The neighbourhood physical environment:

Relationships with physical activity and depression in adults in the United Kingdom

Jemima Stockton

Thesis submitted for the degree of Doctor of Philosophy

University College London

DECLARATION

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Jemima Stockton

ABSTRACT

Background: The impact of neighbourhood physical environments on physical activity and health is widely acknowledged, with much research conducted to identify key factors. Results have been mixed, partly due to inconsistencies in how neighbourhoods are defined. This thesis examines relationships of physical environments with physical activity, and with depression, exploring influence of neighbourhood operationalisation.

Method: Physical activity and depression outcomes were derived from a sample drawn from the seventh wave of the Whitehall II study conducted in 2004/5, and depression outcomes were also taken from the 2008 Health Survey for England. Neighbourhoods were operationalised at three levels of administrative geography and as residential postcode-centred GIS software-computed zones. Four main exposure variables were specified: a greenspace measure was constructed from the Generalised Land Use Database; an objective measure of environmental quality was derived from metadata of the Multiple Environmental Deprivation Index, and a subjective one from the 2008 Place Survey; and a walkability measure was constructed using GIS, drawing on several geographical databases. Multivariate logistic regression was used to measure statistical associations between exposures and outcomes, with adjustment for individual-level sociodemographic factors and area-level deprivation, and multilevel modelling was performed to estimate the contribution of neighbourhood characteristics relative to those of individuals to variation in outcomes.

Results: Neighbourhood physical environments accounted for a small proportion of variation in all outcomes. Nevertheless, significant associations were found between all exposure variables and physical activity, independently of individual-level sociodemographic factors and area-level deprivation, the direction dependent on outcome specification. Only objectively measured environmental quality was significantly and independently with depression, with lower quality giving higher odds of this outcome. Strengths of associations were not substantively affected by neighbourhood operationalisation.

Conclusion: This thesis increases understanding of physical environment attributes relevant to physical activity and depression in a European context and how neighbourhoods in which they are measured may best be defined.

ACKNOWLEDGEMENTS

I am sincerely grateful to my inspiring supervisors – Dr Nicola Shelton, Professor Eric Brunner, Dr Manos Stamatakis and Dr Oliver Duke-Williams – for their guidance in the development and production of this thesis over the last four years. My primary supervisor, Nicola, has been encouraging and enthusiastic about my research throughout. I thank the participants of the Whitehall II Study and of the 2008 Health Survey for England for providing the health data on which this thesis has drawn, and the Economic and Social Research Council for funding my work. Also, many thanks to Professor Muki Haklay (UCL Dept of Civil, Environmental & Geomatic Engineering) for arranging access to UKMap data, to Dr Severine Sabia (UCL Dept of Epidemiology & Public Health) for help and advice in the derivation of the physical activity variables, to Dr Aida Sanchez (UCL Dept of Epidemiology & Public Health) for supplying Whitehall II data and to Anne Patrick (Ordnance Survey) for arranging access to the Urban Paths layer of OSMasterMap. Finally, I thank my loving family for putting up with me, my fellow PhD students – Silvia, Carolina, Tuba and Peggie – for being hilarious and making life in Room 404 so much fun, and my adorable friends Gayle and Sarah for their kindness and wise words.

For

Charlotte Lily Shaw,

Finlay Rocket Shaw

and

Lola Roxana Shaw

LIST OF APPENDICES

LIST OF EQUATIONS

LIST OF FIGURES

LIST OF GRAPHS

LIST OF MAPS

LIST OF TABLES

ABBREVIATIONS

Abbreviation	Definition
CAS ward	Census Area Statistics ward
CB	Circular buffer
CI	Confidence interval
GC200	200m by 200m grid cell
GC2000	2000m by 2000m grid cell
GHQ	General Health Questionnaire
GHQ-12 depression	A score of four or more on the 12-item General Health Questionnaire
GIS	Geographic Information System
GLUD	Generalised Land Use Database
GOR	Government Office Region
GS	Greenspace
GS+DG	Greenspace plus domestic garden
(L)HSE2008	(London only) Health Survey for England 2008
(L)WIIP7	(London only) Whitehall II Study Phase 7
ICC	Intra-class correlation
IMD	Index of Multiple Deprivation
ITN	Integrated Transport Network
J	Junction density only walkability model
JD	Junction density
L1	Land Use Mix 1 only walkability model
L2	Land Use Mix 2 only walkability model
L3	Land Use Mix 3 only walkability model
LA	Local authority
LSOA	Lower Super Output Area
LUM	Land Use Mix
M1	Full walkability model 1
M2	Full walkability model 2
M3	Full walkability model 3
MAUP	Modifiable Areal Unit Problem
MEDIx	Multiple Environmental Deprivation Index
MET	Metabolic Equivalent of Task
MPM	Minor psychiatric morbidity
MSOA	Middle Super Output Area
NLUD	National Land Use Database
NSSeC	National Statistics Socio-economic Classification
OA	Output area
OEQ	Objective environmental quality

Abbreviation	**Definition**
ONS	Office of National Statistics
OR	Odds ratio
PBNB	Polygon based network buffer
PEQ	Perceived environmental quality
R	Residential dwelling density only walkability model
RDD	Residential dwelling density
SD	Standard deviation
SE	Standard error
TTW	Being in the top tertile of WIIP7 for time spent walking per week
WHO	World Health Organisation
WHOPA_E	Meeting the World Health Organisation recommended physical activity level excluding any walking as a contributory physical activity
WHOPA_I	Meeting the World Health Organisation recommended physical activity level including walking for those who reported their normal pace was brisk or fast;
UP	Urban path

1 Introduction

"The potential benefits of physical activity to health are huge. If a medication existed which had a similar effect, it would be regarded as a 'wonder drug' or 'miracle cure'."

Liam Donaldson, Annual report of the Chief Medical Officer, 2009 [1]

Physical activity promotes both better physical and mental health [2] and reduces the risk of several leading causes of mortality and morbidity such as cardiovascular disease, breast and colon cancer, and depression [3]. Clearly bottling physical activity as a prophylactic or curative drug, as described by the Chief Medical Officer, is impossible, yet with most adults (63.3%) in England in 2008 not meeting recommended levels of physical activity and citing lack of time as a main reason [4], it would be convenient. In lieu of "bottled exercise" and in acknowledgement of perceived time pressures, the Chief Medical Officer's report [5] therefore suggests that incorporating physical activity into everyday life such as by walking – the most common form of physical activity [6] – is the easiest way to get it. There is evidence that individual factors such as health status and self-efficacy are determinants of physical activity [7], indicating they may also factor in an individual's propensity to incorporate walking into daily routines. However, another important parameter is likely to be external to the individual; the extent to which the physical environment of everyday life – the neighbourhood – is supportive of walking, making it a more attractive means of getting about than using motorised transport. Sports and recreation may represent a less common, albeit valuable, domain of physical activity for the England's "time-pressed" adults, but it too may be impeded or facilitated by features of the neighbourhood, such as the presence of parks. Thus, whilst putting physical activity in a bottle is unachievable, putting it – or at least the support to get it – on people's doorsteps in their neighbourhoods might be a realistic and important target for public health professionals and urban planners. Indeed, the last twenty years have seen burgeoning interest in the relationship of neighbourhood physical environments with physical activity - as a mediator in the relationship with health – and with health directly, with research conducted in efforts to provide the knowledge base necessary for such intervention. A particular focus has been the role of the natural environments such as greenspace in health, alongside – and partly driven by – concern for the environmental sustainability of urbanisation. However, progress in identifying and characterising neighbourhood physical environments, with respect to greenspace and other attributes, which promote physical activity and improve health has been sluggish and inconsistent evidence has been produced. A main reason is that many cross-sectional studies fail to model expected causal pathways, resulting in the construction of inadequate exposure variables in terms of how and where attributes are measured.

The goal of this thesis is to address this second issue through an examination of relationships of neighbourhood physical environments with physical activity, and with depression – a major cause of morbidity in England – and to explore the influence of definitions of neighbourhood and of physical environments on these associations.

More specifically, the major aims are to:

1. determine whether the proportion of neighbourhood greenspace is associated with physical activity, and with depression, independently of individual-level sociodemographic factors and of area-level deprivation
2. establish whether two "off-the-shelf" indicators of environmental quality – an objective and a subjective one – are independently associated with these outcomes
3. develop a bespoke walkability index for London designed to capture aspects of the neighbourhood salient to physical activity and examine the association of walkability with physical activity, and with depression
4. compare the effect on associations of neighbourhood operationalisation as administrative units at various geographical scales, and as geographical information system (GIS) software computed zones

To examine relationships of neighbourhood physical environments with physical activity, and with depression, and to explore the influence of definitions of neighbourhood and of environment on these associations, it is necessary to draw on large samples from two studies comprising high quality datasets, and on a variety of geographical databases. Physical activity and depression outcomes will be derived from a study sample drawn from the seventh wave of the Whitehall II study, a prospective cohort study that monitors the health and social circumstances of a group of civil servants. Additionally depression outcomes will be taken from the 2008 Health Survey for England, an annual government-commissioned survey to monitor trends in the nation's health. Identifiable to residential postcode level, the Whitehall II study sample will enable operationalisation of neighbourhoods at multiple administrative geographies and as residential postcode centred GIS software-computed zones, unique to each participant. Whilst each participant in the Health Survey for England sample is only identifiable to the administrative unit of local authority, allowing only operationalisation of neighbourhoods at this geographical level, the survey is nationally representative and thereby increases generalizability of findings. This thesis will employ multivariate logistic regression to measure associations between exposures and outcomes. Both studies that are used to provide outcomes as secondary data for this thesis also offer extensive information covering the sociodemographic characteristics of participants which will be exploited to adjust for potential confounders in associations with neighbourhood physical environment exposure variables. People with similar characteristics show geographic clustering which can lead to overestimation of the importance of any associations between outcomes and neighbourhood exposures. To overcome this problem multilevel modelling will be performed, showing how much of the variation in outcomes in attributable to differences in characteristics of individuals relative to that attributable to variation between neighbourhoods. Further, it will show how much of any variation that is found to be attributable to differences between neighbourhoods is accounted for by the measured neighbourhood environment exposures relative to unmeasured attributes.

A variety of sources of geographic data is necessary to construct neighbourhood physical environment exposure variables. The greenspace exposure variable will be constructed from the Generalised Land Use Database which gives proportions of this land use for a range of 2001 administrative geographies for England. The objective measure of environmental quality will be derived from the metadata of the 2001

Multiple Environmental Deprivation Index, a composite measure of air quality, greenspace, temperature, sunlight and proximity to industry, whist the subjective measure will be drawn from the 2008 Place Survey, which gives local authority-level satisfaction with parks and open spaces. The requisite geographical data in construction of the walkability index, in neighbourhood delineation and in walkability score attribution to neighbourhoods includes digitised boundary data, a land use mapping database (UKMap), census household counts and transport networks (Ordnance Survey MasterMap). Also the construction of particular forms of residential postcode-centred neighbourhood zones – polygon based network buffers – using GIS software requires transport network data.

The intention of this thesis, in summary, is to build a more solid foundation for understanding the contribution of neighbourhood physical environments to physical activity and health through appropriate modelling of expected causal pathways.

2 Literature review

2.1 Introduction

The purpose of this study is to investigate the relationships between characteristics – specifically greenspace, environmental quality and walkability – of neighbourhood environments and physical activity, and depression, in adults in the United Kingdom. This chapter provides the conceptual framework in written form as a review of the literature and in graphical form to illustrate the main concepts and variables to be studied and the presumed relationships between them. The literature review identifies findings, including inconsistencies, relevant to the present study, and gaps in the literature and methodological shortcomings that call for follow-up studies. It also provides definitions of terms. The conceptual framework serves to ensure the underlying assumptions are correct [8]. Literature reviewed here is predominantly from works published in English, in peer-reviewed journals relating to physical environments, physical activity and mental health, including depression. Relevant literature has been identified using large citation databases such as Medline and Web of Science, and also through internet searches. Search terms included, but were not limited to, the following: “neighbourhood”; “physical environment”; “natural environment”; “greenspace”; “environmental quality”; “physical activity”; “walkability”; “walking”; “mental health” ; “depression”; and “adults”.

Although this study is specifically concerned with examining relationships between neighbourhood physical environments and physical activity, and between physical environments and depression, the literature reviewed includes studies of relationships between neighbourhood physical environments and health outcomes such as obesity, for which physical activity is hypothesised to play a mediatory role. Inclusion of such studies in the review provides useful insight into specification of environmental exposures and evidence for pathways on which physical activity may lie. The depression study is concerned with outcomes of adults of all ages although, the physical activity outcomes are only studied in adults studied aged between 50 and 72 years, an age grouping which straddles middle and old age but does not fall clearly in either. Thus, this review has no particular focus on adults within a specific age range. Cities of North America and Australia provide the settings for the majority of studies examining associations between physical environments and physical activity and health outcomes, with European cities underrepresented in this field. However, limiting the literature review to studies conducted in Europe, the findings of which may be more pertinent to the present United Kingdom-based, and particularly London-based investigations, would risk missing potentially valuable and applicable evidence only because it has not yet been gathered in European settings. Therefore, North American and Australian studies are reviewed alongside European ones but with acknowledgement and summary of their important contextual differences.

Physical environment factors discussed in this review in relation to physical activity and health outcomes putatively mediated by physical activity fall in two broad categories, perceived and objective measures. The perceived measures are not readily classifiable as they are often constructed from numerous survey items designed to gauge the extent to which individuals feel the environment is conducive to physical activity. For example, participants may be asked to rate an environment for its overall quality, the criteria for which are open to interpretation but may include perceived number of facilities, aesthetic appeal and safety. In contrast, objective measures are subdivided further and considered here as accessibility and as atmospheric conditions. The review begins with an overview of the ecological model of health underpinning this study and the studies on which it draws. Four key concepts are then discussed; the physical environment, neighbourhoods, physical activity and depression. The review then presents evidence of relationships between physical environments and physical activity, together with evidence of relationships between physical environments and health outcomes, including depression. Lastly, studies of associations between physical environments and depression in which physical activity plays no explicit role are reviewed.

2.2 The ecological model of health and health behaviours

A model, or theoretical framework, is useful to understand how concepts relate to one another. The ecological model of health was developed because, at a population level, it is more efficient to prevent health problems than to treat them as they arise in individuals. Using the ecological model of health to give a broad understanding of the interplay between individual, social and environmental determinants of health, the situations that created the health problems can be identified and examined in context. These determinants of health are the characteristics in which living takes place [9], with those considered to merit attention potentially altered by informed action [10]. Sallis *et al* note that through the explicit consideration of multiple levels of influence, ecological models lead to the development of more comprehensive interventions [11]. Numerous ecological models of health have been constructed, largely based on Bronfenbrenner's Ecological Framework for human development, with development defined as the way a person perceives and deals with their environment and wherein multiple factors influence behaviours [12]. Bronfenbrenner saw these influences as layers surrounding the individual, each layer representing a different level of the environment. The individual is influenced by their social environment, such as their family, which is embedded within their physical environment, attributes of which include neighbourhood facilities. The social and physical environments are embedded in the policy environment. However, these layers of influence operate in several directions as there is a reciprocal relationship between an individual and their environments: an individual's behaviour is influenced by environments but there are also environmental impacts of their behaviours. Also, the environments are complex because they can be described in several ways, and these descriptions may be actual or perceived.

The mechanisms by which determinants of health affect health outcomes have been modelled in a multitude of ways over recent years. Dahlgren and Whitehead's diagram shows an individual's general position in relation to determinants of health in terms of the directness of their effects [13]. Thus, genetic factors are positioned next to individuals, surrounded by individual lifestyle factors. Social and community factors, forming the next layer, are surrounded by living and working conditions which, in turn are surrounded by socioeconomic, cultural and physical environmental conditions. Dahlgren and Whitehead posit that the outer layers, are more powerful than, and influence, the inner layers. In contrast, the emphasis of Diderichsen, Evans and Whitehead's model is not on the relative influence of determinants but specifically on the effect of social stratification on health [14]. Social contexts are shown to influence an individual's social position. Those in lower social positions are depicted as being more likely to be exposed to health damaging situations and as more likely to be susceptible, due to individual level factors, to health damage. In addition, lower social positions are shown to increase the adverse effect of the exposure to which individuals are more susceptible due, for example, to worse access to healthcare. Mackenbach's model of "selection and causation" illustrates how health inequities are generated [15]. Certain factors predispose individuals – or "select" individuals for assignment – to particular socioeconomic positions. For example, poverty in childhood might place an individual in a low socioeconomic position in adulthood. The socioeconomic position then "causes" lifestyle factors, such as smoking, and thereby poor health. The model also shows how a factor which selects an individual's socioeconomic position can also be a direct cause of a health problem. The emphasis of Brunner, Marmot and Wilkinson's model of the determinants of health is on how environmental, psychological and behavioural risks are accrued and interact across the life course, resulting in health inequalities [16].

Combined, these models are useful in explaining how individuals become predisposed to poor health and the role of different stages of life, and how circumstances interact to increase health inequalities. However, none have a spatial dimension to illustrate how exposure varies according to physical location. In modelling physical activity as a behaviour, it is useful to emphasise the role of physical environmental factors, with account of other influences, such as social and psychological ones. Whilst individual factors are essentially "stuck" to the individual, and social, cultural and political determinants envelope them, exposure to the neighbourhood physical environment varies according to their movement within it. Given the focus of the present research on physical environmental determinants, a more useful schematic might depict an individual carrying their individual, social, cultural and political "baggage" moving within their neighbourhood environment and across the life course. The nature of interactions between determinants would also vary according to an individual's spatial position. For example, the mobility, and thereby physical activity, of an older person in an area of high traffic density might be more restricted than that of a younger person who might perceive such an environment as less threatening.

2.3 The physical environment

The environment is everything outside the individual, classified as physical, sociocultural, economic and political [17]. The physical environment comprises the built environment and the natural environment. The natural environment encompasses living and non-living things, such as vegetation, microorganisms, rocks and soil, as well as natural resources and phenomena such as air, water, radiation and climate. By contrast, the built environment is defined as aspects of our surroundings that are modified by people, such as buildings and spaces – including homes, schools, workplaces, parks and business areas, energy infrastructure and transport systems [18]. Thus, the definitions of natural environments and built environments are not mutually exclusive: a city park, for example could be considered a part of the built environment if subject to substantial human intervention or, if largely "untamed", as a component of the natural environment. Arguably, all aspects of environments surrounding human settlement, and which may impact human health, are to some extent modified by humans, rendering the term "physical environment" more appropriate than the distinct labels of "natural" or "built".

Data on physical environments collected to investigate relationships with physical activity and health outcomes can be primary or secondary. Primary data is usually collected through surveys of individuals reporting on environmental features and thus constitutes perceived measures. Systematic social observations by field work auditors can be used to collect objective measures [19], although objective measurement of potentially important determinants such as aesthetics is difficult [20]. The advantage of primary data is that it is specifically designed to address particular research questions whereas secondary data, such as that sourced from a census, is limited in terms of its availability, content and scale. However, secondary data is cheaper and usually objective, giving unbiased facts when related to physical activity or health outcomes. Nevertheless, many argue there is need for the investigation of both perceived and objective measures of physical environments [21][22]. There is evidence for concordance between actual and perceived measures of certain physical environment attributes such as walkability[21][23]. Interestingly, a recent study found perceived air pollution in London to be significantly positively correlated with objectively measured traffic-related pollutant, suggesting that exposure to air pollution is perceptible and that perceived measures may be a proxy and cheaper alternative [24]. Many studies, however, suggest considerable discordance between actual and perceived measures [25][26][27] and differential relationships with physical activity outcomes [28] [29][25].

2.4 Neighbourhoods

The investigation of associations between neighbourhood physical environments and health outcomes requires specification not only of the physical attribute of the neighbourhood, such as land used mix, that is likely to influence the outcome of interest, such as physical activity, but where it will exert its effect; how to operationalize the neighbourhood. An early spatial definition of neighbourhood, conceived by urban planner Clarence Perry to guide the development of desirable neighbourhoods in industrialising cities, was a unit that contained four basic elements: a primary school, small parks, small shops, and buildings and streets configured to allow all public facilities to be within safe pedestrian access [30]. A less specific definition that is perhaps more amenable to researchers is as a person's surrounding environment [31], represented as an area unique to each member of a given study population. When the outcome of interest is walking, or indeed being engaged in any physical activity, exposure necessitates that the individual has had contact with the physical environment by moving through it by whatever mode of transport: an individual in a neighbourhood with environmental attributes supportive of walking would be expected to be more physically active within it. However, associations will only be found if the exposures are captured in an area that is sensitive to walking or other physical activity and whilst Perry's neighbourhood unit might describe the scale well, its boundaries are not explicitly defined. Thornton *et al* argue that the geographical extent of an activity space, which they define as the locations visited by an individual within a specified time period, is unique to the individual as a product of both environmental and individual-level factors [32]. The proximity of resources determines how far an individual must travel, for example, whilst his or her age influences their willingness and ability to do so. The location of an individual can be monitored whilst they undertake their physical activity using, for example, a global positioning system (GPS), with the GPS trace representing the exact location of the physical environment exposure of interest [33]. A neighbourhood defined by GPS tracking represents the ideal, especially as the premise that an individual's residence is geographically central to their neighbourhood may be false; it could be that an individual spends a short period of time in their local home environment relative to that spent in their local work one. It reflects true exposures more accurately, improving the specificity between the exposure and outcomes [32]. However, deployment of the requisite equipment and the use of study participants as collectors of their own geographical data make this approach impractical for the large study samples needed to detect small but significant effects. Given this practical constraint, compromise is necessary through selection of the spatial units that best define the neighbourhood for study participants.

