It is clear that individuals choose their mode of travel and route to work for reasons other than those of minimising time and distance. Nevertheless, although using GIS to predict routes and exposures is problematic, it important to recognise that modelled routes can provide indication of the routes that people might follow, and therefore of the environmental characteristics that might influence their propensity to choose particular modes. In such situations, they are not predicated on the need to observe real behavioural patterns.

Given that the GIS modelling largely replicates the shortest path methodologies used by in-car satellite navigation systems, we had expected journeys taken by car to most accurately reflect the GIS modelled routes. However, this was generally not the case, particularly as the propensity to deviate via destinations was greatest for car and motorcycle journeys. It is unsurprising that actual routes taken tended to be less direct than those estimated from the shortest path, as individuals deviate from routes to visit destinations. It may also be that people choose to deviate from theoretically optimal routes for reasons of familiarity, something we were unable to assess. Indeed, qualitative research into commuting behaviour in Cambridge suggests that choice of route is influenced by how pleasant, safe and enjoyable the journey is, as well as by family circumstances (Guell et al., 2012) and feelings of well-being (Guell and Ogilvie, 2013).

We did not attempt to model intermediate via locations when we used GIS software to calculate routes as our intention was to compare an actual route with the shortest path distance between home and work in order to examine the simplified assumptions that are commonly made when predicting journey routes. Indeed, most studies of this nature do not have access to information on use of intermediate destinations and therefore it would not be possible to model them.

4.1. Strengths and limitations

In terms of study strengths, we extracted 276 complete, validated commute journeys from a quota sample of 51 participants and included a range of different travel modes. Care was taken to manually extract, geolocate and clean the data to ensure that only complete, validated routes with an identified mode of travel were included in the analysis. However, it was not always possible to match modes, dates and times provided in the household travel diary and the questionnaire with the data recorded with the GPS device. As a result, seven trips had to be omitted from the analysis due to insoluble conflicts. Due to the time-intensive nature of data processing, GPS data were extracted for only 51 people out of a possible 194 in the entire sample, although these participants did provide a total of 276 individual journeys made through the May to November time period. Walkers were underrepresented in the sample due to the limited number of participants using this exclusive mode of travel for work. This reflected local travel patterns, and the recruitment strategy of the overall study which excluded individuals who lived very close to work (Ogilvie et al., 2010).

In terms of limitations, georeferencing using postcode information is less accurate in rural areas, as each postcode covers a larger geographic area because houses are fewer and more dispersed than in urban areas. However, postcode information is widely used to geo-locate rural residents and was the only option available for 20% of our sample who lived in a rural part of the study area.

Modelling multimodal trips was problematic, as participants did not specify where or for how long they walked or cycled when taking multimodal journey to or from work. This meant the active travel element of the journey could not be modelled. We assumed the motorised mode comprised the largest proportion of the journey, and therefore it was that road network that was used for modelling.

In this study, we have taken GPS to represent the best-available method to objectively record actual routes travelled by commuters. In reality, GPS data is subject to positional error and this has been shown to vary depending on the device, its use, and according to setting, such as the amount of sky or number of satellites visible to the device (Kerr et al., 2011). A recent study by Schipperijn et al. (2014) tested the accuracy of the QStarz BT-Q1000X for recording routes taken by different transport modes in a range of settings, finding an overall median error of just 2.9 m, ranging from 0.7 m in open areas to 5.2 m in heavily built-up urban areas, and from 3.9 m for walking trips to 0.5 m for car journeys. In this analysis, we undertook a comprehensive data cleaning exercise and used buffers to negate such positional inaccuracies, but it is likely that actual routes delineations are subject to a small error. It is also important to remember that while GPS data could help to inform or validate route choice models, its use may be inappropriate in some situations such as scenario planning, as observed behavioural patterns may be determined by factors other than those under study (Chaix et al. 2013).

Our sample worked in the city of Cambridge, UK, and lived within a 30 km radius of the centre. Therefore the findings may not be representative of other areas or of the general population living in the study area. However, commuting is a widespread behaviour with the same common aim and it may be possible to overcome any generalisability limitations by taking into account other factors that may influence commuting behaviour in a particular area. For example, topology may be important if the location is particularly hilly.

A buffer distance of 50 m was chosen for this analysis, as used in a previous study (Harrison 2014), however this raised the question of the appropriateness of this distance. The sensitivity of the model was tested by using a smaller buffer size of 15 m, and the same variables remained significant but the model fit was not as good (indicative R^2 of 0.129 rather than the 0.218 for the 50 m). A larger value than 50 m would not have been appropriate, as this would have resulted in a substantial overlap of roads. We used the best available digital data for the road and pedestrian route networks, by supplementing OS data with mapping held by local authorities, and additional crowd-sourced information. However, whilst these data contain some information on building cut-throughs, short-cuts and other informal paths, this is evidently still incomplete.

In this analysis we included journeys taken by bus. In reality commuters do no influence the route that the bus takes and bus routes are often not distance optimised. However as the purpose of this analysis was to examine the implications of the simplifying assumptions commonly used when modelling routes using GIS, bus trips were retained to specifically investigate the utility of estimating shortest routes for bus commuters.

5. Conclusions

The use of GIS to model routes may be acceptable when an approximate estimate of travel distance is required or when estimates of the features of potential routes that could be taken are needed. This is particularly relevant for active commuting, where actual routes travelled may show high levels of agreement with modelled shortest routes. However, the predictability of the commute route clearly depends on the mode of travel used. Therefore, public health researchers should not rely on uncritical assumptions regarding the validity of GIS-modelled routes as a proxy for the actual routes followed by commuters, particularly for those travelling by car. If we are to quantify exposure to environmental features along routes followed to and from work, or accurately estimate distance travelled by car, our results emphasise the need to obtain information on actual commuting behaviour. To obtain accurate estimates of environmental contexts in which behaviour actually occurs, the use of GPS data is recommended, as is the consideration of appropriate characteristics of the local area.

Acknowledgements

The Commuting and Health in Cambridge study was developed by David Ogilvie, Simon Griffin, Andy Jones and Roger Mackett and initially funded under the auspices of the Centre for Diet and Activity Research (CEDAR), a UKCRC Public Health Research Centre of Excellence (087636/Z/08/Z and ES/G007462/1). Funding from the British Heart Foundation, Economic and Social Research Council, Medical Research Council, National Institute for Health Research and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. The study is now funded by the National Institute for Health Research Public Health Research programme (project number 09/3001/06: see http://www.phr.nihr.ac.uk/funded_projects). David Ogilvie is supported by the Medical Research Council (Unit Programme number MC_UU_12015/6). Jenna Panter is supported by an NIHR post-doctoral fellowship (NIHR-PDF-2012-05-157). The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the NIHR PHR programme or the Department of Health. The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript. We thank all staff from the MRC Epidemiology Unit Functional Group Team, in particular for study coordination and data collection (led by Cheryl Chapman), physical activity data processing and data management.