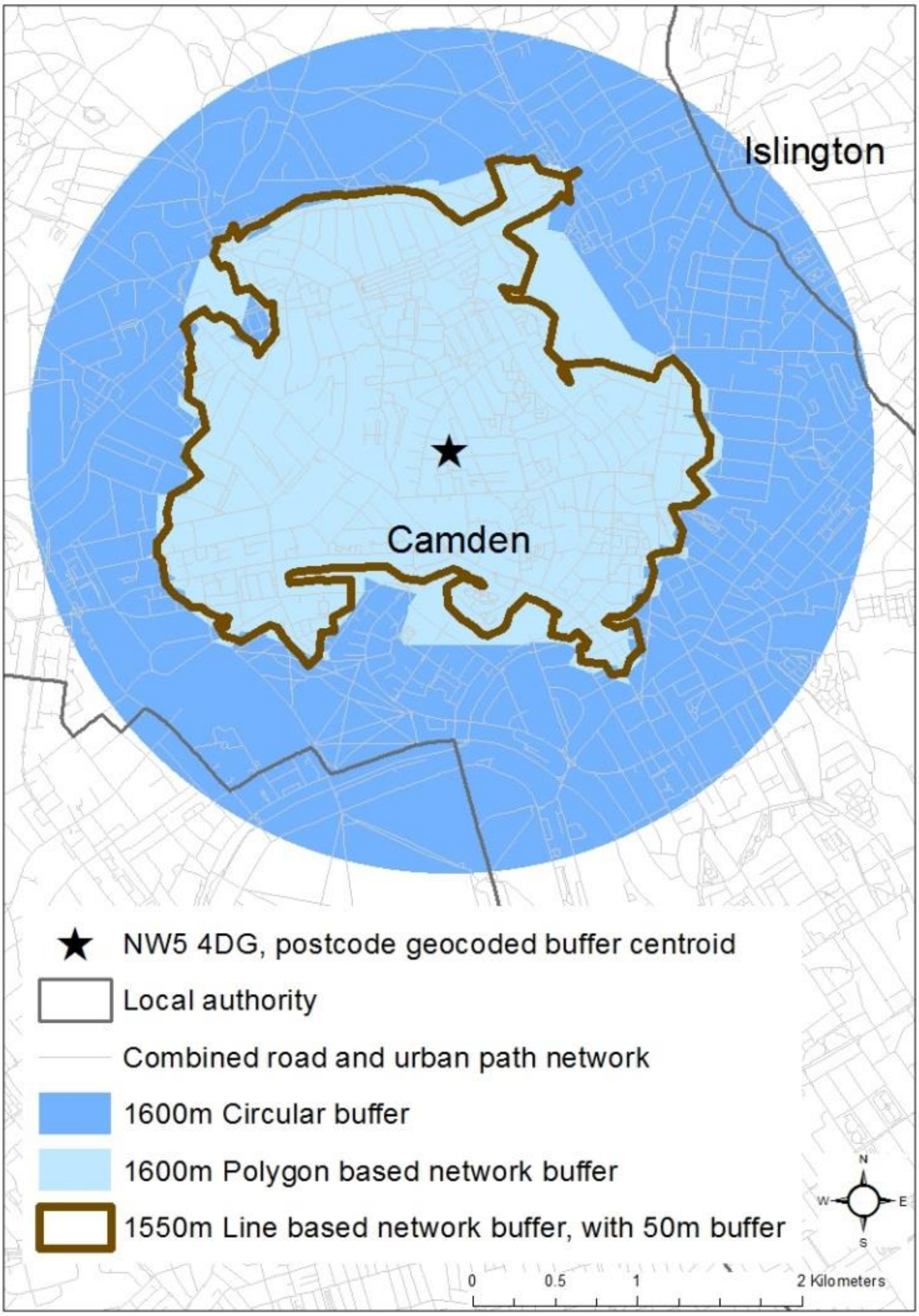
Selection of the optimal spatial units of analysis – the areas in which environmental data pertaining to the attribute are aggregated – to construct an appropriate environmental exposure variable is a critical element of study designs: the choice of units, with respect to number and spatial boundaries, for aggregation of spatial data influences the quantification of the environmental exposure and, therefore, the nature of any apparent relationship with the outcome. The dependencies of the number of units and of the boundaries that spatially define the units on the results of analysis constitute the scaling effect and the zoning effect, respectively, of the Modifiable Areal Unit Problem (MAUP) [34][35]. These dependencies are problematic for researchers investigating neighbourhood effects because the choice of units is limited by the geographical levels at which environmental data and outcome data have been collected.

Identifying the residential location of each study participant is often the first step in defining neighbourhood on the premise that the residence must lie within the neighbourhood. In many studies the residential location is only identifiable to an administrative geography such as a local authority so neighbourhood can only be defined as this geography, or a higher one. Environmental data is often collected within and aggregated to administrative areas, making "off the shelf" measures attractive for researchers [36]. However, there might be considerable variation in an environmental attribute within a larger administrative area such as a local authority that is effectively "ironed out" by aggregation. At larger scales, and thus with fewer units, between-area variation in the values of aggregated environmental data tends to be less – the scaling effect of the MAUP – and the ability to detect any differential relationships of neighbourhood attributes with health outcomes may be reduced. It must be noted that smaller units, between which variation in measures of environmental exposure is greater, do not necessarily constitute better units of for spatial analysis [37]; the optimal units are those that encompass the area where the hypothesised effect on the outcome operates. Defining the neighbourhood as a smaller administrative unit in the case of examining physical environment effects on physical activity, however, is potentially better than defining it at a larger scale because attributes hypothesised to be associated with neighbourhood walking or other physical activity are aggregated in units "closer to home", within walkable distance [38]. If study participants are identifiable to residential postcode level, essentially they can be "pin pointed" on a map as the postcode represents a unique point on the Earth's surface, with a specific longitude and latitude. Neighbourhoods of individuals for whom a residential postcode is available can therefore be defined at any administrative geography, including the smallest in the United Kingdom as Output Areas, identified using GIS as that in which the postcode falls. However, addressing the scaling effect of the MAUP by selecting the most relevant areal size does not necessarily address its zoning effect: at any given scale differential delineation of unit boundaries will change the values of the aggregated data and therefore the results of the analysis. Indeed, it has been argued that of the two the MAUP issues of scaling and zoning, the latter is more important because the shapes of zones into which an area can be divided are infinite [37]. Stafford *et al* investigated the extent to which the zoning effect influenced the estimates of health inequalities – including inequalities in walking and in self-rated health – in two Inner London boroughs [39]. They delineated neighbourhoods using Census Area Statistic (CAS) ward boundaries, using physical and man-made features of the environment, and as socioeconomically homogenous zones, and employed multilevel modelling to compare the level of inter-area variation in each outcome with each method of delineation. Finding the level of inter-area variation to be of similar magnitude, irrespective of neighbourhood definition, Stafford *et al* concluded that despite the weak theoretical basis for administrative boundary-defined neighbourhoods in the study of health inequalities it did not substantively affect estimates. Another study, however, which took into account perspectives of residents, found evidence for the potential of administratively defined neighbourhoods to underestimate neighbourhood effects on health [40]. In response to concern over the possible mismatch between researcher and resident-defined neighbourhoods, Coulton *et al* conducted a qualitative study in the city of Cleveland in the United States to test several methods of defining neighbourhood units. Their comparison of maps drawn by residents with census definitions of neighbourhoods revealed considerable differences with respect to zones covered, although the areal sizes of resident-defined neigbourhoods were similar to those of their census-defined neighbourhoods, suggesting that arbitrary boundaries of administrative areas may inadequately capture the neighbourhood to which can individual is exposed. However, as discussed, for analysis of large study samples it is rarely practicable to use resident-defined neighbourhoods. Therefore, a compromise may be to construct individualized representations by using geographical information systems (GIS) to delineate neighbourhoods as buffers. A buffer is a boundary placed around an area or a point using a predefined scale using either a straight-line (Euclidean) or network distance [32]. Generated using a Euclidean distance, a buffer constitutes a circular area with the centre as the area or point, such as an individual's geocoded residential postcode. The Euclidean distance can be set to a walkable distance so that the circular buffer represents a walkable area, extending to the maximum distance individuals would be prepared to walk from their residential location to reach places. Empirical evidence suggests that 1km is the distance that adults walk to reach places from home [41] and surveys indicate that people perceive areas within 1km from home to be part of the neighbourhood [42].

Circular buffers are commonly used to represent neighbourhoods because they are generated relatively quickly using GIS software but they do not account for restrictions that the road network places on an individual's ability to access destinations. The fact that a circular buffer neighbourhood can be described as walkable as the crow flies, illustrates the flaw in its design. Individuals cannot fly above streets and a circular buffer, which takes no account of road networks, may contain large areas that are not visible or that are at least inaccessible to the individual and which, therefore, are unlikely to contribute to any neighbourhood physical environment effects on their physical activity. The network buffer representation attempts to address this oversight by limiting the neighbourhood to that which is within the specified walkable distance along the road network from the individual's residential location. By taking account of the road network and thereby only including accessible places the network buffer is argued to offer a more sensitive representation of the neighbourhood than a circular buffer [31]. A polygon based network buffer is formed by joining the end point of each road of the specified walkable distance, each end point constituting a vertex to form an irregular polygon, whereas a line-buffered network buffer is delineated as a specific distance around each of the roads of the specified walkable distance. There is some evidence that associations between measures of the physical environment and walking are greater in line-buffered network buffers, possibly because they represent only the neighbourhood that would be visible to an individual along the walking route, providing awareness of opportunity for walking or other physical activity [43]. However, the construction of line-buffered network buffers is computationally more demanding than that of polygon based network buffers, rendering their application in larger studies more onerous if many have to be created. Also, the cruder definition of neighbourhood achieved using the polygon-based network buffer may include areas of land that are beyond the immediate visible vicinity of a road, such as open spaces of which an individual may be aware and which may be important for explaining physical activity behaviour. Thus, polygon-based network buffers neighbourhoods are not without merit. Within dense grid-like urban road networks polygon-based network buffers have been shown to approximate the size and shape of line based ones whereas, within lower-density, irregular suburban road networks, polygon-based network buffers constitute larger, more regular polygons. Decisions regarding the type of buffer and its scale, which determine the shape and size of the neighbourhood, should be based on the hypothesised effect of the physical environment exposure and the outcome of interest [44]. Examples of a circular buffer, a polygon based network buffer and a line-buffered network buffer are illustrated in Map 2.1.

In addition to delineation as administrative units, researchers have operationalised neighbourhoods as grid cells which reduces the subjectivity of boundary selection. Krizek used GIS software to create a neighbourhood accessibility index for a pedestrian scale resolution network of 150 metre by 150 metre grid cells, providing continuous measures of housing density, population and street connectivity for each grid cell [45]. To account for the influence of proximal cells, the score for each grid cell was calculated by averaging values for all grid cells across a one-quarter mile radius of it. Operationalisation of neighbourhoods as per Krizek's method is particularly attractive for the construction of indices such as those for walkability because it enables spatial interpolation, which is unfeasible for neighbourhoods delineated as irregularly shaped and sized administrative units, or as spatially discontiguous individualised buffers.

Map 2.1 Examples of a circular buffer, a polygon based network buffer and a line-buffered network buffer.



2.5 Physical activity

2.5.1 Definitions and health guidelines

Caspersen *et al* differentiate the concepts of physical activity and exercise [46]. They define physical activity as follows:

> *"…any bodily movement produced by skeletal muscles that results in energy expenditure. The energy expenditure can be measured in kilocalories. Physical activity in daily life can be categorized into occupational, sports, conditioning, household, or other activities*.

Exercise is distinguished as:

> *"…a subset of physical activity that is planned, structured, and repetitive and has as a final or an intermediate objective the improvement or maintenance of physical fitness".*

The terms physical activity and exercise are often used interchangeably yet physical activity is not always undertaken in the form of exercise. Caspersen *et al* argue that precise terminology must be used in research because measurement of each concept is unique [46].

As physical inactivity is a major risk factor for morbidity and mortality, guidelines are issued on the frequency, duration and intensity of physical activity necessary to reduce an individual's risk of diseases such as coronary heart disease, stroke, diabetes, breast and colon cancer and depression. The World Health Organisation's (WHO) current recommendation is that adults aged 18–64 should do at least 150 minutes of moderate-intensity aerobic physical activity throughout the week. Alternative targets recommended are at least 75 minutes of vigorous intensity aerobic physical activity throughout the week or an equivalent combination of moderate - and vigorous-intensity activity" [47]. The primary United Kingdom-specific guidelines are identical to those of the WHO but the UK ones additionally include recommendations regarding strength, limiting sedentary behaviour and balance [48]. Physical activity to improve muscle strength on at least two days a week is recommended for adults aged 19–64 who are also advised to minimise the amount of time spent being sedentary – sitting – for extended periods. UK guidelines are also offered specifically for adults aged 65 and over to follow recommendations for people aged 19-64 but additionally the incorporation, on at least two days a week, of physical activity to improve balance and co-ordination is advised for those at risk of falls.

The WHO identifies four physical activity domains as work (including non-manual labour), transport (such as walking or cycling to shop or work), domestic duties (such as housework) and leisure (sports and recreational activities) [47]. Types of physical activities differ in their intensities and the energy expenditure they demand so it is important to measure engagement in physical activity in terms of intensity as well as time. Physical activity may have important effects on health independent of intensity, acting via social contact for example. Therefore, it is also useful to measure physical activity in terms of volume alone as the product of duration and frequency. A physical activity can be assigned a value to indicate its average energy cost, expressed as Metabolic Equivalent of Tasks (MET), as listed in a reference manual developed for standardisation of physical activity measurement [49]. One MET equals the resting energy expenditure rate so an activity of 3 METS, for example, spends three times the energy expenditure of resting.

2.5.2 Non-physical environmental determinants of physical activity

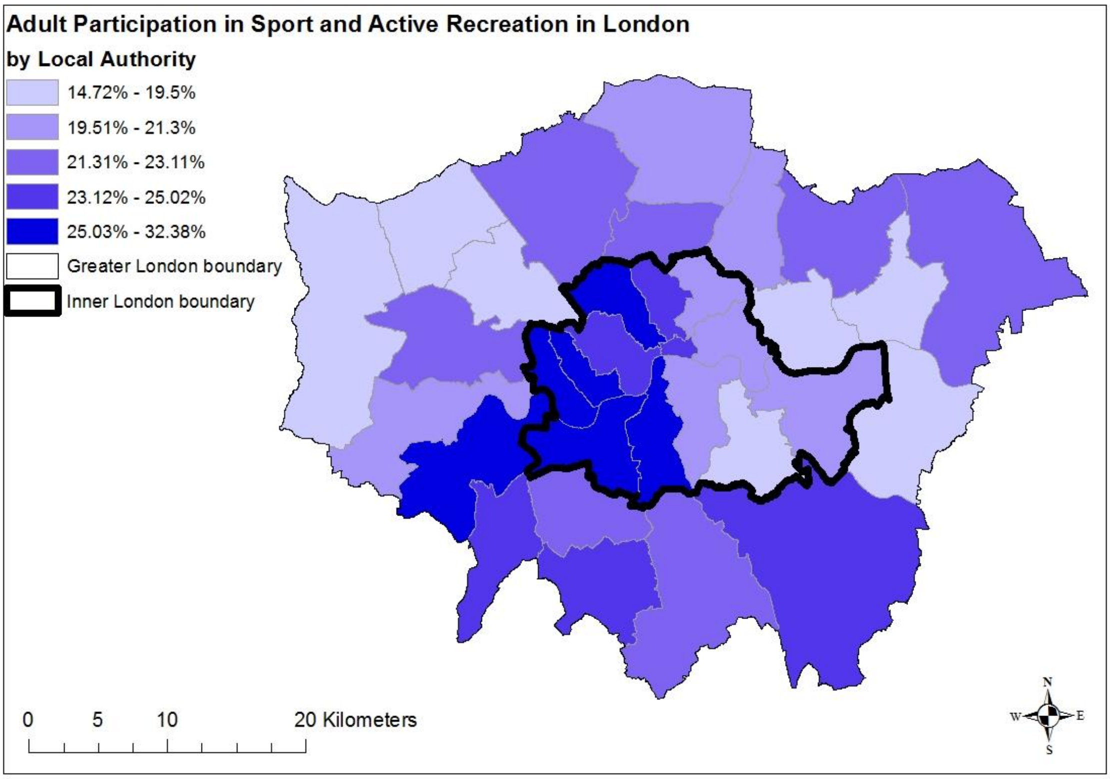
Bauman *et al* summarised present knowledge on correlates of all physical activity in adults based on evidence from systematic reviews, identifying individual and social factors [7]. They found the strongest evidence for health status and self-efficacy as determinant factors and weaker evidence for personal history of physical activity during adulthood as causative as well. Being younger, being male, having a higher level of education and being of white ethnic origin were also found to be associated with higher levels of higher overall physical activity, and being of higher occupational social class to be associated with higher leisure-specific physical activity but being of lower occupational social class with overall physical activity. In addition, Bauman *et al* identified not being overweight, lower perceived effort, and greater social supports as significant non-physical environmental determinant correlates.

Adams used data collected in the 2005 UK Time Use Survey to examine the prevalence and sociodemographic correlates of physical activity specifically as active transport, defined as any walking, jogging or cycling for purposes other than enjoyment [50]. Younger people, those without access to a car, those of lower occupational social class, those not in paid employment and those who left education at an older age were significantly more likely to report spending 30 minutes or more per day engaged in active transport. It is interesting that whilst Adams identified occupational social class to be negatively associated with active transport participation, Bauman *et al* found it to be positively associated with leisure-related physical activity. This may reflect the greater economic means and “free time” associated with higher occupational social class which enable individuals to choose to participate in the “fun” physical activities over undertaking the non-negotiable chore-driven active transport demanded of those of lower occupational social class. In light of Adams findings, the fact that overall physical activity was found by Bauman *et al* to be inversely associated with occupational social class suggests that a large proportion of physical activity is errand- rather than enjoyment-based. However, in terms of undertaking a level of physical activity recommended for health, accounting for both volume and intensity, there is evidence for a positive association with greater wealth. The Health Survey for England Report 2008 noted that equivalised household income was positively associated with meeting the then Chief Medical Officer’s recommendation of taking part in physical activity of at least moderate intensity of 30 minutes or more duration on five or more days [6]. Those in the lowest income quintile were less likely to achieve it than those in the higher income quintiles, although an income dose response was only found in women. Thus, moderate and high intensity physical activities that contribute to achieving a level recommended for health may be more prone to fall in the recreation domain, more accessible to those of greater wealth and higher occupational social class.

2.5.3 Geographic variation in physical activity across the United Kingdom and across London

There is evidence for geographic variation in physical activity across the United Kingdom [51]. Local authority estimates for adult participation in sport and active recreation reveal geographic variation across England. Data collected from April 2011 to April 2013, showed the percentage of the adult population (age 16 years and over) in a local authority who participate in sport and active recreation, at moderate intensity, for at least thirty minutes on at least twelve days out of the last four weeks (equivalent to thirty minutes on three or more days a week), varied between 14.7 and 32.4 per cent. Spatial patterning across England (data not shown) was not apparent. However across London which was home to both the least and the most active English local authorities, spatial patterning was more obvious (Map 2.2). Westerly local authorities in Inner London tended to have higher rates of participation than Easterly ones but there was no discernible geographic patterning in Outer London. There was however no adjustment for social class which would be likely to influence spatial patterning.

Map 2.2 Rates of adult participation in sport and active recreation in London by local authority represented as London-based quintiles, with the lightest shade indicating the lowest quintile. Source of data from which map was constructed: Coyle: Health survey for England 2006 London boost: health and lifestyle in London main report [52].



2.6 Depression

2.6.1 Definition

Mental health is "a state of well-being in which an individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community" [53]. The importance of mental health for individuals and wider society is expressed succinctly by the World Health Organisation constitution which states that there is no health without mental health [54], yet by 2020 depression is expected to be second only to cardiovascular disease among the largest contributors to the global burden of disease [55]. Whilst common, this mental disorder does not simply constitute unhappiness: the Diagnostic and Statistical Manual of Mental Disorders (DSM), a manual which classifies mental disorders provides extensive standard criteria for a diagnosis of depression [56]. Briefly, a person who is depressed has, for at least a two-week period, experienced five or more of the following symptoms:

- Depressed mood
- Markedly diminished interest or pleasure in activities
- Significant weight loss, or decrease or increase in appetite
- Insomnia or hypersomnia
- Psychomotor agitation or retardation
- Fatigue or loss of energy
- Feelings of worthlessness or excessive or inappropriate guilt
- Diminished ability to think or concentrate
- Recurrent thoughts of death

The Structural Clinical Interview (SCID), a clinical interview that uses the DSM criteria for illness, is the gold standard for a diagnosis of depression. However, in research contexts for population studies, the General Health Questionnaire (GHQ) was specifically devised, and is used widely, to measure minor psychiatric morbidity (MPM), including symptoms of anxiety and depression [57]. The questions are designed to capture individuals' levels of happiness, symptoms of depression and anxiety and sleep disturbance over the last four weeks, and the questionnaire exists in a number of forms varying in terms of length and scoring.

2.6.2 Non-physical environmental determinants of depression

There are several factors associated with depression which have been identified through traditional biological research as determinants of the disease, such as specific genes, and a family history accounts for almost 40% of variance in depression, suggesting genetic factors play an important role [58]. Furthermore, physical illnesses, such as diabetes and cardiac disease [59] increase the risk of depression and encourage the search for physical causes. However, the natures of the relationships between the many non-biological factors and depression are unclear: explaining disease through socio-ecological modelling is more complex. A survey of adults aged 16-74 in private households carried out in 2000 for the Office for National Statistics sought to estimate the prevalence of psychiatric morbidity including depression and to establish factors associated with it [60]. Using a sampling strategy similar to that employed in the Health Survey for England, the survey drew a sample representative of all households in England, Wales and Scotland. The response rate was 70% and information was collected from the 8800 participants through a structured interview followed by a clinical interview in a subset of 600. The survey showed that cases of neurotic disorders, including depression, were more likely to be women (59% among cases compared to 48% among non-cases), aged 35 to 54 (45% among cases compared to 38% among non-cases) and separated or divorced (14% among cases compared to 7% among non-cases). In addition a higher proportion of depression cases were economically inactive (39% among cases compared to 28% among non-cases). In a meta-analysis to evaluate associations of socio-economic factors with depression, Lorant *et al* [61] found significant associations with socioeconomic status, education and income. Those in the lowest socio-economic group had higher odds of being depressed (OR=1.81, p<0.001). Dose responses were found for education, with the log odds of being depressed decreasing by 3% for each additional year in education, and for income, with a 0.74% decrease in the log odds of depression for each 1% increase in relative ranking. Using data collected in the 1999 Health Survey for England from 4281 adults aged 16-74 years living in private households in England, Weich *et al* (2004) found common mental disorders including depression to be more prevalent among non-whites than whites [62].

2.6.3 Geographic variation in depression across the United Kingdom

Social and economic resources are not evenly spread across areas of the United Kingdom – some areas have more than others – and there is evidence that suggests people living in more socioeconomically deprived areas are more likely to suffer longer spells of depression [61][63]. However, a nationally representative longitudinal study of UK adults by Weich *et al* found statistically significantly variation (12%) in the onset and maintenance of depression only at the level of the household and the individual [64]. Between CAS wards, the level at which socioeconomic deprivation was measured and of which there are 7969 in England with an average population of 5500, variation was only 0.2% and was not significant after adjusting for individual and household characteristics. Weich *et al* explain the counterintuitive finding that depression onset and maintenance is no more likely in the most deprived wards than the least deprived by the choice of spatial scale. To protect the identity of study participants they were restricted to measurement of deprivation within wards. However, the authors posit that the variation they detected between households may have been attributable to variation between spatial units larger than households but ones smaller than the ward – to which measurement of socioeconomic deprivation was restricted – as there is evidence for associations between common mental disorders and features of the physical environment in smaller areas [65][66]. Another factor proposed to account for failure to detect variation between wards in depression onset and maintenance is that the measure of socio-economic deprivation used does not capture aspects of the environment with the greatest impact on depression. Weich *et al* suggest that other place-based effects, such as residential mobility, may have significant impacts. Others have found morbidity variability to be linked to within ward variation in deprivation [67]. In summary, the spatial scales at which place matters for depression have yet to be determined and analyses of associations with physical environments should employ spatial scales that reflect hypothesized effects.