References

Aldred R., 2014. Modelling on the move: Seminar 2—Public Health and Transport Modelling [Online]. London: University of Westminster. Available: ⟨http://modellingonthemove.org/modelling-on-the-move-write-up/write-up-seminar-2/⟩ [accessed 5 June 2014].
Badland, H.M., Duncan, M.J., Oliver, M., Duncan, J.S., Mavoa, S., 2010. Examining commute routes: applications of GIS and GPS technology. Environ. Health Prev. Med. 15, 327–330.
Badland, H.M., Schofield, G.M., Garrett, N., 2008. Travel behavior and objectively measured urban design variables: associations for adults traveling to work. Health Place 14, 85–95.
Bauman, A.E., Reis, R.S., Sallis, J.F., Wells, J.C., Loos, R.J.F., Martin, B.W., 2012. Correlates of physical activity: why are some people physically active and others not? Lancet 380, 258–271.
Beyer H., 2004. Hawth's Analysis Tools for ArcGIS [Online]. Available: ⟨http://www.spatialecology.com/htools⟩ [accessed 21 June 2011].
Bibby, P., Shepherd, J., 2004. Developing a New Classification of Urban and Rural Areas for Policy Purposes—The Methodology. ONS, London.
Bland, J.M., Altman, D.G., 1999. Measuring agreement in method comparison studies. Stat Methods Med Res 8, 135–160.
Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transport. Res. A—Policy Pract 46, 1730–1740.
Burgoine, T., Forouhi, N.G., Griffin, S.J., Wareham, N.J., Monsivais, P., 2014. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study ([Online], 348 (g1464), 10.1136/bmj.g1464. Available: ⟨http://www.bmj.com/content/348/bmj.g1464⟩). BMJ ([accessed 13 March 2014]).
Burgoine, T., Monsivais, P., 2013. Characterising food environment exposure at home, at work, and along commuting journeys using data on adults in the UK ([Online], 10 (85), 10.1186/1479-5868-10-85. Available: ⟨http://www.ijbnpa.org/content/10/1/85⟩). Int. J. Behav. Nutr. Phys. Act. ([accessed 8 August 2014]).
Cambridge County Council, 2010. Rights of Way, Permissive Access and Registered Common Land (Open Access on Foot) [Obtained via Personal Communication 30 November 2010]. Cambridge County Council, Cambridge.
Chaix, B., Méline, J., Duncan, S., Merrien, C., Karusisi, N., Perchoux, C., Lewin, A., Labadi, K., Kestens, Y., 2013. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health Place 21, 46–51.
Cronberg, A., Bonsall, P., 2006. Analysis of the Accuracy of NTS Respondents' Estimates of Trip Distance and Duration. Technical Annex 1 of Research Project UG599 (Review of the Potential Role of 'New Technologies' in the National Travel Survey). Department forTransport, London.
Department of Health, 2009. Be Active, Be Healthy: A Plan for Getting the Nation Moving. Department of Health, London.
Duncan, M.J., Mummery, W.K., 2007. GIS or GPS? A comparison of two methods for assessing route taken during active transport. Am. J. Prev. Med. 33, 51–53.
Duncan, S., Stewart, T.I., Oliver, M., Mavoa, S., MacRae, D., Badland, H.M., Duncan, M.J., 2013. Portable global positioning system receivers: static validity and environmental conditions. Am. J. Prev. Med. 44, e19–e29.
ESRI, 2009. ArcGIS 9.3. ESRI: California.
Evenson, K.R., Wen, F., Hillier, A., Cohen, D.A., 2013. Assessing the contribution of parks to physical activity using GPS and accelerometry. Med. Sci. Sports Exercise 45, 1981–1987.
Flint, E., Cummins, S., Sacker, A., 2014. Associations between active commuting, body fat, and body mass index: population based, cross sectional study in the United Kingdom ([Online], 349 (g4887), 10.1136/bmj.g4887. Available ⟨http://www.bmj.com/content/349/bmj.g4887⟩`). BMJ ([accessed 27 August 2014]).
Fraser, S.D.S., Lock, K., 2010. Cycling for transport and public health: a systematic review of the effect of the environment on cycling. Eur. J. Public Health 21, 738–743.
Grewal, R., Cote, J.A., Baumgartner, H., 2004. Multicollinearity and measurement error in structural equation models: implications for theory testing. Mark. Sci. 23, 519–529.
Guell, C., Ogilvie, D., 2013. Picturing commuting: photovoice and seeking wellbeing in everyday travel. Qual. Res. 0, 1–18.
Guell, C., Panter, J., Jones, N.R., Ogilvie, D., 2012. Towards a differentiated understanding of active travel behaviour: using social theory to explore everyday commuting. Soc. Sci. Med. 75, 233–239.
Harrison, F., 2014. How well do modelled routes to school record the environments children are exposed to?: A cross-sectional comparison of GIS-modelled and GPS-measured routes to school ([Online], 13 (5), 10.1186/1476-072X-13-5. Available: ⟨http://www.ij-healthgeographics.com/content/13/1/5⟩). Int. J. Health Geogr.hics ([accessed 8 August 2014]).
Humpel, N., Owen, N., Leslie, E., 2002. Environmental factors associated with adults' participation in physical activity: a review. Am. J. Prev. Med. 22, 188–199.
Jones, T., Pooley, C.G., Scheldeman, G., Horton, D., Tight, M., Mullen, C., Jopson, A., Whiteing, A., 2012. Moving around the city: discourses on walking and cycling in English urban areas. Environ. Plann. A 44, 1407–1424.
Kang, B., Moudon, A.V., Hurvitz, P.M., Reichley, L., Saelen, B.E., 2013. Walking objectively measured: classifying accelerometer data with GPS and travel diaries. Med. Sci. Sports Exercise 45, 1419–1428.
Karusisi, N., Thomas, F., Méline, J., Chaix, B., 2013. Spatial accessibility to specific sport facilities and corresponding sport practice: the RECORD study. Int. J. Behav. Nutr. Phys. Act. 10.
Kerr, J., Duncan, S., Schipperjin, J., 2011. Using global positioning systems in health research: a practical approach to data collection and processing. Am. J. Prev. Med. 41, 532–540.
Krenn, P.J., Titze, S., Oja, P., Jones, A., Ogilvie, D., 2011. Use of global positioning systems to study physical activity and the environment: a systematic review. Am. J. Prev. Med. 41, 508–515.
Lin, L.I.-K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268.
Maddison, R., Mhurchu, C.N., 2009. Global positioning system: a new opportunity in physical activity measurement. Int. J. Behav. Nutr. Phys. Act. 6, 73–80.