2.7 Relationship between physical activity and mental health, including depression

The relationships between physical activity and mental health, including depression, are complex and multidirectional and difficult to disentangle, largely due to associations of both constructs with physical health. Bi-directional relationships between physical health and mental health conditions such as depression are well documented. For example, depression has been associated with increased risk of cardiovascular disease [68] whilst cerebrovascular disease has been linked with increased risk of depression [69]. There is evidence to suggest that physical activity has both direct biological and psychological effects on mental health. However, physical health confounds the relationship between physical activity and mental health, with those in better physical health more likely to be physically active [7]. With evidence for biological effects of physical activity that improve physical health, a determinant of mental health, to some extent physical activity accounts for, or mediates, the relationship between physical health and mental health. Those in better mental health as indicated by stress level, however, are more likely to be physically active [7], confounding the relationship between physical activity and physical health.

Hamer *et al* found strong associations between self-reported physical activity, of any type, at a minimum level of at least twenty minutes per week and reduced odds of psychological distress among a representative sample of adults from the 1995, 1998 and 2003 Scottish Health Surveys, as indicated by a GHQ-12 score of four or more [70]. These associations were dose-responsive, with greater risk reduction observed with higher levels of physical activity, particularly sports. Their finding that these associations remained significant after adjustment for long-standing illness and obesity, suggests that they are independent of these factors and physical activity may have direct biological effects on mental health. This theory is consistent with evidence that physical activity reduces dyslipidaemia, glucose intolerance, inflammation and vascular dysfunction, conditions related to mental health disorders such as depression and dementia [71][72][73]. Another biological pathway through which physical activity may operate is regulation of responses to proteins released upon exposure to acute mental stress. Physical activity could lead to physical fitness-associated down-regulation of proteins released upon exposure to stress [74] and thereby reduce risk of psychological morbidity, which is associated with heightened responsiveness to such proteins [75]. In addition, the independent associations found by Hamer *et al* could indicate physical activity acts through psychological and psychosocial pathways. Self-reported psychological wellbeing has been shown to be higher among those reporting higher levels of moderate to vigorous physical activity [76], possibly accounting for the negative association between physical activity and psychological distress. Also, physical activity may reduce risk of psychological distress by improving self-esteem, self-efficacy, perceived behavioural control and cognitive functioning [77][78]. Whilst the associations identified by Hamer *et al* remained significant after adjustment for longstanding illness and obesity, they were reduced. Given that there is evidence that physical activity plays a role in the prevention of many chronic diseases [79], this suggests that physical activity may also reduce the risk of psychological distress through long-standing illness and obesity. For example, physical activity may reduce the likelihood of diabetes in an individual, itself a risk factor for psychological distress.

2.8 Relationships between physical environments and physical activity, and physical activity-mediated health outcomes

2.8.1 Physical environments measured in terms of perceived support for physical activity

Many studies have examined perceptions of the physical environment in relation to mainly self-reported physical activity. In a cross sectional survey of 4265 adults conducted in England, Foster *et al* investigated associations between self-reported walking behaviour and perceptions of the attractiveness and accessibility of the neighbourhood physical environment for walking [80]. They found that only in men was perception of the environment, specifically access to a park, significantly associated with 150 minutes per week in the past four weeks, indicative of meeting the recommended physical activity level [48]. However, others have found stronger evidence for the importance of perceived measures of the environment as determinants of physical activity. In a longitudinal study conducted in Australia, Sugiyama *et al* found significant positive associations between positive perceptions of the presence of, and proximity to, green spaces and walking maintenance over four years as indicated by self-reported walking frequency [81]. Also, in a review of studies – the majority of which employed perceived physical environment measures – Humpel *et al* found that accessibility of facilities, opportunities for activity and aesthetic attributes [82] were associated with self-reported physical activity. Perception of overall environmental quality of neighbourhoods, an attribute that cannot be measured objectively, has been found to be related to transport-related physical activity [83][84].

There is evidence, however, that characteristics of individuals influence how their perception of neighbourhood physical environments relates to their physical activity behaviour. Shigematsu *et al* surveyed 1623 adults in Seattle, United States, using the validated Neighbourhood Environment Walkability Scale questionnaire [85] to collect perceptions of physical environment attributes of neighbourhoods, defined as the area within a fifteen minute walk from home, and the International Physical Activity Questionnaire [86] to gather self-reports of frequency and duration of walking [87]. Attributes included residential density, proximity to non-residential land uses, ease of access to non-residential uses, street connectivity, walking and cycling facilities, aesthetics, pedestrian traffic safety, crime safety, and proximity to recreation facilities. Participants were purposefully selected from a range of areas differing in objectively measured walkability to ensure wide variability in the physical environments. Recreational walking was not significantly related to any perceived attributes for any age group but almost all perceived attributes were significantly associated with walking for transport in 20–39 year olds, the youngest age group. In contrast, only proximities to facilities such as shops and parks were associated to walking for transport in the oldest age groups, comprising those aged 66 and over, and the associations were the strongest found in the study. Given that older people are more likely to have physical impairments that render them less mobile this finding makes sense. However, a small study of older adults (n=60) found that perceived access to services was only marginally significantly associated with physical activity objectively measured by pedometers, after adjustment for individual level factors [88]. It is important for public health and urban planning professionals to understand and consider the influence of perceived support for physical activity within the neighbourhood on physical activity before making any public health promoting modifications. Safety concerns in a neighbourhood for example, might deter physical activity within and thereby reduce the public health value afforded by a greenspace. However, the pattern of results of a study examining interactions between perceived safety and built environment variables in explaining physical activity among adults suggested safety related to crime, traffic, or pedestrian infrastructure did not have a moderating effect [89].

2.8.2 Objectively measured physical environments

2.8.2.1 Physical environments measured in terms of accessibility

Accessibility is the ability to reach opportunities (desired goods, services, activities and destinations) and can be described as a product of mobility and proximity which is enhanced by either increasing the speed of getting between point A and point B (mobility) or by bringing points A and B closer together (proximity) [90]. Factors that affect accessibility include the geographic distribution of activities and destinations, and the directness of links and density of connections in path and road networks [91]. As accessibility reflects opportunity costs to individuals in terms, for example of time or money, it influences travel behaviour and therein physical activity. Measures of accessibility reviewed here include density, diversity of land use (land use mix) and presence of, and distance to, opportunities, and street connectivity (junction density). Accessibility is reviewed in relation to physical activity and health outcomes based on the *a priori* hypothesis that if it is objectively easier to access services and recreational facilities by foot or bicycle people will be more likely to do so. Also, a greater density or higher presence of a recreational facility such as a park within a neighbourhood could facilitate physical activity within it, even if reached by car.

One of the simplest and commonly used measures of accessibility reflecting the geographic distribution of activities and destinations is the mean distance between destinations [92]. In addition, measures of accessibility commonly used that reflect the geographic distribution of activities and destinations are presence or proportion of a particular land use within an area, such as residential density, and land use mix (LUM), which may be included as factors in composite walkability indices. Another measure of accessibility frequently studied in isolation, or within a composite walkability index, is street connectivity [93]. This measure indicates accessibility as reflected in the directness of links and density of connections in path and road networks. Whilst many studies evaluate a range of individual environmental attributes as exposure variables in relation to a physical activity or health outcome, others include them in a single composite measure, such as the aforementioned walkability index. Here, studies evaluating individual environmental attributes as exposure variables in relation to a physical activity and health outcomes will be reviewed first, followed by those examining composite measures in relation to these outcomes.

2.8.2.1.1 Distance to, and presence of, greenspace

Natural environments, particularly those in urban areas, are often regarded as recreational stomping grounds, and planning policy [94] requires purposeful incorporation of greenspaces, like parks, into cities. Despite a lack of robust evidence for the health benefits of greenspace, a recommendation is that people in urban areas should be able to access a green space of at least two hectares within a five minute-walk of their home [95]. Therefore, access to greenspace, as indicated by proximity to an individual, or as the proportion within the neighbourhood area, has been the focus of many studies seeking to identify physical environment determinants of physical activity, and health outcomes likely to be mediated by physical activity. As a resource that is theoretically accessible to all, and free to use, there has been much interest in examining putative benefits to socioeconomically disadvantaged populations. Maas *et al* found physician-assessed morbidity to be negatively associated with amount of objectively-measured greenspace within a radial area around participants' GP practices [96]. The association was stronger within a one kilometre radius than within a three kilometre radius, for lower socioeconomic groups. However, in the most urban areas – those with the highest residential density – there was no relationship. The authors posited that urban greenspace is lower quality in the most densely populated, explaining the absence of positive health effects. They hypothesised that the stronger association within the smaller radial area was attributable to higher exposure to greenspace closer to home, and that the stronger association for poorer people was because they are more likely to be economically inactive, spending more time exposed to greenspace in their neighbourhoods. Maas *et al*'s findings are echoed by those of Mitchell and Popham who investigated the impact of objectively-measured greenspace on income-related health inequalities in an ecological study of the population of England (n=40,813,236) [97]. Their study showed that populations exposed to the highest quintiles of greenspace proportion in their neighbourhood, operationalised as the Lower Super Output Area of residence, had the lowest incidence rate ratio for mortality from circulatory diseases, suggesting that access to greenspace counters the negative health effects associated with low socioeconomic status. Due to the cross-sectional design of Mitchell and Popham and Maas *et al*'s studies, it is not possible to infer that greenspace causes better health outcomes and, as Lee and Maheswaran conclude in their review, whilst most studies show greenspace is positively associated with health, evidence for a causal link is weak [98]. Nevertheless, the hypothesis that greenspace improves health by facilitating physical activity is an attractive one, which has led many researchers to investigate associations between greenspace and physical activity.

Coombes *et al* examined the relationship of access to objectively-measured greenspace with self-reported use of it, and with self-reported physical activity, as indicated by meeting physical activity guidelines [48], among 6821 adults in Bristol, United Kingdom [99]. Self-reported use declined as objectively measured distance increased and, even after adjustment for individual factors and area deprivation, odds of meeting the physical activity guidelines declined as distance to formal parks increased. Their findings suggest that greater use of greenspace was facilitated by greater proximity to greenspace, and that the higher odds of meeting physical activity guidelines were attributable to greater proximity to, and possibly use of, greenspace, specifically as parks. A mediatory role of physical activity in the relationship between greenspace and health is plausible in light of Maas *et al*'s finding that greenspace close to home was more strongly associated with better health.

A study of 10,286 adults conducted in Brisbane, Australia, examined network distance to the nearest park in relation to self-reported minutes spent walking in the previous week [100]. After adjustment for individual factors, the likelihood of walking for more than 30 minutes (relative to less than 30 minutes) was not significantly higher for those closer to the nearest park. Another smaller study (N=380) conducted in Ontario, Canada, examined the association between distance to parks from residents' homes and using parks for physical activity [101]. It found that whilst there was no significant association between the distance to a park, or indeed its size, and park use for physical activity, there was a significant positive association between physical activity and the quality of the park's objectively audited features, particularly its footpaths. The importance of accounting for factors that may interact with proximity to parks as a putative determinant of physical activity is suggested by research by Giles Corti *et al* [102]. They constructed gravity-based models [103], which take account of both access and desirability of destinations, to examine accessibility in relation to park use. Their study, conducted in Perth, Australia, of 1803 adults found that the fullest model, which adjusted for distance, size and attractiveness was most strongly associated with undertaking six or more walking sessions per week totalling more than 180 minutes using parks for physical activity. Those in the top quartile for access as defined by this model were more likely to undertake this volume of physical activity than those in the bottom quartile.

In an effort to further characterise associations between physical activity and neighbourhood greenspace, many have investigated the relative influences of greenspace on recreational and transport-related physical activity. Hanibuchi *et al*, in a Japan-based study, operationalised neighbourhoods of older adults (n = 9,414) as network buffers with radii of 250 to 1000 metres[104]. Their study showed the presence of greenspace, regardless of neighbourhood size, to be positively associated with frequency of recreational but not transport-related physical activity. However, a study of middle aged adults (n = 4950) conducted in Norwich, United Kingdom found no significant association between access to neighbourhood greenspace - in terms of distance, size and quality – within the neighbourhood with self-reported hours per week recreational activity in 4950 middle-aged adults living in Norwich. Whilst the specification of greenspace exposures and physical activity outcomes in these studies were not identical and may account for the contrasting findings, it may be that the recreational physical activity of older adults is more sensitive to neighbourhood greenspace than younger adults who are more likely to be economically active and spend less time within their residential neighbourhood.

Reviewing the literature Leslie *et al* noted that whilst studies tended to indicate perceived neighbourhood greenness was positively associated with physical activity, there was less consistent evidence for positive associations between objective neighbourhood greenness and physical activity [105]. Leslie *et al* compared perceived neighbourhood greenness scores, derived from questionnaire responses from 94 individuals, to objectively-measured greenness covering the neighbourhood (a 400 metre buffer around each residential parcel) as measured by satellite imagery and found only a weak correlation. They hypothesise that the ground-level evaluation of greenness perceived by individuals may constitute a more salient measure of the environment than objective aerially measured greenness with respect to effects on physical activity. They argue that how residents see the neighbourhood will have a bigger impact on how they use it, explaining the trend for studies to report positive associations between physical activity and perceived but not objectively-measured greenness.

Barton and Pretty sought to establish the optimal dose of physical activity in natural environments – green exercise – for mental health through a meta-analysis of studies involving a total of 1252 participants [106]. Their results indicated that green exercise improved self-esteem and mood, particularly with those with a starting health status of mental ill health, but with diminishing returns as duration and intensity increased.

2.8.2.1.2 Residential density

Residential density can be defined as the number of residential dwellings per unit area of land. It has been hypothesised that high residential density creates a critical mass of people at which seeing people walk encourages others to walk and that it increases walking for transport by increasing pedestrian and motorised traffic congestion such that walking or taking public transport become more attractive ways to travel than driving [107][93]. Indeed, in North America rates of car ownership have been found to be higher in areas of higher residential density [108]. Residential density data for administrative areas is readily available and may be a proxy for other variables such as neighbourhood income[92], for which privacy issues often prevents access and for which associations with physical activity are also expected. Thus, residential density is one of the most common measures investigated in relation to physical activity [92], although it is argued that variation in definitions compromises the comparability of studies [109].

Glazier *et al* investigated residential density as a single factor, and in combination with other physical environment attributes, in relation to active travel in the population of the Canadian city of Toronto [110]. Using census data for the city's 2.6 million residents along with transportation and national health surveys, they found that compared with those living in areas that were less walkable as indicated by either lower residential density or lower availability of destinations, individuals in more walkable areas were significantly more likely to walk for transport – the average number of daily walking and cycling trips per person was more than twice as high – and were less likely to own a car. Upon modelling residential density and availability of destinations together as indicators of walkability, these associations were strengthened suggesting an additive effect. The authors concluded that residential density and the availability of walkable destinations should be used by urban planners as measures of walkability. However, Glazier *et al* noted that a lack of data precluded adjustment for individual-level factors such as age, sex, ethnicity or socioeconomic status. Had they adjusted for socioeconomic status, the significant associations may have been reduced or eliminated because their results showed that individuals in the least walkable areas had a higher average household income. It may be that higher income as opposed to physical environment factors accounted for higher car ownership rates in lower walkability areas.

Data from the Twin Cities Walking Study [111], conducted in Minnesota, United States, was used to examine residential density in relation to walking for various purposes by Forsyth *et al* [107]. The design of the study allowed the researchers to compare associations between various measures of residential density, taken at different scales, and walking measured both objectively and self-reported. Their main findings were that, irrespective of how neighbourhood was operationalised and how residential density was specified, there were no significant associations between residential density and total walking but that density was associated with specific walking purpose: higher density areas had more transport-related walking whilst lower density ones were associated with more recreational walking. Thus, there was a "zero-sum" effect of residential density on overall walking, highlighting the importance of simultaneous consideration of multiple factors of the built environment. In contrast to Forsyth *et al*'s study, a study conducted in Japan found higher residential density to be positively associated with recreation physical activity but not transport-related physical activity among 9141 older adults [104]. The difference in findings may be attributable to differences between the studies in terms of neighbourhood operalisation as well as contextual differences between the countries studied.

2.8.2.1.3 Street connectivity

Street connectivity is hypothesised to be a determinant of walking or cycling for transport by affecting the directness of travel and therefore its efficiency and attractiveness, with a higher number of routes where street connectivity is greater providing more interest and, potentially, safety [112]. There is indeed evidence to support the notion that street connectivity promotes physical activity [113][41][112]. However, it has been proposed that social stratification defines neighbourhoods with, for example, poor people living in certain areas and rich people in others [114] which affects the interpretation of physical environment neighbourhood associations with health. This argument led Oakes *et al* to employ a sampling strategy in their cross-sectional study of the relationship between street connectivity – operationalised as block size, with smaller block size indicative of greater street connectivity – and physical activity that did not rely heavily on statistical adjustment for confounding [115]. Participants were drawn from residential areas selected from an environmentally diverse but demographically homogenous part of a North American city. After adjustment for socioeconomic factors they found street connectivity was not significantly related to distance walked per day. The authors attributed this finding, inconsistent with previous research, to their study's sampling design which mitigated residual confounding by socioeconomic status. Thus, whilst relationships between environmental attributes such as street connectivity with physical activity may be intuitive and plausible, their importance relative to individual factors may be overestimated by studies which fail to account adequately for confounding variables.

2.8.2.1.4 Land use mix

It has been suggested that the last fifty years of the implementation of public health-driven policies to separate residential areas from industrial areas, and thereby people from harmful pollution, has inadvertently fuelled the current obesity epidemic [116] by reducing land use diversity. Diversity in land use – land use mix (LUM) – is posited to enable better access to services and employment, and to induce shorter within-neighbourhood travel by foot and by bicycle as a range of destinations are located near residences [117]. Also, in areas where different destinations such as restaurants and workplaces are co-located walking is likely more time-efficient than using public or private motorised transport to access them. Thus, LUM encourages physical activity, reducing risk of obesity and associated disease. LUM is a measure of the physical environment that gives the heterogeneity of land uses, typically including those that are residential, commercial, institutional and recreational, in geographically-defined areas [118]. A variety of LUM measures exist, the deployment of which is recommended to be based on the number of land uses of interest and the scale of land use variation [119]. Batty *et al* note that variation in LUM at a very fine scale may be entirely eliminated, so there is no multi-functionality whatsoever [120]. Entropy measures of diversity, wherein higher values more varied land use environments [121], are widely used in travel studies. Entropy is a concept which has been described as a measure of "mixed-up-ness" [122].

Whilst mixed land use reflects the availability of destinations to which residents can walk or cycle, and many studies have found the expected relationships between LUM and physical activity, some studies have not [123][124]. Duncan *et al* investigated whether the scale at which LUM was measured and the land uses included, reflecting the specificity of the hypothesised environment behaviour, accounted for differences in associations in the walking for transport among 2,506 adults in 154 Census Collection Districts (CCDs) in Adelaide, Australia. They found that changing either the LUM score to account for the relative sizes of CCDs, or refining the LUM score to exclude land uses of low theoretical relevance to walking for transport, gave measures that had significant positive associations with the frequency of walking for transport, after accounting for sociodemographic factors, whereas the original did not. Duncan *et al* also found that study participants' perceived proximity to destinations were more strongly positively correlated with the LUM scores that accounted for the relative sizes of CCDs, and for land use relevance to walking for transport, than they were with the original ones. Thus, in the context of the Australian city studied, it appears that the physical activity outcomes are sensitive to LUM only if account is made of geographic scale and the relevance of land uses. However, the authors of the study caution that their approach may not yield better LUM measures for cities where large differences in land use distribution or in size of administrative areas exist. They suggest that rather than simply eliminating less relevant land uses, researchers should weight land uses according to their relative relevance to walking to yield LUM measures of better quality as per Frank *et al* [125].

2.8.2.1.5 Walkability

Walkability is the extent to which a place lends itself to walking and cycling as physically active forms of transport and recreation. As discussed there is much evidence that factors such as greater street connectivity and higher residential density constitute physical environmental attributes that support physical activity such as walking and cycling. Historically many researchers quantified a single attribute, such as residential density as a proxy for walkability, arguing that factors positively associated with walking and cycling tend to co-exist but there is a growing consensus that such factors should not be measured in isolation because often they do not occur together [45]. Individually these attributes may be insufficient to promote physical activity. As Frank and Engelke note, for example, individuals are likely to make use of greater street connectivity only if they have a range of places with complementary uses to go, where there is greater LUM [126]. Thus a true indicator of a neighbourhood's walkability may be a measure that incorporates a range of factors, with a highly walkability "score" given to a neighbourhood that was simultaneously highly connected and residentially dense whilst also having a high degree of LUM. Walkability indices are designed to capture the multiple attributes of a place, rather than individual features, for which there is some empirical evidence for a positive association with walking or cycling in order to reflect the multiple dimensions of walkability. Over the past decade walkability indices have been constructed and tested, at various spatial scales and in different settings, for a wide range of populations with researchers tailoring components and their quantification, and units of analysis to fit their hypotheses [127]. Because each is designed specifically with the research population and setting in mind, there is no "off-the-shelf" walkability index that is universally applicable. However, three core dimensions are salient across populations and form the basis of a majority of walkability indices. They are net residential density (the number of residential units per unit area of residential land), street connectivity (the number of junctions per unit area) and LUM (the evenness of distribution of land area specific to a range of uses such as residential, retail, office and recreation). It must be noted that although distance between places, as discussed in this review, is an individual indicator of accessibility for an individual, it is not an areal measurement so cannot be applied to an area to indicate relative walkability of areas.

One of the first walkability indices was developed by Frank *et al* [113]. They found objectively measured LUM, residential density and street connectivity in the neighbourhood, defined as a 1-km road network-based buffer around each participant's place of residence, to be positively associated with objectively measured minutes of physical activity per day. Given that these measures of urban form often co-vary across space – higher residential density tends to occur where street networks are more connected and land uses more mixed [128][117] – the authors integrated the three factors into a single index, avoiding the difficulty of identifying the separate influence of each factor. They found that the combined walkability index explained more of the variance in physical activity, even after adjustment for sociodemographic variables than any of constituent physical environment factors alone. Of the three core dimensions of walkability, LUM is the most complex in terms of possible compositions and computations, and therefore, that in which walkability indices tend to differ. Recently, a longitudinal study examined LUM in relation to the efficacy of a walkability index [129]. It showed how variation in the specific land use types included in the LUM dimension affected the ability of the model to explain walking behaviour in a population (n=1798) in the Perth metropolitan area, Western Australia. Applied within a 1.6km street network buffer of each participant's residential location, the walkability index better explained recreational activity when public open space was included in the model, for example, highlighting the importance of fitting the model to explain the behaviour. Validation of walkability indices by street level audits, the gold standard in environmental assessment, is rarely done [130]. Nevertheless, many reviews have concluded there is evidence that composite measures of walkability are positively associated with physical activity [131][112][132], and with walking [133][134], supporting the use of walkability indices as indicators of health-related behaviours.

In conclusion, there is much evidence that accessibility measures of physical environments are associated with physical activity outcomes but those that combine multiple aspects, such as walkability indices, may constitute better indicators.

2.8.2.2 Physical environments measured in terms of atmospheric conditions

Atmospheric conditions were components of the objective environmental quality exposure variable investigated in relation to physical activity outcomes in Chapter 5, hence their inclusion in this review. The atmospheric conditions reviewed here in relation to physical activity and health outcomes are air quality, and weather (a measure of atmospheric conditions over a short period of time) and climate (a measure of atmospheric behaviour over a relatively long period of time). Unlike the predominantly built physical environment factors discussed this far, atmospheric conditions lack discrete spatial boundaries. The basis of the review of studies examining atmospheric conditions as physical environment exposures is the hypothesis that higher levels of pollution, harsh climates and inclement weather are deterrents to outdoor physical activity.

2.8.2.2.1 Air quality

Disruption of the atmosphere's natural gaseous system through the introduction of chemical, biological or particulate matter has the potential to adversely affect human health. Many of the neighbourhood physical environment attributes theoretically or evidentially linked with greater levels of physical activity and thereby better health outcomes support higher densities of activity, including vehicular traffic. Greater street connectivity, for example, may be attractive for pedestrians but as it represents higher road density it is accommodating of cars as well. Given that traffic is a leading source of air pollution [135] and that high levels of air pollution are a major risk to health, and a leading environmental cause of cancer deaths [136], air quality merits consideration alongside the "fixed" and built physical environments associated with health outcomes.

In urban settings there is substantive evidence for positive associations between concentrations of air pollutants, particularly particulate matter (PM_{10}) and ozone (O_3), carbon monoxide (CO), sulphur dioxide (SO_2) and nitrogen dioxide (NO_2), and respiratory and cardiovascular disease, and total mortality[137][138][139][140][141]. In a recent review two mechanisms by which air pollutants could increase risk of cardiovascular disease proposed were causing inflammation and deposition of fatty plaques on arteries, and causing subtle changes in the heart rhythm [142]. Although the review mentioned lack of physical activity as a risk factor for cardiovascular disease, it did not suggest that physical activity may mediate the relationship between air pollution and cardiovascular disease. Theoretically, air pollution would only cause cardiovascular disease through lack of physical activity if it were perceptible and, thereby, a deterrent to being active outdoors. Given that air pollution tends to be imperceptible at the level associated with increased risk of cardiovascular disease – non-particulate pollutants are invisible and odourless – it is likely that physical activity does not lie on the causal pathway. However, with evidence that perceived and objective air pollution measures are positively correlated [143], a role for physical activity as a mediator to some health outcomes is plausible.

Hypothesising that walkability would be differentially associated with air quality dependent on neighbourhood income, Marshall *et al* investigated the interaction between neighbourhood walkability and air pollution exposure as indicated by levels of NO_2 and O_3 in Vancouver, Canada [144]. NO_2 is a surrogate for all traffic-related combustion products [145] and commonly used indicator of traffic volume whereas O_3 is a secondary pollutant, non-emitted pollutant which occurs downwind of high density areas. They found that, compared to higher income areas, lower income areas tended to have higher walkability together with higher NO_2 concentrations but with lower O_3 concentrations. Neighbourhoods that had high walkability but low concentrations of both pollutants were mostly higher income areas. Their results suggest that whilst walkability may have health benefits through promotion of physical activity, it may coexist with factors which have health costs, particularly in lower income areas. Thus, it may be that the positive health outcomes found to be associated with high accessibility environments are reduced by the negative ones associated with poor air quality.

2.8.2.2.2 Climate and weather

The human thermoregulatory system enables the body to cope with thermal stress within limits but outside this comfortable range – at both low and high temperatures – thermoregulation may be impaired or fail. Reviewed evidence suggests that mortality, particularly from cardiovascular and respiratory disease, increases substantially during heat waves [146] and cold snaps [147] but more subtle associations exist with behaviour. It is intuitive that daylight, extreme temperatures and precipitation levels influence outdoor physical activity such as neighbourhood walking and, for example, a study of adults found that administrative areas where the rate of meeting the recommendations for physical activity was in the top quartile had the highest percentage of days with dry moderate conditions whereas those where the rate was in the bottom quartile had the highest percentage of days with moist tropical conditions [148]. However, Tucker and Gilliland argue that precipitation and other potential climate and weather-related determinants of physical activity have received inadequate research attention [149]. Their review of studies of the effect of seasonality identified poor or extreme weather as a barrier to participation in physical activity among various populations and concluded that failure to account for seasonality in studies examining other physical environment attributes as determinants compromised the validity of their conclusions. However, Humpel *et al* found that weather tended to be pooled with other factors as a an environmental influence, potentially masking any strong associations with physical activity [82]. Therefore, whilst researchers may account for weather they need to include it as an individual factor when modelling the environmental effects. Global warming is sparking growing interest in the relationships between climate and physical activity, and a recent scientific article highlighted the risks of an increase in average ambient temperature, including the possibility that there will be increased sedentary behaviour as people are forced to spend more time indoors [150].

2.8.3 Europe-specific studies

Most studies examining associations between physical environments and physical activity and putative physical activity-mediated health outcomes have been undertaken in cities of North America and Australia rather than European cities. However, European cities tend to be more compact [151], making average trip lengths shorter [152]. Thus, the nature of physical environment associations with physical activity and health outcomes are likely to differ and there is some evidence that this is the case with respect to particular physical environment attributes.

Van Holle *et al*'s review summarizing European studies found results to be generally in accordance with those of North America and Australian studies [153]. They found convincing evidence that walkability was positively associated with total physical activity, and also with walking and cycling for transport. This evidence was largely consistent with non-Europe specific reviews which also concluded there was a positive association between walkability and physical activity but not specifically that which was recreational [133][154][112]. Also, Van Holle *et al*'s conclusion that in European settings access to services was positively associated with active travel echoed those in non-Europe-specific reviews of positive associations with total walking [133], active travel [154] and walking for transport [155]. Given the consistency of findings it is suggested that walkability and access to services are pertinent to physical activity in the domain of transport but not recreation, possibly because these physical environment measures capture convenience. This attribute of physical environments is likely to relate more to getting from A to B than to jogging in a park, for example. Van Holle *et al* posit that whilst most measures of physical environment are taken within the residential neighbourhood, however that is defined, recreational physical activity may be undertaken outside that environment. This would explain the lack of evidence their review identified for positive associations between access to recreational facilities and recreational physical activity. However, positive associations between these variables were found in reviews of non-European-specific studies [82][156][17][154], suggesting that in non-European settings individuals' recreational physical activity may be more sensitive to residential neighbourhoods. This is plausible because, given the sprawl that is characteristic of many non-European cities, the administrative units (as which neighbourhoods are commonly operationalised in research) may be substantially larger, making the likelihood that a given individual will undertake recreational physical activity, and be sensitive to any effects of the provision of recreational facilities within it higher.

2.9 Direct relationships between physical environments and mental health, including depression

To date few studies have examined direct relationships between physical environments and mental health outcomes such as depression. However, more studies have focused on the relationships between physical environments and quality of life, subjective wellbeing and self-rated health. Whilst these outcomes are not explicit indicators of mental health status, there is evidence that such measures may constitute valid means of identifying people at increased risk of depression [157][158]. Therefore, studies which specify quality of life, subjective wellbeing and self-rated health as outcomes are included in this section of the literature review.

The idea that the physical environment may affect physical activity is readily understandable if not intuitive, making its investigation in relation to physical activity a logical step in gaining an insight into factors that may be important determinants. However, the rationale for examining the physical environment in relation to mental health is less obvious. As discussed previously, a mediatory role for physical activity in the relationship between physical activity-supportive neighbourhood attributes and health outcomes such as obesity is likely, so it is plausible that one exists for the relationship between physical activity-supportive neighbourhood attributes and mental health. Therefore, studying the associations between the physical environment and mental health outcomes, such as depression, and understanding the extent to which physical activity mediates them is important. Indeed, associations between levels of depression and neighbourhood traffic, public transportation, green space, and services have been reported [159], and it is possible that physical activity mediates these relationships. However, a recent study found an association between the level of physical environmental support for walking within the neighbourhood and depressive symptoms in older men that was independent of self-reported walking frequency [160]. This suggests that more walkable neighbourhoods may confer psychosocial benefits independent of physical ones. Social connections between neighbours are protective against depression [159], whilst sense of community, a related social concept, is associated with the extent to which the neighbourhood supports walking, independent of demographic factors [161], lending weight to the theory that certain aspects of the physical environment engender a greater sense of community affording the neighbourhood social connections that protect against depression.

A recent government-commissioned report which reviewed literature about the physical environment and its impacts on mental health identified factors relating to sensory stimulation as particularly important [162] with, for example, the layout of the built environment (associated with way-finding), a determinant of perception of safety. Access to natural environments, however, was highlighted as an especially important factor associated with mental health. Whilst research interest in the role of natural environments in human health has grown in recent years alongside the broader sustainable development agenda, robust evidence for beneficial effects is lacking. However belief in, and theories of, the benefit of natural space for human health are not new, emerging alongside urbanisation in the mid-nineteenth century. Disease and crime rates were often high in urban areas and it was also apparent that towns were devoid of the natural space afforded to people living in rural areas. The prescriptive redesign of neighbourhoods to incorporate greenspace and optimize the wellbeing of inhabitants was advocated [163]. Whilst appreciating the importance of nature and greenspace, others argued that propagation of more complexity and chaos in cities was necessary to reflect and complement that inherent within individuals and their communities [164].

The biophilia hypothesis, which posits that there is an instinctive bond between humans and the natural ecosystem of which we are a constituent [165], often forms the basis of expectations that nature is good for human wellbeing. Stemming from this hypothesis, another theory proposes that nature provides the necessary environmental stimuli to allow restoration from fatigue from attention to tasks [166]. It is suggested, for example, prolonged computer data entry that leads to attention fatigue (which may reduce mental wellbeing) can be overcome by a walk in a park. Support for this theory is limited but a recent exploratory study that employed mobile electroencephalography devices to monitor brain wave activity found evidence suggestive of a reduction in stress levels when participants (n=12) walked from non-green areas to green areas within an urban environment [167]. Another recent, larger study conducted in Wisconsin, United States also lends support for the restorative benefits of exposure to natural environments [168]. Beyer *et al* examined the relationship between environmental green space and depression outcomes among 2,479 adults. They found higher levels of greenspace within the neighbourhood, operationalised as census blocks, was associated with better mental health, independent of individual level and neighbourhood level socioeconomic factors. Overall, evidence points to a positive relationship between natural environments and mental health and wellbeing, although definitions of the exposure and outcome vary [169][170][171][172]. Fully exposing research study participants to natural environments without requiring them to be physically active to at least some degree is challenging. Therefore, it is difficult to conduct experiments to establish whether natural environments act independently of physical activity. As discussed previously, an independent effect is plausible but there is evidence to suggest that natural environments and physical activity have a synergistic effect on quality of life. Thompson-Coon *et al* systematically reviewed nine randomized and non-randomized controlled trials comparing mental wellbeing effects of participation in physical activity in natural environments with those of participation in physical activity indoors [173]. They found compared with exercising indoors, exercising in natural environments was associated with greater decreases in depression.

Vegetation can affect health-relevant air quality both beneficially and adversely through meteorological effects, and by sequestering and emitting pollutants and allergens [174] and a review of the benefits of absorption of the air pollutants SO_2 and PM_{10} by trees concluded that this vegetation could extend life expectancy and reduce hospital admissions [175]. Given the relationships between vegetation, air quality and health, together with research interest in the effects of natural environments on mental health, it is perhaps unsurprising that many studies have investigated air quality in relation to quality of life and other such measures. Studies indicate that objectively measured air pollution is negatively associated with quality of life [176], subjective wellbeing [177] and life satisfaction [178][24]. It is apparent that the designs of many of these studies do not deal adequately with the methodological challenges inherent in the objective measurement of air pollutants which, as noted previously, lack spatial boundaries. However, their findings are supported by studies using subjective measures: significant negative associations between perception of air pollution and subjective wellbeing [179] and between perception of air pollution and life satisfaction [24] have been found.

2.10 Summary

The review of evidence on associations between physical environments and physical activity has revealed a vast body of research but studies are spread thinly over the plethora of physical environment attributes examined, rendering the strength of evidence for specific associations weak. Associations between attributes of the physical environment with depression have been the focus of few studies although many have investigated associations with related outcomes such as subjective wellbeing, and there is some evidence suggestive of both physical activity-mediated and independent relationships. It is perhaps because interest in the role of the physical environment as a determinant of physical activity and health outcomes is relatively new, emerging over the last thirty years or so, that many studies have been exploratory and that exposure and outcome measures have not been standardised. There is much variation in the operationalisation of neighbourhoods between studies, many of which use large administrative units, seemingly for convenience rather than a hypothesis that such areas represent neighbourhoods where effects of exposure are likely. Inconsistencies in exposure and outcome variable specification and in neighbourhood operalisation limit comparability of study findings, and may account for the failure of many studies to detect significant associations. Additionally, it must be noted that many studies reviewed here have employed cross-sectional designs and are thus subject to the structural confounding issue of segregation wherein, for example, depressed people are likely to be socially drawn or pushed into neighbourhoods of worse quality. Potential reverse causality precludes causal inferences to be made from cross-sectional associations between neighbourhood physical environment and “outcomes”. Also, the vast majority of studies to date have been conducted in North America and Australia, preventing generalizability of evidence. Nevertheless, evidence is accruing for associations between particular physical environment attributes. It is apparent that access to greenspace is important in relation to physical activity and to depression outcomes. Walkability, a composite indicator of accessibility attributes, also emerges as a factor that seems to promote specific domains of physical activity, especially walking for transport. This review also concludes that both perceived and objective measures of the physical environment merit attention. Perceived measures may be particularly useful for gauging overall quality, an attribute difficult to measure objectively. Whilst the reviewed evidence shows many studies have used large representative population samples, it is clear that individual level factors such as age and socioeconomic status moderate relationships and that adjustment for them is not always made due to data limitations.

2.10.1 Conceptual framework

In light of the evidenced reviewed here, a conceptual framework was constructed to guide this study, illustrated in Figure 2.1. The neighbourhood physical environment affects physical activity with, for example, individuals living in neighbourhoods where there is a higher proportion of greenspace being more likely to meet the WHO recommended physical activity level. It also affects mental health through physical activity and directly. However, individual level factors exert influence on physical activity and on health with, for example, younger people more likely to meet the WHO recommended physical activity level and employed people to be less likely to have GHQ-12 depression. In addition, individual factors predispose individuals to certain physical environments. Those without car access, for example, are more likely to live in more walkable areas. Physical activity has a reciprocal relationship with physical and mental health, as does mental health with physical health.

2.10.2 Research gaps

Many gaps in the research exist, some of which are addressed in this study. Firstly, there is a need to examine the physical environment specifically in relation to depression. The influence of physical environment on physical health and health-related behaviours has received considerable research attention relative to its influence on mental health. However, the conceptual framework (Figure 2.1) constructed from the reviewed literature shows the potential importance of the physical environment in relation to depression. Empirical evidence for relationships between the physical environment and wellbeing indicators such as quality of life and self-rated health is building but these outcomes do not constitute explicit indicators of mental health. With depression a growing burden of disease, particularly in aging populations, it is imperative that relationships of this specific outcome are investigated.

Given the preponderance of studies in North America and Canada, there is a need to examine the relationship of attributes of the physical environment – especially greenspace, overall perceptions and accessibility – with depression, and also with physical activity, in case studies in Europe where environmental context is very different. In particular, there is a need to construct walkability indices, which vary in the land uses incorporated, in the context of a city in the United Kingdom, and to test for associations with physical activity and depression. Innovatively, this study harnesses the power of rich geographical databases and GIS software, enabling the street connectivity dimension of the walkability index to take account of a surface transport network that includes footpaths, relevant to walking behaviour outcomes. It also tests the validity of the walkability indices through visual inspection, an auditing procedure rarely done in relation to objectively measured walkability. Finally, neighbourhood can be and has been operationalised in a multitude of ways, accounting for inconsistencies in associations found between physical environment attributes and physical activity and health outcomes so there is a need to determine which operationalisation of neighbourhoods are most appropriate for particular physical environment exposures.

2.10.3 Research questions

In order to address the research gaps identified by this review of the literature, this thesis poses the following questions:

1. Is physical activity independently associated with depression outcomes in the Whitehall II phase 7 (WIIP7) study sample? (Study 1)

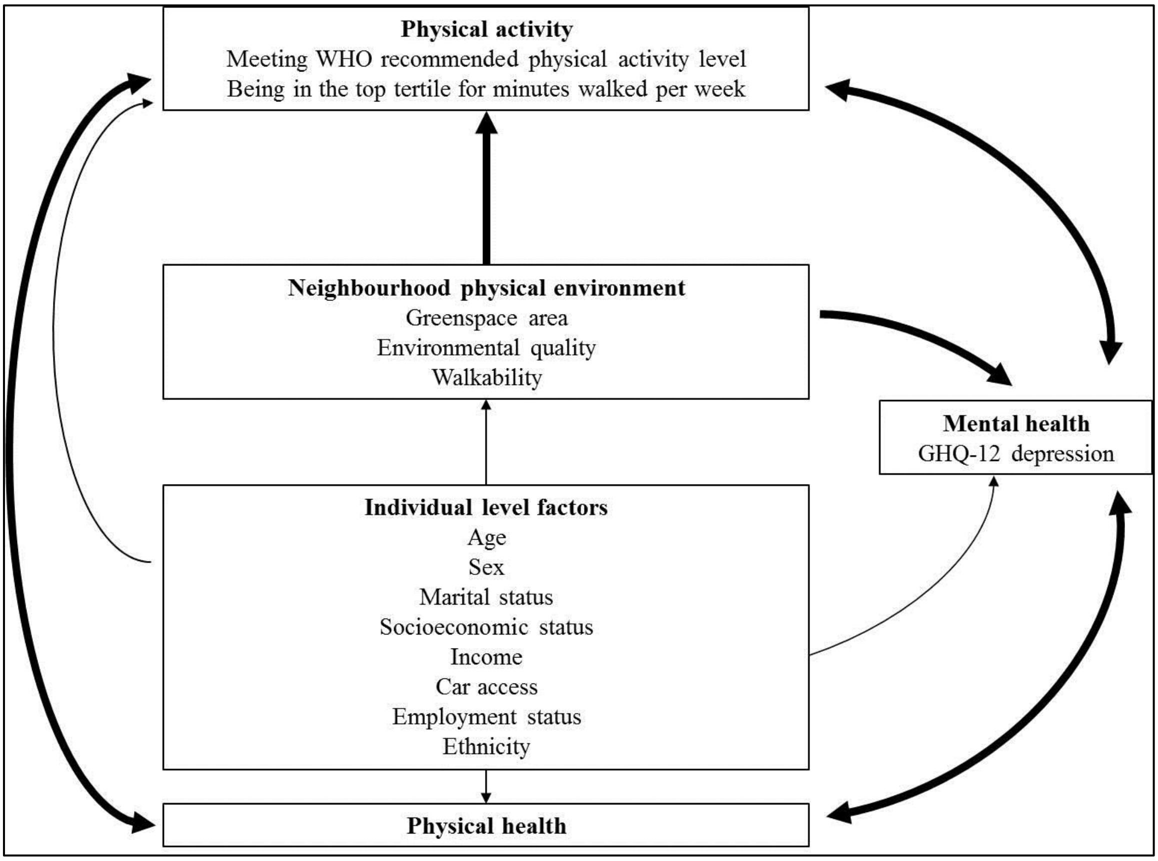
2. Is proportion of neighbourhood green space and domestic garden independently associated with physical activity in WIIP7, and with depression outcomes in WIIP7 and Health Survey for England (HSE) 2008 study samples, and how does the operationalisation of neighbourhood affect associations (Study 2)?

3. Are an objective composite indicator and a subjective measure of neighbourhood environmental quality independently associated with physical activity in the WIIP7 sample, and with depression outcomes in WIIP7 and HSE2008 study samples, and how does the operationalisation of neighbourhood affect associations? (Study 3)

4. What is the spatial variation in walkability across London, and how do the specification of the walkability model and the operationalisation of neighbourhood affect this variation? Do walkability scores attributed to output areas match subjectively assigned scores based on visual inspection? (Study 4)

5. Is neighbourhood walkability independently associated with physical activity in WIIP7, and with depression outcomes in WIIP7 and HSE2008 study samples, and how do the specification of the walkability model and the operationalisation of neighbourhood affect associations? (Study 5)

Figure 2.1 The conceptual framework for the present study.



3 Data and methods

3.1 Introduction

This chapter firstly provides an overview of the datasets from which the study samples were drawn for the present research, detailing what and how data was collected, and contrasting the two study samples with respect to their characteristics and geographical distributions. As a focus of this thesis was the effect of neighbourhood delineation as different levels of administrative geographies on associations between neighbourhood exposures and outcomes, the chapter then outlines the spatial characteristics of administrative units across the United Kingdom and highlights important inter-regional differences. Descriptive statistics for the environmental measures used to construct neighbourhood exposure variables are then presented for administrative levels across the whole of the United Kingdom. These statistics, together with the spatial characteristics of administrative units, are used to justify the specification of the exposure variables. Also, the descriptive statistics for each specified exposure variable are presented for the study samples. Next, this chapter provides descriptive statistics for the data collected in the study samples pertaining to the outcomes of interest, justifies the specification of the outcome variables and presents descriptive statistics for the outcomes in the study samples. Finally, this chapter outlines the design and statistical methods employed in this thesis.

In this chapter, Stata 12.1 statistical software [180] was used for statistical analyses as an exploration of the data informing decisions pertaining to the derivation of outcome and exposure variables and appropriate statistical testing. This involved basic descriptive analyses including calculation of median, means and standard deviations for continuous variables and proportions for categorical variables. Statistical software was also used to construct the study samples and the outcome variables. Neighbourhood exposure variables were derived using statistical software in conjunction with ArcMap 10.1 geographical software [181].

3.2 Study samples

Here, an overview of the datasets from which the study samples were drawn for the present research is given. The study samples were from two sources: the 2008 Health Survey for England (conducted between January 2008 and April 2009), and the seventh phase (2003-2004) of the Whitehall II study cohort.

3.2.1 Whitehall II

3.2.1.1 Whitehall II Background

The Whitehall II study (WII) is a prospective cohort study that monitors the health and social circumstances of a sample of civil servants to understand the social distribution of their health conditions. In 1984 all people between the ages of thirty-five and fifty-five employed in the London offices of the British Civil Service were invited to participate in the study. Seventy three per cent agreed to participate, giving a study sample size of 10,308[182]. Until the 1980s most cohort studies investigating physical health outcomes tended to focus on biological risk factors, excluding potentially interacting social factors such as diet and lifestyle from the causal framework. The original aim of WII was to examine the effect of the work environment on chronic disease, the moderating role of social support and the interaction between these psychosocial factors and known biological and behavioural risk factors. In addressing these novel research questions, WII sought to understand how social factors influence biological pathways. The original outcomes specified included coronary disease, respiratory illness, 'neurotic disorder', and sickness absence. They have since been widened to include health-related quality of life with the aim of examining inequalities in health in the aging population. At five-yearly intervals, the whole WII cohort is invited to a research clinic at which the study's nurses conduct physical examinations such as the height, weight and blood pressures of participants, and take various biological specimens such as blood and urine for analysis of markers of disease, such as lipids and carbohydrates. Between these clinic phases a postal questionnaire is mailed to participants to collect non-biological data. WII also collects geographic residential data to enable maintenance of contact with participants. Whilst this data was originally collected for purely administrative purposes, it is also useful for examining relationships between health conditions and environmental factors which have an inherent spatial dimension.