McBride G., 2005. A Proposal for Strength-of-agreement Criteria for Lin's Concordance Correlation Coefficient. NIWA Client Report: HAM2005-062. Hamilton, New Zealand: New Zealand Ministry of Health.
McGuckin, N., Zmud, J., Nakamoto, Y., 2005. Trip-chaining trends in the United States: understanding travel behavior for policy making. Transport. Res. Rec. 1917, 199–204.
Murray, H., Wharf Higgins, J., Temple, V., Tuokko, H., Porter, M., Rhodes, R., 2012. Measurement of dog walking behaviour: strengths and weaknesses of various methods. Health Fitness J. Can. 5, 16–26.
Noland, R.B., Thomas, J.V., 2007. Multivariate analysis of trip-chaining behavior. Environ. Plann. B 34, 953–970.
Ogilvie, D., Griffin, S., Jones, A., Mackett, R., Guell, C., Panter, J., Jones, N., Cohn, S., Yang, L., Chapman, C., 2010. Commuting and health in Cambridge: a study of a 'natural experiment' in the provision of new transport infrastructure. BMC Public Health 10.
Oliver, M., Badland, H., Mavoa, S., Duncan, M., Duncan, J., 2010. Combining GPS, GIS, and accelerometry: methodological issues in the assessment of location and intensity of travel behaviors. J. Phys. Act. Health 7, 102–108.
OpenStreetMap, 2010. OpenStreetMap: The Free Wiki World Map [Online]. Available: 〈http://www.openstreetmap.org〉 [accessed 20 December 2010].
Ordnance Survey, 2012. OS Code Point 2009–2011 [Online]. Southampton: OS. Available: 〈http://www.ordnancesurvey.co.uk/business-and-government/products/code-point.html〉 [accessed 30 May 2012].
Ordnance Survey, 2013. OS MasterMap Integrated Transport Network (ITN) Layer [Online]. Southampton: OS. Available: 〈http://www.ordnancesurvey.co.uk/oswebsite/products/os-mastermap/itn-layer/index.html〉 [accessed 25 November 2011].
Panter, J., Griffin, S., Jones, A., Mackett, R., Ogilvie, D., 2011. Correlates of time spent walking and cycling to and from work: baseline results from the commuting and health in Cambridge study. Int. J. Behav. Nutr. Phys. Act. 8.
Panter, J., Jones, A., Van Sluijs, E., Griffin, S., 2010. Neighborhood, route, and school environments and children's active commuting. Am. J. Prev. Med. 38, 268–278.
Panter, J.R., Jones, A.P., 2010. Attitudes and the environment as determinants of active travel in adults: what do and don't we know? J. Phys. Act. Health 7, 551–561.
Pascual, C., Regidor, E., Arco, D.A., Alejos, B., Santos, J.M., Calle, M.E., Martínez, D., 2013. Sports facilities in Madrid explain the relationship between neighbourhood economic context and physical inactivity in older people, but not in younger adults: a case study. J. Epidemiol. Community Health 67, 788–794.
PointX Ltd 2010. Ordnance Survey Points of Interest [Online]. Southampton: OS. Available: 〈http://www.ordnancesurvey.co.uk/oswebsite/products/points-of-interest/index.html〉 [accessed 14 March 2011].
Prato, C.G., 2009. Route choice modeling: past, present and future research directions. J. Choice Modell. 2, 65–100.
Ramaekers, K., Reumers, S., Wets, G., Cools, M., 2013. Modelling route choice decisions of car travellers using combined GPS and diary data. Netw. Spat. Econ. 13, 351–372.
Rodríguez, D.A., Joo, J., 2004. The relationship between non-motorized mode choice and the local physical environment. Transport. Res. D—Transp. Environ. 9, 151–173.
Schipperijn, J., Kerr, J., Duncan, S., Madsen, T., Klinker, C.D., Troelsen, J., 2014. Dynamic accuracy of GPS receivers for use in health research: a novel method to assess GPS accuracy in real-world settings ([Online], 2 10.3389/fpubh.2014.00021. Available 〈http://www.frontiersin.org/Journal/Abstract.aspx?s=503&name=public_health_education_and_promotion&ART_DOI=10.3389/fpubh.2014.00021〉). Front. Public Health ([accessed 2014-March-10]).
Stata Corp, 2013. Stata/SE 13.0 for Unix. Stata Corp: College Station, Texas.
Stopher P., Jiang Q., FitzGerald C., 2005. Processing GPS data from travel surveys. In: 28th Australasian Transport Research Forum, September 2005. Sydney, Australia.
Sustrans, 2012. About Sustrans [Online]. Bristol: Sustrans. Available: 〈http://www.sustrans.org.uk/about-sustrans〉 [accessed 3 October 2012].
Winters, M., Teschke, K., Grant, M., Setton, E., Brauer, M., 2010. How far out of the way will we travel? Transport. Res. Rec. 2190, 1–10.
Ye, X., Pendyala, R.M., Gottardi, G., 2007. An exploration of the relationship between mode choice and complexity of trip chaining patterns. Transport. Res. B—Methods 41, 96–113.