3.2.1.2 Defining the Whitehall II study sample

The study sample was defined as all individuals participating in the seventh phase of Whitehall II for whom a valid United Kingdom residential postcode was available, henceforth referred to as WIIP7. This phase was selected because it was relatively recent (the most recent being Phase 11) and the postcode data was more accurate than for other phases. Details of the study sample size and response rate from recruitment to the phase used in this project are given in Table 3.1. Of the 6967 participants of the seventh phase, aged between fifty and seventy-four, a valid postcode was available for 6885 (99%). A subset of WIIP7 was created, LWIIP7, which comprised only the 3020 individuals with a residential postcode in London (44% of WIIP7), for London-specific exposure variable analysis. A summary of the non-biological data collected by self-administered questionnaire for WIIP7 is provided in Table 3.2. Individual-level characteristics of WIIP7 and LWIIP7 are charted in Table 3.3.

3.2.2 Health Survey for England

3.2.2.1 Health Survey for England Background

The Health Survey for England (HSE) is an annual government-commissioned survey to monitor trends in the nation's health. Started in 1991, the survey is commissioned by the Health and Social Care Information Centre and carried out by the Joint Health Surveys Unit comprising the Health and Social Survey Research group at the Department of Epidemiology at University College London and NatCen Social Research, an independent social research agency with registered charity status. The survey is designed to be representative of all adults aged 16 and over living in private households in England. Individuals identified through by stratified two-stage random sampling from the Postcode Address File are invited to participate, as detailed by Craig *et al*[183]. Although institutionalized individuals who are generally older and less healthy are excluded making extrapolation of findings to the whole nation potentially erroneous, the sampling technique is designed to maximize the sample's representation of people living in private households in England. Each year there is a focus on a specific theme which is then repeated after an appropriate number years to enable monitoring of particular population groups, diseases or conditions. The survey comprises a researcher-delivered questionnaire-based interview with core questions and questions specific to the particular year's theme, followed by a nurse visit.

3.2.2.2 Defining the Health Survey for England study sample

The study sample for the present research comprised all individuals aged sixteen and over who participated in the HSE conducted between January 2008 and April 2009 and for whom no items were missing on the General Health Questionnaire (GHQ) section, and thus for whom a GHQ-12 score could be computed, henceforth referred to as HSE2008. In 2008, there was a household response rate of 64 per cent with 15,102 adults interviewed [184], 14,221 (94%) of whom had complete GHQ-12 data. LA codes were provided for HSE2008 enabling area level indicators such as deprivation and land use to be attached to individuals. Attachment of LA codes also enabled identification of the Government Office Region (GOR) of residence, permitting analysis specifically for London, the predominate region of residence of participants in the Whitehall II study sample. A subset of HSE2008 was created, LHSE2008, which comprised only the 1540 individuals with LA codes corresponding to the London region (12% of HSE2008), for London-specific exposure variable analysis. In 2008 the core questions of HSE covered a range of self-reported health measures and health-related behaviours, individual socio-demographic and household characteristics, and during the nurse visit a range of biological measurements were taken. Data collected is detailed in Table 3.4, and Table 3.5 charts individual-level characteristics of HSE2008 and LHSE2008.

3.2.2.3 Comparison of WII and HSE study samples

HSE serves to monitor health trends in the English population and WII, principally, to show the role of social factors in shaping health outcomes. Thus, whilst both HSE and WII are relatively large surveys valuably providing statistical power, their aims are distinct and, consequently, their designs very different. As a result there are important differences in the types of data collected and the distribution of the characteristics of the study samples, rendering physical activity and mental health outcomes incomparable between the two samples. The over representation of men and of older individuals in WIIP7 (Table 3.3), and the specific employment sector from which participants were drawn for this study was reflected in disparities in the frequencies of other sociodemographic variables with the nationally representative HSE2008 (Table 3.5). Frequencies of participants in higher income brackets, occupational social classes and levels of education were higher in WIIP7 than HSE2008. In addition, higher proportions of WIIP7 were married and white, although approximately the same proportion in each sample was not working due to being out of work or retired.

The entire WIIP7 sample is, or has been, employed by the Civil Service so many of the socioeconomic indicators are particular to this specific part of the public sector. For example, occupational social classes are limited to classification as the lowest, “Clerical”, “Professional/Executive” and, the highest “Administrative”, all of which are office-based. Conversely, within HSE2008 the range is broader and includes individuals in non-office based and, therefore, more physically active jobs. However, it has been argued that the job grading system unique to the Civil Service and by which WII participants are categorised, produces more homogeneity within groups than those of occupational social classification systems, such as the National Statistics Socio-economic Classification, used in many surveys including the HSE [185]. More homogenous social groups are produced in the WII Study because at recruitment into the study all participants were working in one location and as sedentary office-based employees. This shared environment meant that within groups there would likely be less difference in terms of material circumstances than would be seen within social groups drawn from diverse locations and different types of jobs, albeit classified as the same grade. The implication, and strength, of this is that where socioeconomic differences in health and health-related outcomes exist they are potentially more accurately detected and pronounced within the WII study.

As a longitudinal survey, WII’s cohort is aging and inevitably subject to attrition. Despite diligence in retaining contact and steps taken to encourage continued participation such as free health checks and regular newsletters, 68 per cent of Phase 1 responders participated in Phase 7 almost twenty years later as shown in Table 3.1; individuals at Phase 7 would likely be underrepresented by those of poorer health for whom participation would be prohibitively onerous. Thus, the Phase 7 cohort is likely to be slightly over-represented by healthy individuals relative to the cohort at recruitment. The “staying-power” of Phase 7 responders must be considered when evaluating associations between physical activity and mental health outcomes and environmental exposure variables. Attrition does not occur in HSE due to its cross-sectional design.

Finally, there were marked differences in the geographical spread between the two study samples. WIIP7 comprises individuals who were all London-based workers in 1985. Consequently at Phase 7, the highest proportion resided in London (44%) with the rest living in other parts of the United Kingdom, the vast majority in England (97%) as shown in Map 3.1. It is apparent that whilst London did not remain home to all participants, those who did leave the capital tended to remain nearby, 30% residing in the South East. All participants in HSE2008 resided in England and individuals were nationally and regionally representative in their geographical distribution across England, with approximately one quarter of the proportion represented by WIIP7 in London being represented in HSE2008 (12%). Within London, HSE2008 participants were relatively evenly distributed across LAs in accordance with the sampling strategy for national representativeness as each administrative unit has approximately the same population. However, there was more variation in the distribution of LWIIP7 participants across London, with notably higher proportions residing in the two adjacent south east Outer London boroughs of Bromley and Croydon at 11.3% and 9.1% respectively as seen in Map 3.2. Approximately, half the proportions for these two LAs were represented in LHSE2008.

Table 3.1 Details of the Whitehall II study sample size and response rate from recruitment to the phase used in this project.

Phase	Dates	Type	Number of participants	Response Rate
1	1985-1988	Screening / questionnaire	10,308	73% of those invited
2	1989-1990	Questionnaire	8,132	79% of Phase 1 responders
3	1991-1994	Screening / questionnaire	8,815	86% of Phase 1 responders
4	1995-1996	Questionnaire	8,628	84% of Phase 1 responders
5	1997-1999	Screening / questionnaire	7,870	76% of Phase 1 responders
6	2001	Questionnaire	7,355	71% of Phase 1 responders
7	2002-2004	Screening / questionnaire	6,967	68% of Phase 1 responders

Table 3.2 Summary of non-biological data collected in in Phase 7 of the Whitehall II study. Adapted from Marmot and Brunner [182].

	Data category	Data topics relevant to present study
Non-biological	Demographic & socioeconomic data	Education; Household composition; Income; Financial assets; Work and work change (retirement)
	Area-level indicators	Deprivation; Classification of area
	Psychosocial/work exposure	n/a
	Health behaviours	Physical activity (Walking; Sports; Gardening; Housework; Do-it-yourself)
	CVD	n/a
	General health (subjective)	n/a
	Mental health (subjective)	General Health Questionnaire (GHQ) (anxiety, depression)
	Health outcomes (objective)	n/a

Table 3.3 Overall frequencies of socio-demographic variables including missing data for WIIP7 (N=6885) and LWIIP7 (N=3020).

		WIIP7		LWIIP7	
Socio-demographic factor		N	%	N	%
Sex	Male	4835	70.23	1869	61.89
	Female	2050	29.77	1151	38.11
Age group	50y to <56y	1670	24.26	800	26.49
	>=56y to <60y	1611	23.40	672	22.25
	>=60y to <66y	1765	25.64	771	25.53
	>=66y to 75y	1839	26.71	777	25.73
Economic activity	Remaining in Civil Service	1975	28.69	999	33.08
	Working outside Civil Service	1415	20.55	565	18.71
	Not working - retired	3073	44.63	1260	41.72
	Out of work	277	4.02	124	4.11

		WIIP7		LWIIP7	
Socio-demographic factor		N	%	N	%
	Not working (long-term sick)	118	1.71	60	1.99
	Missing	27	0.39	12	0.40
Income	<£20,000	1256	18.24	637	21.09
	£20,000 to <£40,000	2043	29.67	853	28.25
	£40,000 to <£60,000	1232	17.89	475	15.73
	>=£60,000	821	11.92	358	11.85
	Missing	1533	22.27	697	23.08
Car availability	Car available	5796	84.18	2274	75.30
	No car available	887	12.88	638	21.13
	Missing	202	2.93	108	3.58
Occupational social class	Administrative	3064	44.50	1103	36.52
	Professional or executive	2951	42.86	1344	44.50
	Clerical	752	10.92	517	17.12
	Missing	118	1.71	56	1.85
Education	No academic qualifications	589	8.55	307	10.17
	O-level	1534	22.28	646	21.39
	A-level	1658	24.08	647	21.42
	BA/BSc	1314	19.08	597	19.77
	Higher degree	821	11.92	370	12.25
	Missing	969	14.07	453	15.00
Marital status	Married	4956	71.98	1961	64.93
	Cohabit	281	4.08	131	4.34
	Single	1059	15.38	614	20.33
	Divorced	492	7.15	257	8.51
	Widowed	81	1.18	47	1.56
	Missing	16	0.23	10	0.33
Ethnicity	White	6315	91.72	2538	84.04
	Non-white	570	8.28	482	15.96

Table 3.4 Data from the household and individual interview, and nurse visit available for adults (aged 16 years and over) in the core general study sample for HSE 2008. Adapted from Mindell *et al* [184].

Data category	Data topic relevant to present study
Health measures (self-reported)	Acute illness in the past 2 weeks; General health and longstanding illness; Limiting longstanding illness; General Health Questionnaire (GHQ)-12 ; EuroQoL-5D
Health-related behaviours (self-reported)	n/a
Biological measurements	n/a
Individual socio-demographic characteristics	Economic status; Occupation; Marital status; Benefit receipt; Educational attainment; Ethnic group
Household characteristics	Household income; Economic status of household reference persona ; Occupation of household reference persona; Car ownership

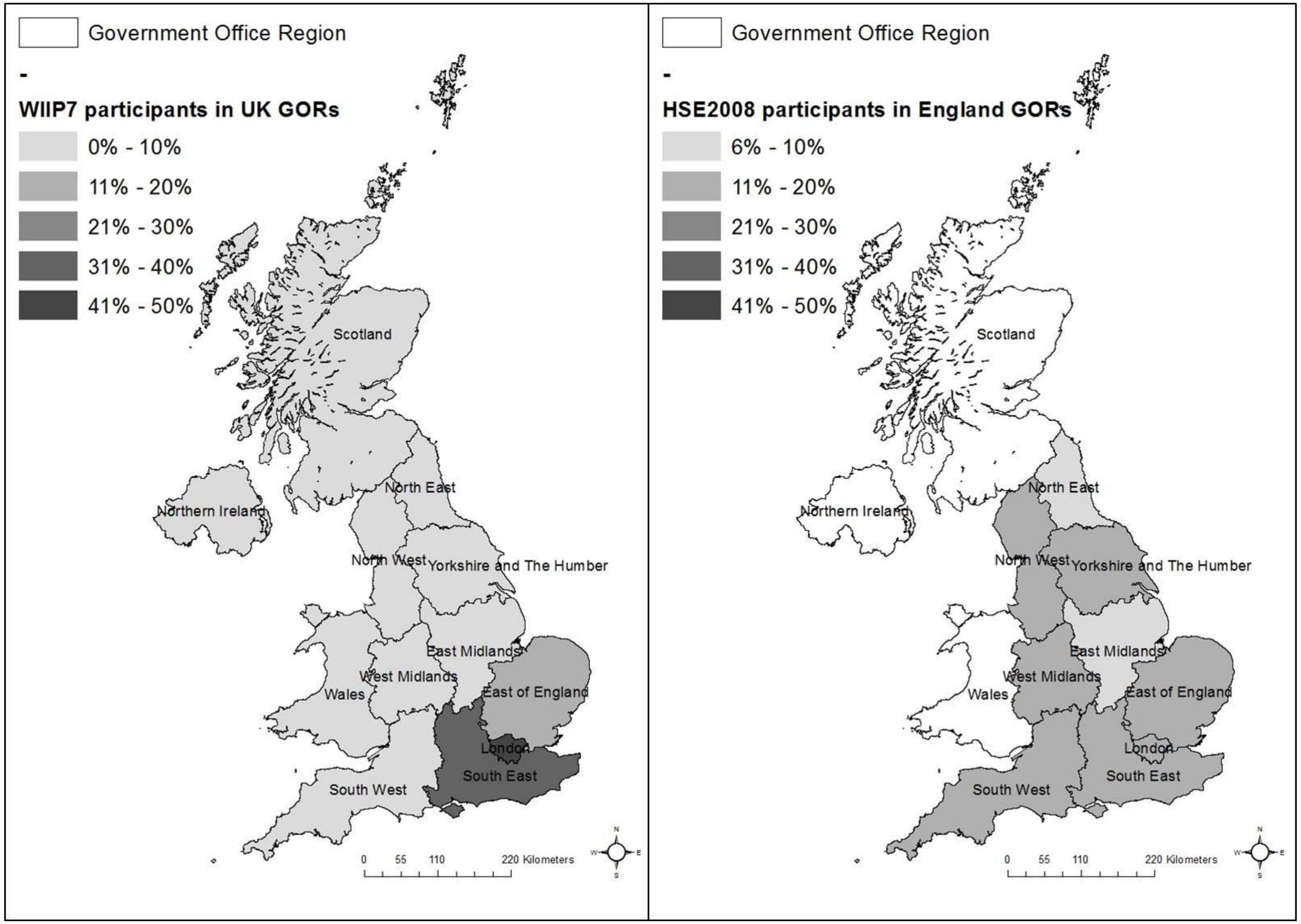
[a]The household reference person (HRP) is the person in whose name the person the property is owned or rented; if there is more than one, the person with the highest income. If two householders have the same highest income the HRP is the oldest one.

Table 3.5 Overall frequencies of socio-demographic variables including missing data for HSE2008 (N=14,221) and LHSE2008 (N=1540).

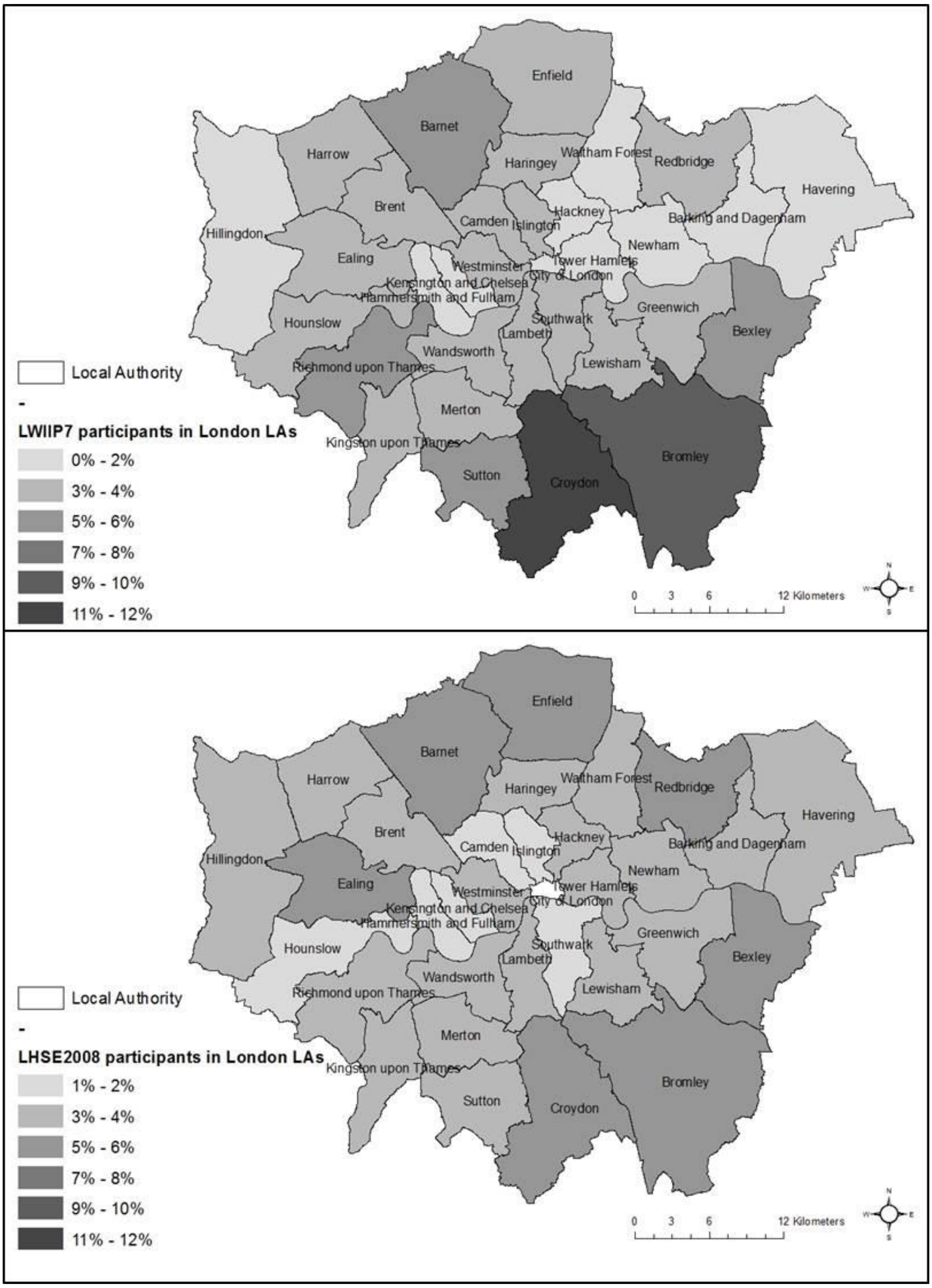
Socio-demographic factor		N	%	N	%
Sex	Male	6331	44.52	694	45.06
	Female	7890	55.48	846	54.94
Age group	16y-34y	3634	25.55	499	32.40
	35y-54y	4922	34.61	570	37.01
	55y+	5665	39.84	471	30.58
Economic activity	Employed	7838	55.12	862	55.97
	Unemployed	634	4.46	100	6.49
	Retired	3564	25.06	289	18.77
	Other	2171	15.27	284	18.44
	Don't know	5	0.04	0	0.00
	No answer/refused	9	0.06	5	0.32
Income	<=£14,918	3477	24.45	387	25.13
	>£14,918-£31,871	3906	27.47	316	20.52
	>£31871	4128	29.03	347	22.53
	Not applicable	2710	19.06	490	31.82
Occupational social class	Managerial and professional	4696	33.02	4	0.26
	Intermediate	3113	21.89	564	36.62
	Routine and manual	5640	39.66	331	21.49
	Other	762	5.36	491	31.88
	Not applicable	10	0.07	150	9.74
Education	No qualification	3618	25.44	377	24.48
	Up to and including O-level	3840	27.00	322	20.91

Socio-demographic factor		N	%	N	%
	A-level	2111	14.84	214	13.90
	Above A-level below degree	1607	11.30	133	8.64
	Degree	2802	19.70	478	31.04
	Foreign/other	237	1.67	16	1.04
	No answer/refused	1	0.01	0	0.00
	Not applicable	5	0.04	0	0.00
Marital status	Married	7529	52.94	743	48.25
	Cohabit	1602	11.27	163	10.58
	Single	2690	18.92	393	25.52
	Divorced	1285	9.04	135	8.77
	Widowed	1115	7.84	106	6.88

Map 3.1 Percentage of WIIP7 (left) and HSE2008 (right) residing in each GOR. HSE2008 were confined to England by virtue of the study design.



Map 3.2 Percentage of LWIIP7 (top) and percentage of LHSE2008 (bottom) in London residing in each LA.



3.3 Variables and descriptive analysis

3.3.1 Geographies of the United Kingdom as units of analysis

The effect of neighbourhood delineation as different levels of administrative geographies on associations between exposures and outcomes was a major focus of this thesis. It is important to outline the spatial characteristics of administrative units by means of descriptive statistics, identifying inter-regional differences, as this information is pertinent to exposure variable construction and to hypothesised associations between neighbourhood exposures and outcomes. Although there are additional levels of geographies in the United Kingdom, such as lower layer super output areas and middle layer super output areas, three levels were selected for investigation in this thesis, representative of the top, middle and bottom of the hierarchy; local authorities (the largest), Census Area Statistics wards and output areas (the smallest). The distances between these three levels on the hierarchy were judged sufficient for detection of differential effects of scale on associations between environmental exposures and health and health-related behaviour outcomes.

3.3.1.1 Local authorities

Of the United Kingdom's 434 local authorities (LAs), 354 were in England, 22 in Wales, 32 in Scotland and 26 in Northern Ireland. In London there were 33 LAs, the central twelve of which constitute Inner London with the remainder comprising Outer London.

3.3.1.2 Census Area Statistics wards

Below LAs in the administrative geography hierarchy are Census Area Statistics (CAS) wards. In 2001 there were 7969 CAS wards in England, 881 in Wales, 1222 in Scotland and 582 in Northern Ireland. There were 633 CAS wards in London of which 230 were in Inner London and 403 in Outer London.

3.3.1.3 Output areas

Positioned below CAS wards in the administrative geography hierarchy, the output areas (OAs) discussed in the present research were defined for the 2001 Census , with 165,649 OAs for 2001 Census output in England, 9,769 in Wales, 42,540 in Scotland and 5,015 in Northern Ireland. In London there 24,141 OAs, 9,032 in Inner London and 15,106 and Outer London, respectively. OAs for England and Wales had a minimum size of 40 households or 100 resident people, but with a recommended size of 125 households. Postcodes defined on Census Day in 2001 formed the basis of the OAs of England and Wales and all OAs were delineated to fall within the 2003 electoral ward boundaries. The same specifications applied for the OAs of Northern Ireland but the postcodes on which the OAs were based derived from January 2000 and fitted within the 2001 electoral ward boundaries. The threshold residential populations specified for OAs in Scotland was considerably lower: the minimum was 20 households or 50 resident people, with a target of 50 households. The postcodes on which the OAs were based were those of December 2000 which related to the electoral ward boundaries of 2001 but not all OAs fitted within them.

3.3.1.4 Descriptive statistics for units of analysis

Statistical analyses were performed to quantitatively describe the units of analysis. The descriptive statistics derived are graphically illustrated in Appendix 3.1, Appendix 3.2, Appendix 3.3, Appendix 3.4 and Appendix 3.5. LAs in London tended to be much smaller than those of other GORs with the median LA area in London of 38.7Mm^2 around an eighth that for the median for the whole of the UK of 288Mm^2. Within London there was a notable difference between Inner and Outer London in the median sizes of LAs, that of Outer London at 56.0Mm^2 over twice that of Inner London (21.8Mm^2). In contrast to the very large difference in median LA area for London compared to that for the median for the whole of the UK, the median CAS ward area difference was less stark: the median CAS ward area for London at 1.8Mm^2 was around a third that of the median for the whole of the UK at 4.9Mm^2. Nevertheless, median CAS ward area for London was substantially lower than that for the UK. Variation in area between London wards was lower than for other regions of the UK, such as Scotland which had many extreme outliers. Whilst CAS ward area size was more uniform within London than in other regions (Appendix 3.2), there was much variation in median CAS ward size between LAs (Appendix 3.3) which upon further analysis reflects differences between wards in Inner and Outer London. As for LAs, within London there was a notable difference between Inner and Outer London in the median sizes of CAS wards, with that of Outer London at 2.3Mm^2 over twice that of Inner London (1.1Mm^2). There was variation in OA areas between GORs which, in contrast to variation in LA areas between GORs and to variation in CAS ward areas between GORs, was modest: the interquartile ranges of OA areas in almost all GORs included the median OA area for the UK. Only the interquartile range of London fell just below the UK median of 60,812m^2. The median OA area for London was 34,850m^2 whilst those for Inner and Outer London were 21,885m^2 and 43,807m^2 respectively.

In summary, univariate descriptive analysis for administrative areas reveals not only large differences in areas between the hierarchical levels but also substantial inter and intra-regional differences within each hierarchical level, be it LA, CAS ward or OA, particularly between London and other regions. This is an important consideration when analysing the associations between the environment measured within an individual's administratively defined neighbourhood and health outcomes. For example, a neighbourhood in the London region operationalised as a local authority could be of similar size to a neighbourhood defined as a CAS ward outside, masking any effect of differential neighbourhood delineation. Therefore, it was deemed necessary to conduct sub-analyses of associations between exposures and outcomes for WIIP7 and HSE2008 living in this region. Also, given the comparatively high density of administrative areas in London, it was considered important to construct quantile variables from environment exposures measured in administrative units in London alone in order to generate sufficient heterogeneity in levels of exposure.

3.3.2 Exposure variables

The exposures investigated in this thesis were as follows:

- physical activity as indicated by: meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity; meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; and being in the top tertile of the WIIP7 sample for time spent walking per week (Chapter 4, Study 1)
- greenspace and domestic garden as recorded in the Generalised Land Use Database (Chapter 5, Study 2);
- objectively measured environmental quality as indicated by the Multiple Environmental Deprivation Index (Chapter 6, Study 3);
- perceived environmental quality in terms of satisfaction with local parks and open spaces as indicated by the 2008 Place Survey (Chapter 6, Study 3); and
- walkability (Chapter 8, Study 5), the derivation of which is detailed in Chapter 7

The administrative area in which a study participant resided represented the neighbourhood of exposure. (Additionally, GIS-computed zones around an individual's residential postcode represented the neighbourhood of exposure for walkability but the modelling of walkability is addressed in Chapter 7 rather than here.) None of the environmental exposures subject to investigation in this thesis were measured de novo. Therefore, the administrative level in which environmental exposure data was originally measured constituted the lowest neighbourhood representation, or operationalisation, possible as charted in Table 3.6. Whilst disaggregation of environmental exposure measures to lower levels was not possible, it was possible and indeed necessary to re-aggregate data to a higher administrative geography if study participants were not identifiable at the level of original measurement. For the WIIP7 study sample, identifiable to postcode level, there was no lower limit to the administrative scale at which neighbourhoods could be delineated. However, it was necessary to re-aggregate data and delineate larger neighbourhoods for the HSE2008 study samples for exposures which had been measured within administrative units of lower level than that at which this study sample was identifiable, the local authority. Given that local authority was the highest administrative geography of neighbourhood operationalisation in this thesis, it was the only one for HSE2008.

Although some of the original environmental data from which exposures were derived was continuous rather than discrete, all neighbourhood environmental exposures were constructed as categorical variables. Categorisation was done to facilitate interpretation of results of logistic regression: the odds of an outcome being two fold higher for one level of exposure relative to another, for example, is easier to comprehend than the odds of the outcome being two fold higher for every percent increase in the exposure. Results that elicit straightforward interpretation are more useful for policy development. A qualitative and quantitative examination of environmental measures for the administrative areas of original data collection was necessary to guide optimal construction of categorical exposure variables. Here an overview of each environmental measure is given and descriptive statistics presented.

Table 3.6 The administrative level at which environmental exposure data was originally measured.

Chapter/Study	Environmental measure	OA	CAS ward	LA
5/2	Greenspace and domestic garden	✓	✓	✓
6/3	Objectively measured environmental quality		✓	
6/3	Satisfaction with local parks and open spaces			✓

3.3.2.1 Greenspace as recorded in the Generalised Land Use Database

The United Kingdom's National Land Use Database (NLUD) version 4.4 establishes the national standard classification system for naming and defining land[186]. Despite its name NLUD classifies groups of both land use and land cover, distinct but related terms. Land cover is "the observed (bio)physical cover on the earth's surface" whereas land use refers to "the arrangements, activities and inputs by people to produce, change or maintain a certain land cover type"[187]. Thus, land cover is restricted to the terrestrial surface whilst land use essentially describes humans' interaction with land cover, which may occur above the surface, such as in a shopping centre, or deep below, such as in a mine. The Generalised Land Use Database (GLUD) is a classification system derived from NLUD which allocates all identifiable land features in the form of land cover in England, on the national large-scale digital topographic mapping database provided by Ordnance Survey (OS) MasterMap into ten simple categories: "domestic buildings"; "non-domestic buildings"; "roads"; "paths"; "rail"; "gardens"; "green space"; "water"; "other land uses"; and "unclassified"[188]. The OS MasterMap Address layer is used in conjunction with the Topography layer to differentiate domestic and non-domestic buildings. Within OSMasterMap, real world objects such as buildings have attributes associated with them that are encoded as sixteen-digits called TOIDs, making them uniquely identifiable and readily classified within GLUD. The production of GLUD served primarily to support policies that would increase quality of life in urban areas by, for example, enabling statistics to be generated on availability of green recreational space. This explains why land covers of greenspace and domestic gardens are distinct categories despite having a common natural form. Land classified as "other land uses" refers predominantly to hardstanding such as that used for vehicle parking and tarmacked tennis courts whilst "unclassified" land cover is that which cannot be categorized using the GLUD automated methodology. GLUD details the areas of land within various administrative units, the lowest in the hierarchy being OAs and the highest GORs. The components detailed in GLUD of interest as exposure variables in the present study were the natural land uses classified as greenspace and domestic garden.

GLUD provides the total area of the administrative unit, enabling users to calculate proportions of each land category within a given administrative unit. In addition it gives the total land area attributed, or allocated, to that administrative unit and the percentage difference between the total allocated land and the actual land area that is administratively defined. Whilst in most cases the total land allocated is approximately the same as the total administratively defined, in some cases there is a considerable difference. If the administratively-defined land area were used in the calculation of the percentage of a given GLUD land use for that administrative area, a large area of the particular land use could fall outside the administrative boundary but would be counted, and contributing to the numerator, result in an inaccurately large percentage, possibly exceeding 100%. However, if total land allocated by GLUD were used as the denominator, the percentage might reflect that of an area smaller or larger than the actual administrative unit. It was important to choose and consistently use one denominator type and attach percentages derived from this denominator type to the study samples. Paired t tests were performed at OA, CAS ward and LA level to determine whether, at each level the difference between the total mean area of land allocated to a given administrative area was significantly different from the actual administrative area. At all administrative geography levels across England the total mean area allocated by GLUD to an administrative area significantly exceeded the mean area defined by administrative boundary. The t-values were 3.403, 3.279 and 3.209 for LAs, CAS wards and OAs, respectively all with p=<0.001. Conversely, at all administrative geography levels specifically within London the mean area defined by the administrative boundary exceeded the total mean area allocated by GLUD although the differences were not significant. The fact that significant mean differences were observed for all administrative levels across England, but were not seen specifically within London, may be attributable to the smaller land parcels catalogued by GLUD within the high density London region. Whilst t tests revealed significant differences across England but not specifically within London, it was important to select a consistent denominator for the calculation of land use percentages and therein construction of the greenspace use exposure variables at all geographical scales and for all areas. Total allocated land was chosen over administratively defined areas because this was considered to better represent actual environmental exposure. Total percentages of domestic garden and greenspace were calculated for all administrative units across England at LA, CAS ward and OA level.

3.3.2.1.1 Univariate descriptive statistics for greenspace across England and London

There were 354 LAs in England for which areas of domestic garden and greenspace were detailed in GLUD. The median percentages of domestic garden and greenspace within these administrative areas were 6.6 and 82.4, respectively. Within London, however, a different pattern was apparent: their respective median percentages were 22.8 and 31.4. Greenspace accounted for the greatest proportion of most LA areas in England but in London it accounted for a much smaller proportion, with domestic garden accounting for a larger proportion. Areas of domestic garden and greenspace were detailed in GLUD for 7969 CAS wards in England of which 633 were in London. The median percentages of domestic garden and greenspace for all England CAS wards were 12.8 and 63.7, respectively. Within London, their respective median percentages were 25.5 and 21.9. The patterns of the median percentages for these land uses were the same as for LAs, with greenspace accounting for the greatest proportion of most CAS ward areas in England but in London accounting for a lower proportion in most. Again, domestic gardens accounted for a larger proportion within London than England-wide. There were 165438 OAs in England, with 24122 of these in London, for which areas of domestic garden and greenspace were detailed in GLUD. The England-wide median percentages of domestic garden and greenspace within these administrative areas were 33.1 and 19.5, respectively, whilst those of London alone were 33.8 and 7.7, respectively. Boxplots illustrate the differences at the three administrative levels across England in the relative proportions of greenspace and of domestic garden in Appendix 3.8. Proportions of greenspace in CAS wards across England and proportions of greenspace and domestic garden across London are mapped in Appendix 3.9.

3.3.2.1.2 Construction of greenspace, and green space plus domestic garden, exposure variables

The disparities between the median proportions of greenspace and domestic garden at each administrative level for England and for London alone warranted sub-analysis of London for associations with outcomes in the study samples. In the derivation of the categorical exposure variables, greenspace, and green space plus domestic garden, percentages were transformed into quintiles because five categories were considered to encompass sufficient levels of exposure. They were derived at each administrative level from all units across England and also from all units across London and attributed to participants of each study sample, at LA, CAS ward and LA level in WIIP7 and at LA level only in HSE2008.

3.3.2.1.3 Univariate descriptive statistics for greenspace and domestic garden in WIIP7

The frequencies of WIIP7 participants in each England-derived quintile for proportion of greenspace, and for proportion of greenspace plus domestic garden, at each administrative level are presented in Table 3.7. They are also presented for each London-derived quintile. It is notable that at LA and CAS ward level that WIIP7 were overrepresented in the lowest England-derived quintiles for greenspace with frequencies of 43.5% and 38.6% respectively, a consequence of their overrepresentation in the London region. The relatively more even distribution of the WIIP7 study sample across the London-derived quintiles at these administrative levels highlights the value of generating London-specific quintiles for examination of associations with outcomes.

3.3.2.1.4 Univariate descriptive statistics for greenspace and domestic garden in HSE2008

Whilst not equally represented across England-based and across London-based LA-level quintiles of greenspace and domestic garden, the HSE2008 study sample were more evenly distributed than the WIIP7 study sample (Table 3.8). This reflects the nationally representative sampling design of the study from which the HSE2008 was drawn. However, the finding that HSE2008 participants were not exactly distributed across the quintiles can be attributed to the fact that this data is not weighted: those who responded to the survey (64%) were not evenly distributed across LAs.

Table 3.7 Frequencies of LA, CAS ward and OA-level quintiles of greenspace (GS), and greenspace plus domestic garden (GS+DG) of WIIP7 participants across England and across London.

Admin level	Quintile	Frequency of England-based LA-level land use quintiles (%)				Frequency of London-based LA-level land use quintiles (%)			
		GS		GS+DG		GS		GS+DG	
		n	%	n	%	n	%	n	%
LA	1	2905	43.5	2491	37.3	381	12.6	295	9.8
	2	1351	20.2	1584	23.7	465	15.4	561	18.6
	3	986	14.8	882	13.2	550	18.2	511	16.9
	4	989	14.8	1067	16.0	916	30.3	642	21.3
	5	446	6.7	653	9.8	708	23.4	1011	33.5
	Pearson chi2(4)*		2600		1600		298.2		443.4
CAS ward	1	2578	38.6	2119	31.7	499	16.5	366	12.1
	2	1436	21.5	1638	24.5	553	18.3	429	14.2
	3	1354	20.3	1436	21.5	673	22.3	640	21.2
	4	837	12.5	923	13.8	589	19.5	689	22.8
	5	472	7.1	561	8.4	706	23.4	896	29.7
	Pearson chi2(4)*		1900		2.9×10^{9}		48.0		299.8
OA	1	1853	27.8	1046	17.2	712	23.7	343	11.4
	2	1333	20.0	1262	20.8	682	22.7	427	14.2
	3	1163	17.5	1164	19.1	591	19.6	611	20.3
	4	1122	16.8	1524	25.1	505	16.8	719	23.9
	5	1192	17.9	1086	17.9	519	17.3	907	30.2
	Pearson chi2(4)*		4×10^{7}		4.6×10^{7}		58.0		340.0

Admin level	Quintile	Frequency of England-based LA-level land use quintiles (%)				Frequency of London-based LA-level land use quintiles (%)			
		GS		GS+DG		GS		GS+DG	
		n	%	n	%	n	%	n	%

*All chi2 test statistics for a difference between the observed frequencies and the expected frequencies of 20% for each quintile were significant with p<0.001.

Table 3.8 Frequencies of LA-level quintiles of greenspace (GS), and greenspace plus domestic garden (GS+DG) of HSE2008 participants across England and across London.

Quintile	Frequency of England-based LA-level land use quintiles (%)				Frequency of London-based LA-level land use quintiles (%)			
	GS		GS+DG		GS		GS+DG	
	n	%	n	%	n	%	n	%
1	3708	26.1	3620	25.5	212	13.8	217	14.1
2	3395	23.9	3468	24.4	269	17.5	269	17.5
3	2807	19.2	2849	20.0	345	22.4	313	20.3
4	2412	17.0	2405	16.9	326	21.2	357	23.2
5	1899	13.4	1879	13.2	388	25.2	384	24.9
Pearson chi2(4)*		749.3		743.8		61.1		58.5

*All chi2 test statistics for a difference between the observed frequencies and the expected frequencies of 20% for each quintile were significant with p<0.001.

3.3.2.2 Objectively measured environmental quality (OEQ) as indicated by the Multiple Environmental Deprivation Index (MEDIx)

The Multiple Environmental Deprivation Index (MEDIx), a UK-wide Census Area Statistics (CAS) ward level summary environmental index based on health-relevant dimensions of the physical environment, was developed by Richardson *et al* to enable evaluation of the effects of complex physical environments on health in the UK [189]. In summary, higher levels of air pollution, lower temperatures and greater proximity to industry were classified as health detrimental, or pathogenic, whilst higher levels of UV radiation and greater areas of greenspace were classified as health beneficial, or salutogenic. Population-weighted scores were generated by summing component scores, giving a final score ranging from -2 to +3, with higher scores indicative of worse overall environmental quality. Calculation of MEDIx scores is summarised in Table 3.9.

3.3.2.2.1 Univariate descriptive statistics for OEQ across the UK, England and London

Across all UK wards, the frequencies of ward level MEDIx showed an uneven distribution with only 3.2% of UK CAS wards scoring -2, indicative of the highest environmental quality, and 0.4% scoring +3, indicative of the lowest (Table 3.10). The highest frequencies were for the middle scores of 0 and +1. Across all England wards only, the pattern of frequencies was similar, with 4.3 and 0.4% of English CAS wards scoring -2 and +3 respectively, and the highest frequencies for the middle scores. However, a distinctive pattern emerged for London, with overrepresentation of CAS wards of worse quality: 86.6 and 11.2% scored +1 and +2 respectively.

3.3.2.2.2 Construction of OEQ exposure variable

As evidenced by the results of univariate statistical analysis of MEDIx scores (Table 3.10), the construction of the index was such that MEDIx scores were not evenly distributed across UK CAS wards. This was problematic because it was more likely that individuals within the study samples would reside in CAS wards falling in the mediocre categories of 0 and +1 by virtue of the fact that they were more frequently represented. Thus, differentiation of individuals based on multiple deprivation would be limited. The polarisation of scores was particularly marked within London, with the vast majority of fairly low environmental quality by UK MEDIx standards and none of the highest (as indicated by a score of -2) or lowest (+3). For the present research it was important to generate MEDIx scores specific to certain regions and at different geographical scales relevant to, and available in, the study samples: the HSE2008 study sample was restricted to England with respect to spatial coverage and neighbourhoods could only be operationalised as local authorities, and the WIIP7 study sample was particularly concentrated in London and the South East. Using the metadata kindly provided by Richardson (*pers comm* [190]), individual MEDIx component scores were calculated for all CAS wards in the UK. This enabled indices for England only, and for London only, to be developed. In addition it allowed the derivation of local authority-level scores. Creating scores for these administrative units was not simply a case of calculating the mean score of the wards they contained: it was essential to weight the components with respect to population size. The population of each LA was computed as the sum of those its constituent wards. Although it was possible to obtain LA population figures directly, it was considered more accurate to sum the populations of constituent wards to obtain the LA populations because LA boundaries change more frequently than CAS ward boundaries which can change approximately every ten years at the time of the census. The MEDIx score of a constituent CAS ward of a LA was weighted by the proportion of the LA population represented by the ward to give the proportion of the score attributable to the LA. A whole score for each LA was then calculated as the sum of the population-weighted constituent CAS ward scores. If data on one or more of the CAS ward populations within a LA was missing it was not possible to construct a score for that LA so MEDIx score coverage was not as comprehensive at this administrative level as it was for CAS ward.

As each MEDIx score was not equally represented by CAS wards or by LA wards whether on a UK, England or London only based index, there was not an even distribution of relative levels of environmental quality exposure. Also, as a consequence of their calculated MEDIx scores at LA level were not integers, but a continuous variable. Therefore, MEDIx scores were dichotomised at the median, with those scoring equal or greater than the median classified as of "low" objective environmental quality (OEQ) and attached to WIIP7 and HSE2008 participants via their administrative codes. Henceforth the objective environmental quality exposure is referred to as OEQ.

3.3.2.2.3 Univariate descriptive statistics for OEQ in WIIP7

Among all WIIP7 participants across the UK, just over 50% resided in CAS wards with low OEQ, and within England only approximately the same proportion lived in low OEQ CAS wards as indicated by the UK-based MEDIx score (Table 3.11). However, within London only, the vast majority of WIIP7 participants (94.0%) resided in CAS wards of low OEQ as indicated by the UK-based MEDIx score whilst 33.4% lived in low environmental quality CAS wards as indicated by the London-based MEDIx score. A similar pattern was evident for LA level wards (Table 3.11). This highlights the importance of environmental quality benchmarks for London and the derivation of a London-specific MEDIx index to detect variation in OEQ for this region.

3.3.2.2.4 Univariate descriptive statistics for OEQ in HSE2008

A slightly higher proportion (60.0%) of HSE2008 participants across England than WIIP7 participants across England resided in LAs of low OEQ as indicated by the UK-based MEDIx score (Table 3.12). All HSE2008 participants in London resided in LAs of low OEQ as indicated by this variable but only 41.6% lived in LAs of low OEQ as indicated by the London-based MEDIx score. The fact that a slightly higher proportion – almost half – of London-dwelling HSE2008 participants lived in low OEQ LAs as indicated by the London-based MEDIx score than the proportion of London-dwelling WIIP7 participants reflects the national representativeness of the former study sample. The findings for the HSE2008 sample with respect to MEDIx exposure variables again highlight the importance of the London-specific MEDIx index.

Table 3.9 Calculation of MEDIx score.

	Component	Score
Detrimental	Air pollution	+1 (highest quintile) or 0
	Proximity to industry	+1 (highest quintile) or 0
	Temperature	+1 (lowest quintile) or 0
Beneficial	Greenspace availability	-1 (highest quintile) or 0
	UV level	-1 (highest quintile) or 0
MEDIx score		-2 (best) to +3 (worst)

Table 3.10 Frequencies of ward level MEDIx scores for CAS wards in the UK, England, and London derived from ward component scores for the UK, with a score of -2 indicative of the best qaulity and +3 of the worst.

	CAS ward score frequency (%)		
UK-based MEDIx score	UK	England	London
-2	3.2	4.3	0.0
-1	18.1	21.8	1.0
0	34.5	32.2	1.3
+1	35.1	33.1	86.6
+2	8.7	8.2	11.2
+3	0.4	0.4	0.0

Table 3.11 Frequencies of WIIP7 participants living in low OEQ CAS wards and LAs as indicated by MEDIx scores.

Admin level	Exposure	UK participants	England participants	London participants
CAS ward	UK-based low OEQ	52.8	53.2	95.3
	England-based low OEQ	N/A	56.1	94.0
	London -based low OEQ	N/A	N/A	33.4
LA	UK-based low OEQ	56.6	56.8	100.0
	England-based low OEQ	N/A	59.2	100.0
	London -based low OEQ	N/A	N/A	35.2

Table 3.12 Frequencies of HSE2008 participants living in low OEQ LAs as indicated by MEDIx scores.

Exposure	England participants (%)	London participants (%)
UK-based low OEQ	60.0	100.0
England-based low OEQ	60.3	100.0
London -based low OEQ	N/A	41.6

3.3.2.3 Perceived environmental quality (PEQ) in terms of satisfaction with local parks and open spaces as indicated by the 2008 Place Survey

In 2008 the government commissioned the Place Survey, a postal questionnaire administered by every LA in England to gather information about residents' perceptions of their local areas with the aim of tracking performance on National Indicators and other important issues. As one of the largest surveys in Europe, with responses from over half a million adults, it provided a wealth of information on public perceptions on topics such as satisfaction with neighbourhoods, attitudes towards public services and levels of civic participation. The survey's design was criticised for potential bias: although most LAs reached the target of 1,100 responses, the response rates were generally low with thirty LAs recording rate of less than 30 per cent. Nevertheless, the 2008 survey provided a useful measure of perceived environmental quality as an exposure measure that was independent of the present study's participants and therefore could be linked to them at the level of the LA without confounding outcomes. Level of satisfaction with parks and open spaces was calculated as the total percentage of respondents reporting they were either *"Very satisfied or "Fairly satisfied"* in answer to the question *"How satisfied or dissatisfied are you with [Parks and open spaces] provided or supported by [your local council]?*".

3.3.2.3.1 Univariate descriptive statistics for PEQ across England and London

There were 325 LAs in England in which satisfaction with local parks and open spaces was measured by the Place Survey. The median percentage of residents who were satisfied in these English LAs was 69.0%, ranging from 47.2 to 92.5%, and the mean was 68.7% (SD:7.9). Among the thirty three London LAs alone, the median percentage of residents who were satisfied was higher at 72.2%, ranging from 57.1 to 92.5% and the mean was also higher at 72.9% (SD:8.6). The slightly higher median level of satisfaction – perceived quality – for London LAs may be explained by higher objective quality due to greater investment councils on greenspace as a resource because it is more limited in area. Mapping of quintiled percentages derived from all English LAs across England, revealed higher levels of satisfaction in the south of England (Appendix 3.10A). Across London, mapping of quintiled percentages derived from London only LAs showed geographical clusters of LAs tended to have similar quantile scores, perhaps a reflection of their smaller size: the relatively small areas of London LAs may lead respondents to evaluate the quality of parks and open spaces outside of their own LAs (Appendix 3.10B).

3.3.2.3.2 Construction of PEQ variable

By attaching variables derived from LA level satisfaction with local parks and open spaces to the study samples, it was possible to investigate perceived environmental quality in relation to the outcomes independent of the participants' characteristics which could influence their own perceptions. Given the slightly higher median LA-level satisfaction with local parks and open spaces, it was deemed necessary to derive a London-specific variable for perceived environmental quality as well as an England-wide one. Dichotomous exposure variables were constructed for perceived environmental quality, consistent with the derivation of the objective environmental quality exposure variables, with dichotomisation at the median, and attached to the WIIP7 and HSE2008 study participants via the LA code. The variables were reverse coded such that being exposed constituted living in a LA where less than 69.0 % of residents were satisfied, for the England-based measure, and living in a LA where less than 72.2 % of residents were satisfied, for the London-based measure. Henceforth the perceived environmental quality exposure is referred to as PEQ.

3.3.2.3.3 Univariate descriptive statistics for PEQ in WIIP7

5.5 per cent of the Whitehall study sample resided in English and non-English LAs in which the Place Survey was not conducted. However, all LAs in London collected Place Survey data so data was available for all individuals living in this GOR. Chi-squared tests for the univariate frequency distributions were carried out. They revealed that there was a significant difference (chi2(4) = 2.1×10^3) between the observed and expected distribution of PEQ as indicated by satisfaction with English LA-based parks and open spaces quintiles for the Whitehall II study sample, assuming an equi-probability model. Also, there was a significant difference (chi2(4) = 304.91) between the observed and expected distribution of PEQ as satisfaction with London LA-based parks and open spaces quintiles for the London contingent of this study sample, again assuming an equi-probability model. These findings highlighted the fact that WIIP7 was not a geographically nationally representative sample. Of the 6508 individuals in the WIIP7 sample in England in LAs where the Place Survey was conducted, 28.5% resided in LAs classified as low PEQ. The proportion of WIIP7 sample resident in London (N=3020) residing in low PEQ LAs was 28.3% for the England-wide median cut-off and 43.1% for the London-based cut-off.

3.3.2.3.4 Univariate descriptive statistics for PEQ in HSE2008

13155 individuals (92.5%) in the HSE2008 sample lived in LAs where the Place Survey was conducted. There was a significant difference (chi2(4) = 107.96) between the observed and expected distribution of PEQ as indicated by satisfaction with English LA-based parks and open spaces quintiles for the HSE2008 study sample, assuming an equi-probability model. Again, assuming an equi-probability model, there was a significant difference (chi2(4) = 56.01) between the observed and expected distribution of PEQ as indicated by satisfaction with London LA-based parks and open spaces quintiles for the London contingent of this study sample. Although the deviations from the expected distributions were significant, the fact that the chi2 values were much lower than for those of the equivalent tests performed in WIIP7, reflects the national representativeness of the HSE2008 study sample. The national representativeness of the HSE2008 is again apparent within England, with approximately half (53.0%) residing in LAs classified as low PEQ, the cut-off for which was the median percentage of satisfaction across all LAs. The proportion of WIIP7 sample resident in London (N=1540) residing in low PEQ LAs was 32.8% for England-wide median cut-off and 54.1% for the London-based cut-off.

3.3.2.4 Implications of choice of exposure variable measurement

The choice of measurement in terms of quantification and of spatial units of analysis of the three exposure variables described here – greenspace and domestic garden, objectively measured environmental quality and perceived environmental quality in terms of satisfaction with local parks and open spaces – had several implications. Whilst MEDIx was an index and thus could only be treated as a categorical OEQ variable, greenspace and satisfaction with local parks and open spaces (PEQ) were measured on continuous scales (as percentages) but were constructed as categorical variables. Categorisation of exposure variables has advantages and disadvantages [191]. An advantage is the relative ease of interpretation of statistical results with relative risk presented for categorical exposure variables as opposed to regression coefficients for continuous ones. However, the choice of cut-offs in the derivation of categorical variables may influence findings, for example, in the construction of a dichotomous exposure variable of perceived environmental quality where low satisfaction with local parks and open spaces was operationalised as below the median percentage of satisfaction for LAs. It may be that this threshold is too high with this subjective measure of environmental quality only affecting outcomes when level of satisfaction is much lower. Another disadvantage is that statistical efficiency is lost when continuous measures are categorised: a bigger sample size is necessary to detect a significant association when the exposure variable is categorical. The decision to use only categorical exposure variables in the present study was made on the basis of graphical and statistical examinations of the exposure measures, and also for consistency across variables.

Investigating the relationship between neighbourhood physical environment exposures and the GHQ-12 depression outcome in HSE2008 was limited to LAs, the administrative geography at which participants were identifiable. However, the additional operalisation of neighbourhoods at the lower administrative geographies of CAS wards and OAs for examining relationships between exposures and the physical activity and depression outcomes in WIIP7 enabled comparisons to be made of scales of analysis.

3.3.3 Outcome variables

The outcome variables in this thesis were measures of depression in HSE2008 and in WIIP7, and of physical activity in WIIP7 only. The exact specification of the outcome variables was informed by availability of, and by evaluation of statistical analysis of relevant data collected in the study samples. Here descriptive statistics are presented and justification for the specification of the outcomes is given, and descriptive statistics for the specified outcomes are finally presented.

3.3.3.1 Depression

The 12-item General Health Questionnaire (GHQ-12) [57] (Appendix 3.6) was used to derive the depression outcome in this thesis because this tool was an integral component of both the HSE2008 and WIIP7 study questionnaires that was used to evaluate mental health. Higher scores were indicative of worse mental health.

3.3.3.1.1 Descriptive statistics for GHQ-12 scores WIIP7

GHQ data was missing for 259 (3.8%) participants in the WIIP7 study sample, including 132 (4.4%) of the London contingent. Although these individuals lacked information for this outcome variable, they were retained in the study for the analysis of the other outcome variable theme, namely physical activity. Also, it was useful to retain them for the development of the walkability index. GHQ-12 scores in both the whole WIIP7 study sample and the London contingent showed a reverse J-shaped distribution for the whole WIIP7 study sample (Graph 3.1). The majority of the sample, 62% in the whole (N=6626) and 61% in the London contingent (N=2888), scored 0, resulting in extreme positive skews.

3.3.3.1.2 Descriptive statistics for GHQ-12 scores in HSE2008

GHQ-12 scores showed similar distributions in HSE2008 to those seen in WIIP7: in both the whole HSE2008 study sample and the London contingent there was a reverse J-shaped distribution for the whole HSE2008 study sample (Graph 3.1). The majority of the sample, 63% in the whole (N=14221) and 61% in the London (N=1540) contingent, scored 0, giving the extreme positive skews.

3.3.3.1.3 Construction of GHQ-12 depression outcome variable

A diagnosis of depression can only be ascertained by a clinical interview that uses the Diagnostic and Statistical Manual of Mental Disorders criteria for illness [56]. Thus, a GHQ-12 score cannot be used as evidence of depression caseness. Nevertheless, the GHQ-12 is commonly used to create a binary caseness variable, with psychiatric caseness indicating increased risk of an individual receiving further psychiatric attention if s/he presented in general practice [192]. Therefore, in this thesis a variable was created in which depression was defined as a score of four or more in accordance with studies validating the GHQ-12 against standardised psychiatric interviews [193]. The bimodal scoring method was employed which has a 4-point scale, scored as follows dependent on the presence of symptoms: not at all = 0, same as usual = 0, more than usual = 1 and much more than usual = 1. With each item scoring either 0 or 1, the total score varies between zero and twelve. In acknowledgement that its derivation does not meet the strict criteria for a diagnosis of depression, and for brevity, the depression outcome is henceforth referred to as GHQ-12 depression.

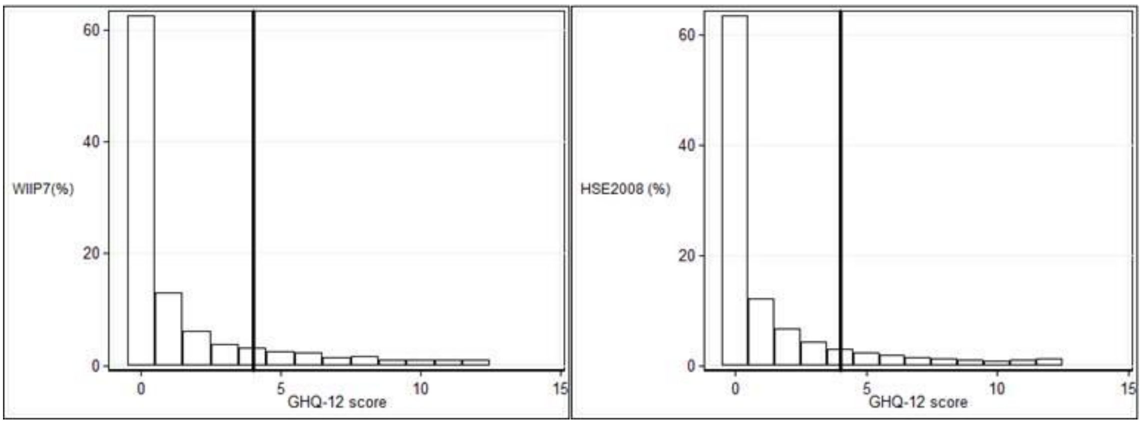
3.3.3.1.4 Descriptive statistics for GHQ-12 depression outcome in WIIP7

In the whole WIIP7 study sample prevalence of GHQ-12 depression, as indicated by a GHQ-12 score of 4 or more, was 14.6% and within the London contingent it was 15.1%.

3.3.3.1.5 Descriptive statistics for GHQ-12 depression outcome in HSE2008

In the whole HSE2008 study sample prevalence of GHQ-12 depression was 13.7% and within the London contingent it was 15.7%.

Graph 3.1 Histogram of GHQ-12 scores in the whole WIIP7 (left) and HSE2008 (right) study samples. The vertical lines mark the cut-off score of 4 for GHQ-12 depression caseness.



3.3.3.2 Physical activity

Self-reported physical activity was collected from twenty items on the self-administered WIIP7 questionnaire (Appendix 3.7). The types of physical activity for which data was collected in WIIP7 were walking and cycling, sports (football, golf, swimming and up to two other self-specified sports), gardening (mowing, weeding, hoeing, pruning and one other, self-specified gardening activity), housework (carrying heavy shopping, cooking, hanging out washing and up to two other, self-specified housework activities), do-it-yourself (manual car washing, painting, decorating and one other, self-specified do-it-yourself activity) and up to two additional, self-specified physical activities. With the exceptions of walking and cycling, WIIP7 participants' physical activities were self-reported in terms of the total volume (number of hours undertaken) in the last four weeks and as the total frequency (number of occasions) in the last four weeks. In the case of walking and cycling, participants reported average weekend and weekday daily volumes (number of hours per day) walked and cycled. It was not possible to differentiate walking and cycling undertaken as transport from walking and cycling undertaken as leisure.

3.3.3.2.1 Descriptive statistics for physical activity in WIIP7

3.3.3.2.1.1 Participation in physical activities

In preliminary analysis, variables were constructed to identify individuals as participants in a given physical activity if they reported that participation on at least one occasion in the last four weeks. Frequencies of participation among the WIIP7 study sample are presented in Table 3.13. Virtually all WIIP7 participated in walking and housework: participation frequencies for these activities were 99.4% and 98.1% respectively. Lower majorities of WIIP7 participated in gardening, sports and do-it-yourself but only 13.0% participated in cycling. It must be noted that for each physical activity data was not collected for the whole study sample, as indicated by the response rate in Table 3.13, and it may be that individuals with missing data did not actually participate, making true frequencies lower.

Table 3.13 Frequencies of participation among the WIIP7 study sample in physical activity domains for which data was collected.

Physical activity domain	Frequency (%)	Response rate (%) (N=6885)
Walking	99.4	95.2
Cycling	13.0	73.6
Gardening	74.6	91.0
Do-it-yourself	62.0	88.3
Housework	98.1	95.3
Sports	64.8	78.4
Other physical activity	22.2	47.8

3.3.3.2.1.2 Walking volume (duration) and speed (self-reported pace) of walking

In light of the high rate of participation in walking found in the study sample, together with the quality of data collected, and the importance of this physical activity in relation to neighbourhood environmental exposures, the characteristics of walking among WIIP7 participants merited further exploration.

Variables were constructed to give the total time spent walking per week as the sum of average weekend and weekday daily volumes, number of hours per day. Table 3.14 gives descriptive statistics for total time spent walking, stratified by self-reported walking pace. At any self-reported walking pace, the mean hours per week spent walking among the WIIP7 study sample for whom data was available (N=6409) was 5.4 (SD:4.2) and the median 4.5. Hours of self-reported walking per week ranged from 0 to 63. The mean and median hours per week spent walking among the London only contingent of the study sample, LWIIP7, for whom data was available (N=2756) were not markedly different at 5.5 (SD: 4.1) and 4.5, respectively. Although physical activity outcomes in HSE2008 were not investigated in this thesis, it is valuable to contrast these findings for the WIIP7 study sample with published findings for the nationally representative study from which the HSE2008 sample was derived. In the 2008 Health Survey for England the mean walking hours per week reported was 2.2 among men (N=7315, weighted) and 1.9 among women (N= 7676, weighted) with the median in both sexes 0.0 [6]. This contrasts with the higher durations of walking reported in WIIP7 and is counterintuitive given the older median age of the WIIP7 study sample. However, it is explained by the fact that the statistics for the 2008 Health Survey for England included only walking for participants reporting their usual pace as at least brisk. As shown in Table 3.14 over half of all WIIP7 participants reported their walking pace as less than brisk, at slow or steady average, and the median hours spent walking per week were 3.5 and 4.5 for slow and steady walkers, respectively. Thus it is likely if less than brisk walkers were included in the derivation of walking statistics for the 2008 Health Survey for England, results would be more similar and, given the younger median age, mean and median time spent walking probably higher. Another reason for the disparity between self-reported volumes of walking in the WIIP7 study sample and the 2008 Health Survey for England relates to the questionnaires items used to collect data. The Whitehall II questionnaire asked participants to report the average number of minutes spent on any walking on a weekday/ on a weekend day. As it is difficult to recall all walking, including for example walking around the house, without specific question prompts, over-reporting by WIIP7 participants was probable. More accurate walking volume data may have been collected in the 2008 Health Survey for England questionnaire in which physical activity was self-reported in response to questions pertaining to specific purposes for walking, such as climbing stairs at work, door to door sales and country walks.

Table 3.14 Total time spent walking per week in the WIIP7 study sample by self-reported walking pace

Walk pace	n	%	Min	Max	Mean	SD	Median
Any	6409	100.0	0.0	63.0	5.4	4.2	4.5
Slow (<3mph)	603	9.4	0.0	48.0	4.3	3.8	3.5
Steady average	2934	45.8	0.0	63.0	5.5	4.4	4.5
Brisk	2523	39.4	0.0	51.0	5.4	3.9	4.5
Fast (>4mph)	218	3.4	0.8	29.0	6.3	5.0	4.8
Missing	131	2.0	0.3	42.2	6.3	6.0	4.7

3.3.3.2.1.3 Physical activity hours per week as a function of intensity

As intensity is an important dimension of physical activity in terms of its health benefits, the WIIP7 dataset contained variables derived for previous research including some by Sabia *et al* [194]. Sabia *et al* assigned energy cost values as METs to each physical activity in the study questionnaire. They then categorised activities assigned values of less than 3 METS, 3 to 5.9 METS, and 6 METS or more, as mild, moderate and vigorous intensity, respectively. Sabia *et al* assigned a value of 2.9METs to walking because the walking pace reported by the majority of participants was less than brisk. Variable were then created that gave the total number of hours spent per week spent for each category of physical activity intensity and also the total number of occasions per week. In addition variables were also constructed in the dataset that assigned a higher MET value to walking, placing it in the moderate intensity category. Table 3.15 gives descriptive statistics for total hours per week spent for each intensity of physical activity in the WIIP7 study sample. Average hours of physical activity reported were higher in the lower intensity categories: the mean hours spent per week of mild intensity physical activity was 8.0 (SD: 5.0) with a median of 7.2 whereas the mean for vigorous intensity physical activity was 0.5 (SD: 1.0) and the median less than one hour. However, when the moderate intensity category included walking the median and mean hours per week spent walking, at 13.6 (SD: 11.9) and 11.3, respectively, exceeded the mean and median for mild intensity physical activity.

Table 3.15 Total hours per week of each physical activity intensity in the WIIP7 study sample

Physical activity intensity	N	Min	Max	Mean	SD	Median
Mild	6692	0.0	63.4	8.0	5.0	7.2
Moderate (excluding walking)	6694	0.0	33.6	3.2	3.0	2.4
Moderate (including walking)	6664	0.0	83.5	13.6	11.9	11.3
Vigorous	6705	0.0	8.1	0.5	1.0	0.0

3.3.3.2.2 Construction of physical activity outcomes

Three physical activity outcomes were specified:

- meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E);
- meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I); and
- being in the top tertile for time spent walking per week (TTW).

3.3.3.2.2.1 Construction of meeting the World Health Organisation (WHO) recommended physical activity level variables

Dichotomous outcome variables were constructed as meeting current recommendations for physical activity because this was indicative of behaviour likely to be associated with better health, as illustrated in the conceptual framework. Previous research using the dataset from which the WIIP7 study sample was drawn had specified meeting the World Health Organisation (WHO) recommended physical activity level as the outcome, informing the decision to use these guidelines for the present study [194]. If an individual met the WHO's minimum physical activity recommendation – at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or at least 75 minutes of vigorous intensity aerobic physical activity throughout the week, or an equivalent combination of moderate and vigorous-intensity activity – they were assigned a score of 1. If they did not achieve these recommendations but had complete data for the constituent variables, they were assigned a score of 0. In the previous research using the dataset from which the WIIP7 study sample was drawn, Sabia *et al* did not classify walking as a moderate intensity physical activity because over half of the study sample reported their usual walking as less than brisk [194]. Therefore, walking even among "brisk" and "fast" walkers did not contribute to meeting the WHO recommended level. (However, cycling fell in the moderate intensity category and, as such, contributed to meeting the WHO recommendation.) For the present research an outcome variable was constructed that was almost identical to Sabia *et al*'s, using derived variables kindly provided by them, differing only with respect to specification of an extra 15 minutes of vigorous physical activity, in accordance with more recent WHO guidelines. This outcome variable of meeting the WHO recommended level excluding all walking as a contributory factor is henceforth referred to as

WHOPA_E. Given that almost half of the WIIP7 study sample reported a brisk or fast usual walking pace, it was considered important to construct an outcome variable derived from a variable that included walking as a moderate intensity physical activity. Whilst inclusion of all walking regardless of pace would likely overestimate the prevalence of meeting the WHO recommended level in the WIIP7 study sample, exclusion of all walking as per Sabia *et al* could underestimate the prevalence. Therefore, an additional variable was created for the present study which differed from WHOPA_E only with respect to including walking as a contributory factor in meeting the WHO recommended level if reported as at least brisk. Henceforth this variable is henceforth referred to as WHOPA_I.

3.3.3.2.2.2 Construction of being in the top tertile for time spent walking per week variable

In addition to the intensity depended WHO recommended physical activity level outcomes, time spent walking, regardless of intensity, was considered an important outcome in the present study because the literature review suggested this physical activity was particularly sensitive to neighbourhood physical environments. A variable was constructed to categorise WIIP7 participants as low, medium and high walkers in terms of hours spent on this physical activity per week: the lowest tertile ranged from 0 up to and including 3.5 hours per week, the medium from over 3.5 to 6 hours inclusive and the high from over 6 to 63 hours inclusive. A dummy variable was then constructed from this which constituted the physical activity outcome of being in the top tertile for time spent walking per week for the WIIP7 study sample. Henceforth this outcome variable is referred to as TTW.

3.3.3.2.3 Descriptive statistics for physical activity outcomes

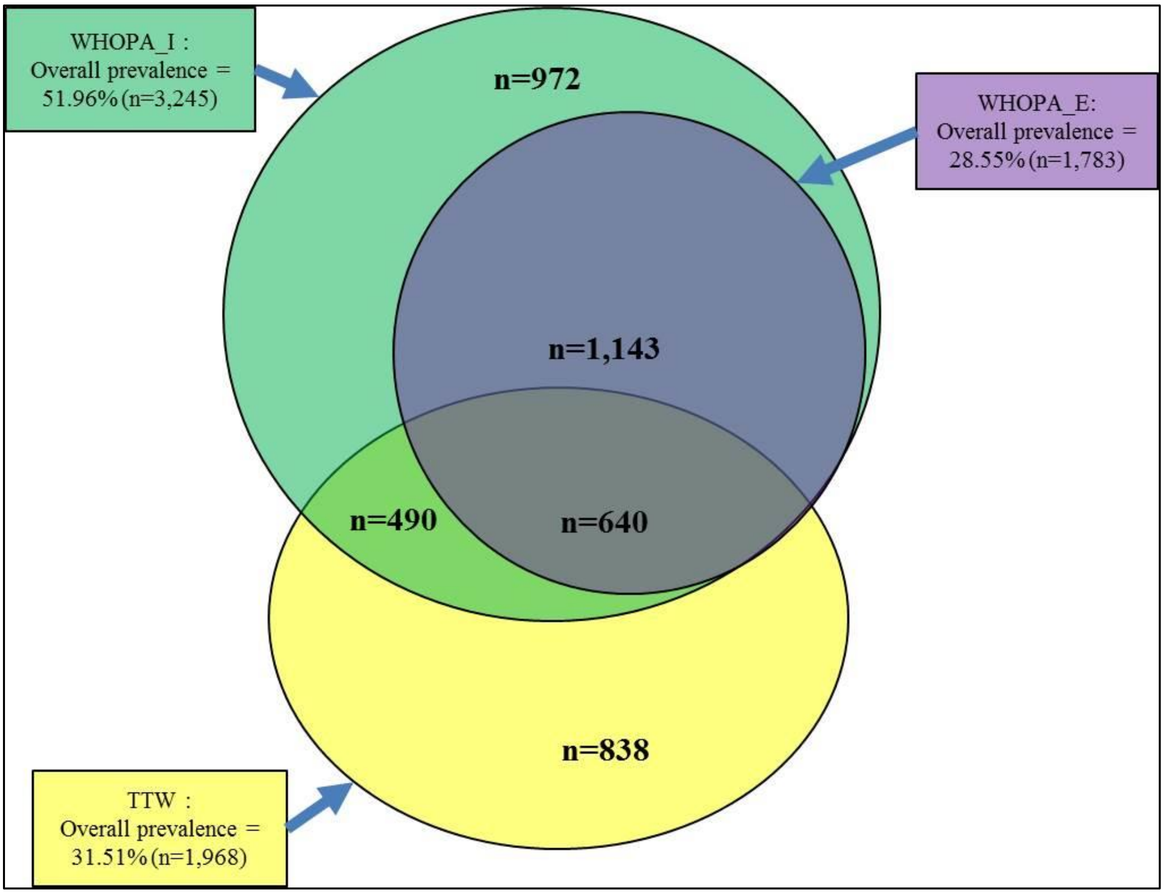
Frequencies of the three physical activity outcomes are presented in Table 3.16. Whilst less than one third of the study sample met the WHO recommended physical activity level when walking was excluded as a contributory physical activity, over half met the recommendation when walking at a brisk or fast pace was classified as a moderate intensity activity and therefore contributed to attainment of this outcome. This reflected the large amount of time spent walking relative to other physical activities in the study sample. The higher prevalence of WHOPA_I than the TTW outcome in the study sample reflected the fact that, even if an individual’s only physical activity contributing to WHOPA_I was moderate intensity-classified brisk walking, they only needed to accrue 150 minutes (2.5 hours) per week, which was less than the walking time of more than 6 hours per week necessary to “achieve” the TTW outcome. The relationships between the physical activity variable outcomes in the WIIP7 study sample are illustrated in the Venn diagram in

Figure 3.1. It shows the relative frequencies of the physical activity variable outcomes in the WIIP7 study sample including only individuals for which data was available for all 3 outcomes (N=6245).

Table 3.16 Frequencies of the physical activity outcomes in the WIIP7 study sample.

Physical activity outcome	N	Frequency (%)
WHOPA_E	6,694	1873 (28.0)
WHOPA_I	6,512	3333 (51.2)
TTW	6,409	2032 (31.7)

Figure 3.1 The relationships of the physical activity variable outcomes in the WIIP7 study sample including only individuals for which data was available for all 3 outcomes (N=6245).



3.4 Statistical analysis

In this thesis the depression and physical activity outcomes variable were dichotomous, occurring in one of two states, and all independent variables, the predictors and the correlates, were also categorical. Given the categorical form of the variables, multivariate logistic regression was selected as the most appropriate method of modelling statistical associations in examination of possible relationships with the neighbourhood physical environment exposures and the outcomes. In this type of logistic regression the outcome, the dependent variable, has a value of 0 or 1 and the relative frequencies are used to predict the probability of the outcome being 0 or 1 when the exposure, the predictor or explanatory variable, and any other independent variables, have particular values. Results, in this thesis, were computed as odds ratios, the ratio of odds of exposure in those with the outcome, such as having depression, to that of those without the outcome. Results were presented as odds ratios alongside their 95% confidence intervals to indicate with 95% confidence the range within which the true value of the odds ratio lay.

Studies 1, 2, 3 and 5 of Chapters 4, 5, 6 and 8, respectively, in this thesis examined relationships between a particular exposure variable or set of variables and each of the specified outcomes, and in each study unadjusted bivariate logistic regression models were first specified which included only the outcome variable and exposure variable of interest to determine the presence and strength of significant associations. Where significant associations were detected, adjustment was subsequently made for factors which could confound the relationship by constructing multivariate logistic regression models to gauge the extent to which the relationships between exposures and outcomes were independent of other factors. In Study 2, an examination of relationships between proportions of greenspace and domestic garden within neighbourhoods and the specified outcomes, the exposure variables were constructed as quintiles – equal ordered subgroups – representing ordered levels of exposure. Also, in the examination of relationships between neighbourhood walkability and the outcomes in Study 5, the exposures were constructed as quartile and decile scores, as detailed in Chapter 7. Thus it was useful to perform tests for trend to detect any significant dose effects of the greenspace and domestic garden and of the walkability exposures on the outcomes.

Logistic regression enabled identification of the strength and significance of associations between exposures and outcomes among the WIIP7 and HSE2008 samples, and the independence of these associations from individual level factors and area deprivation. However, this single level modelling approach was unable to show the level of variation between areas – be it LAs, CAS wards or other operationalisations of neighbourhood – attributable to unmeasured area level variables, the random effects. It was important to identify the level of unaccounted area level variation in order to determine the relative contribution of the area level exposure variables: single level logistic regression only identified the contribution of any residual, unmeasured effects regardless of whether it was operating at the level of the individual or the area. Therefore, in Study 3, multilevel modelling was employed to quantify area level variation attributable to unmeasured area level variables for which single level modelling could not account. The specifications of the multilevel models constructed are detailed in Chapter 6.

In Study 4 in which the walkability of London neighbourhoods was modelled, statistical software was used in conjunction with GIS software to construct a walkability index. Statistical analyses were then performed to describe the size of the areas, and numbers, of GIS-computed and administrative spatial units defining neighbourhoods. Correlation analyses were done to investigate the similarity of walkability scores between models and also for a given model between different spatial units of enumeration.

3.5 Summary

The present research analysed GHQ-12 depression in two large study samples, HSE2008 and WIIP7, and three physical activity outcomes – WHOPA_E, WHOPA_I and TTW – in WIIP7 in relation to several physical environmental exposure variables measured at various geographical scales. This gave rise to a large number of permutations of exposure and outcome variables which have been summarised in Table 3.17. The multiple permutations of exposure and outcome variables analysed in this thesis, together with differences in the geographical distribution of the study samples, had implications for selection of participants and for derivation of variables. Participants of WIIP7 were distributed across the United Kingdom but with a geographical focus in London whereas participants of HSE2008 were restricted to England but their geographical spread was even, owing to the study design for national representation. Selection for inclusion of the participants of the study samples on which this thesis drew for each study was based not only on the availability of data for the outcome but on that for the exposure.

Whilst a central aim of this thesis was to determine the differential effect of neighbourhood operalisations on the effect of associations between neighbourhood physical environment exposures and outcomes, the number of operationalisations possible was limited by the geographical level at which participants were identifiable and also that at which the data was collected from which exposure variables were derived. Table 3.18 summarises the geographical levels at which neighbourhoods were operationalised in the examination of associations between particular exposures and outcomes.

Table 3.17 Summary of outcome and exposure variables used in each study.

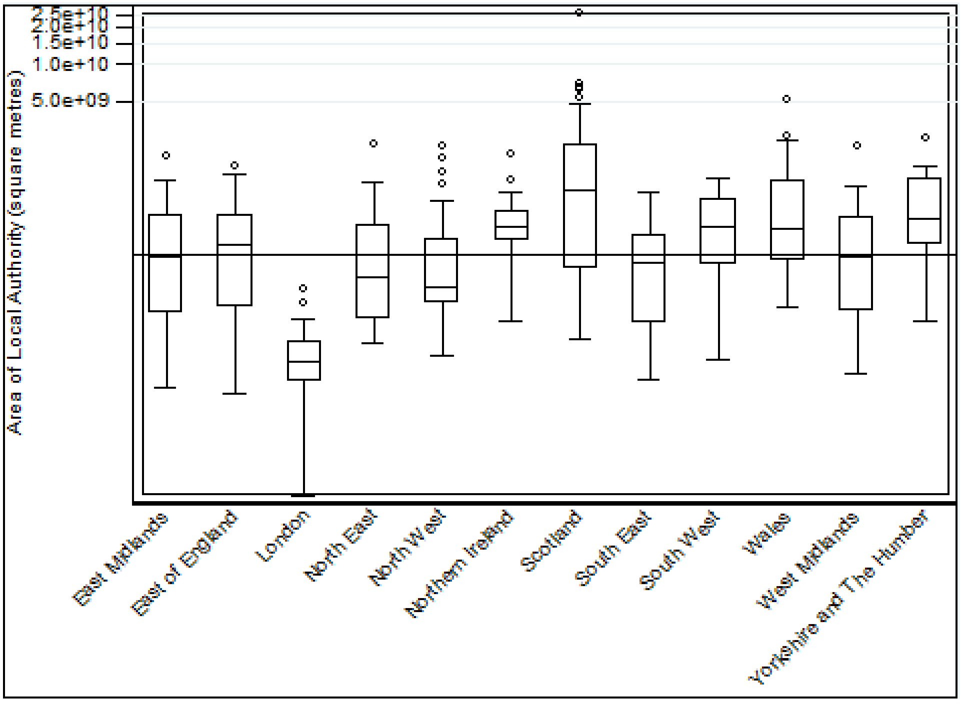
Chapter (Study)	Exposure variables	Outcome variables	Correlates
4(1A)	WIIP7: Sex; Age; Economic activity; Income, Marital status; Ethnicity; Area deprivation HSE2008: Sex; Age; Economic activity; Income; Occupational social class; Marital status; Area deprivation	GHQ-12 depression	N/A
4(1B)	Economic activity; Marital status; Ethnicity; Car availability; Area deprivation	WHOPA_E	N/A
4(1C)	WHOPA_E; WHOPA_I; TTW	GHQ-12 depression	WIIP7: Sex; Age; Economic activity; Marital status; Ethnicity; Area deprivation
5(2A)	Greenspace; Greenspace plus domestic garden	WHOPA_E; WHOPA_I; TTW	WIIP7: Sex; Age; Economic activity; Marital status; Ethnicity; Car availability; Area deprivation
5(2B)	Greenspace; Greenspace plus domestic garden	GHQ-12 depression	WIIP7: Sex; Age; Economic activity; Marital status; Ethnicity; Area deprivation HSE2008: Sex; Age; Economic activity; Income; Occupational social class; Marital status; Area deprivation
6(3A)	OEQ; PEQ	WHOPA_E; WHOPA_I; TTW	WIIP7: As for study 2A
6(3B)	OEQ; PEQ	GHQ-12 depression	WIIP7: As for study 2B HSE2008: As for study 2B
8(5A)	Walkability (As detailed in Chapter 7)	WHOPA_E; WHOPA_I; TTW	WIIP7: As for study 2A
8(5B)	Walkability (As detailed in Chapter 7)	GHQ-12 depression	WIIP7: As for study 2B HSE2008: As for study 2B

Table 3.18 Summary of the geographical levels at which neighbourhoods were operationalised in the examination of associations between particular exposures and outcomes.

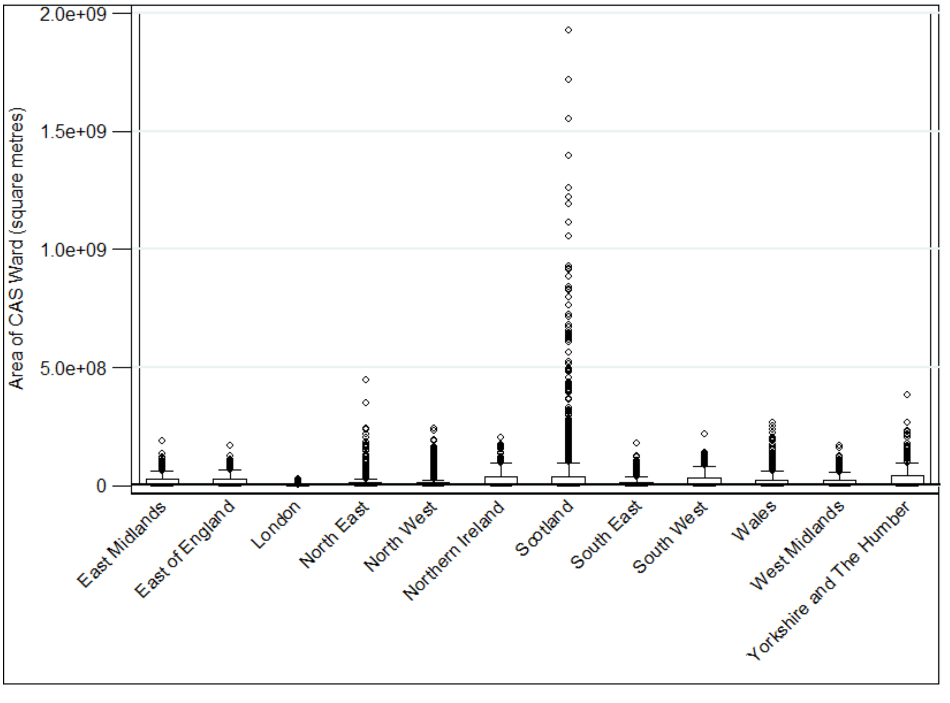
Chapter (Study)	Exposure variables	Outcome variables	Study sample	Neighbourhood operationalisation			
				LA	CAS ward	OA	GIS-computed
4(1C)	WHOPA_E; WHOPA_I; TTW	GHQ-12 depression	WIIP7	✓	✓	✓	
5(2A) 5(2B)	Greenspace; Greenspace plus domestic garden	WHOPA_E; WHOPA_I; TTW	WIIP7	✓	✓	✓	
		GHQ-12 depression	WIIP7	✓	✓	✓	
			HSE2008	✓			
6(3A) 6(3B)	OEQ	WHOPA_E; WHOPA_I;TTW	WIIP7	✓	✓		
		GHQ-12 depression	WIIP7	✓			
			HSE2008	✓			
6(3A) 6(3B)	PEQ	WHOPA_E; WHOPA_I; TTW	WIIP7	✓			
		GHQ-12 depression	WIIP7	✓			
			HSE2008	✓			
8(5A) 8(5B)	Walkability (As detailed in Chapter 7)	WHOPA_E; WHOPA_I; TTW	WIIP7	✓	✓	✓	✓
		GHQ-12 depression	WIIP7	✓	✓	✓	✓

3.6 Appendices

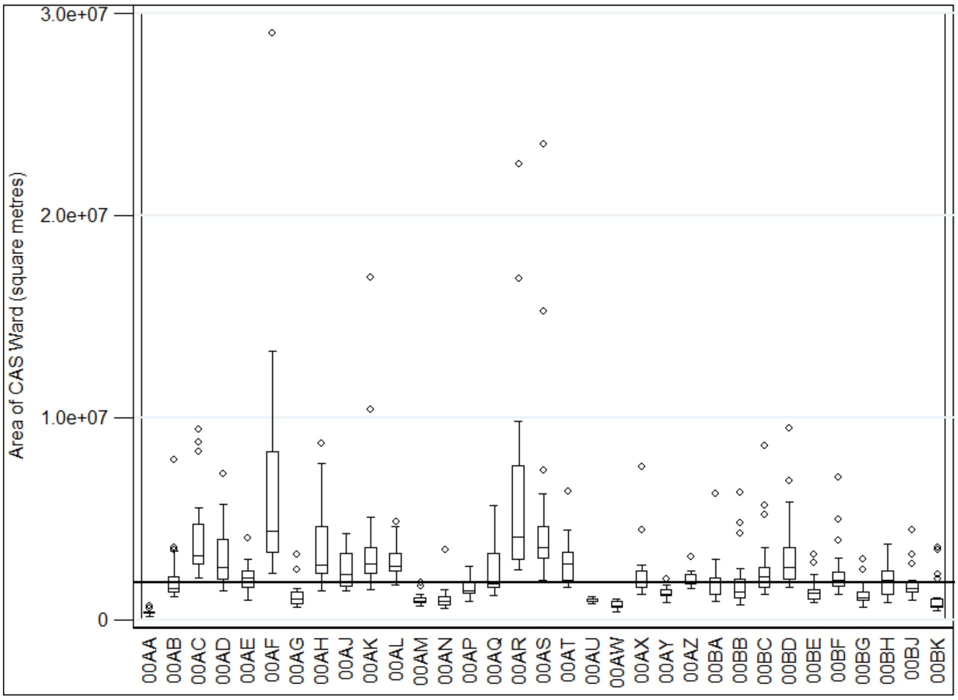
Appendix 3.1 Area of LA as a function of GOR. The black line indicates the median LA area for the UK, 288Mm^2.



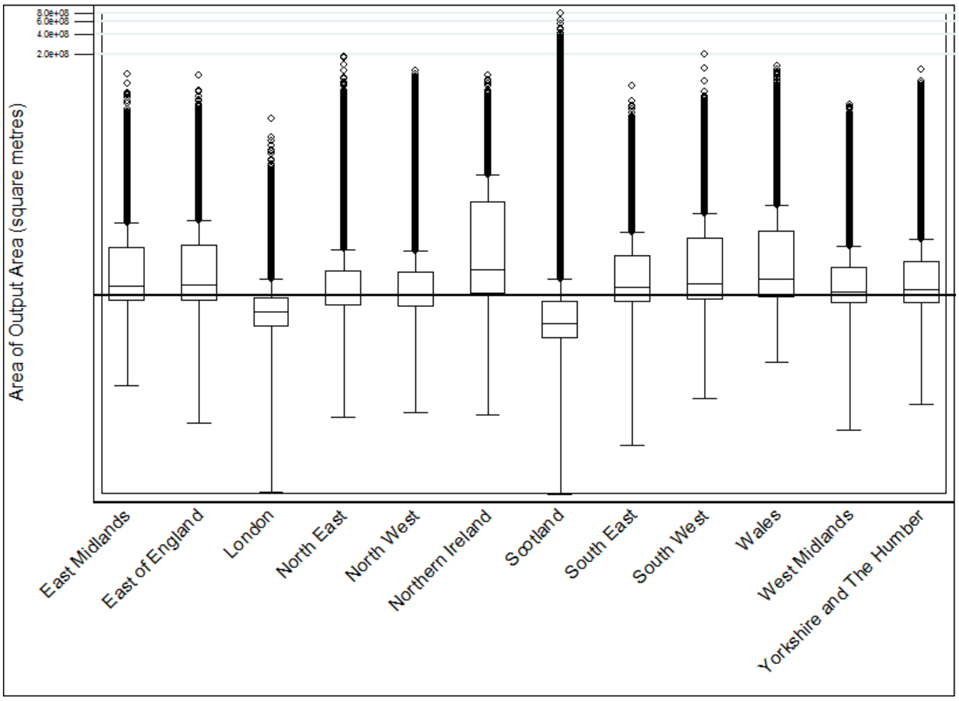
Appendix 3.2 Area of Census Area Statistics ward as a function of GOR. The black line just above the x-axis indicates the median CAS ward area for the UK, 4.9Mm^2.



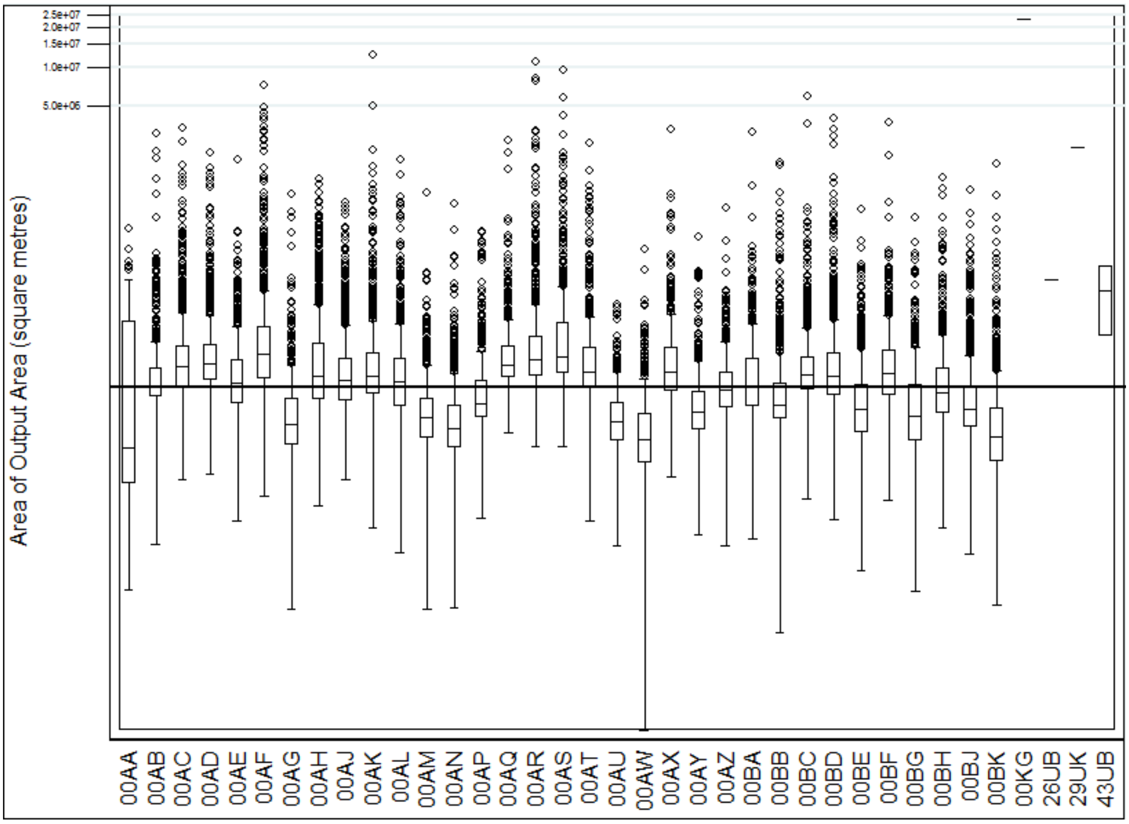
Appendix 3.3 Area of Census Area Statistics ward as a function of LA in London. The black indicates the median CAS ward area for London, 1.8Mm^2.



Appendix 3.4 Area of OA as a function of GOR. The black indicates the median area of OA for the UK, 60,812m^2.



Appendix 3.5 Area of OA as a function of LA in London. The black indicates the median area of OA for London, 34850m^2.



Appendix 3.6 The 12-item General Health Questionnaire

How has your health been in general over the past few weeks. Have you recently:

1	Been able to concentrate on whatever you're doing?	Better than usual	Same as usual	Less than usual	Much less than usual
2	Lost much sleep over worry?	Not at all	No more than usual	Rather more than usual	Much more than usual
3	Felt that you were playing a useful part in things?	More so than usual	Same as usual	Less useful than usual	Much less useful
4	Felt capable of making decisions about things?	More so than usual	Same as usual	Less so than usual	Much less capable
5	Felt constantly under strain?	Not at all	No more than usual	Rather more than usual	Much more than usual
6	Felt that you couldn't overcome your difficulties?	Not at all	No more than usual	Rather more than usual	Much more than usual
7	Been able to enjoy your normal day-to-day activities?	More so than usual	Same as usual	Less so than usual	Much less than usual
8	Been able to face up to your problems?	More so than usual	Same as usual	Less able than usual	Much less able
9	Been feeling unhappy and depressed?	Not at all	No more than usual	Rather more than usual	Much more than usual
10	Been losing confidence in yourself?	Not at all	No more than usual	Rather more than usual	Much more than usual
11	Been thinking of yourself as a worthless person?	Not at all	No more than usual	Rather more than usual	Much more than usual
12	Been feeling reasonably happy, all things considered?	More so than usual	About same as usual	Less so than usual	Much less than usual

Appendix 3.7 Physical activity section of the self-administered WIIP7 questionnaire

SECTION 4 - This section is about your lifestyle

Exercise

We would like to know about your activities at work and in your free time that involve physical activity.

Q81. Thinking about the days of the **PAST WEEK.**

a. On average, for how many minutes did you **walk** outside your home/workplace?

on each weekday

Minutes ☐☐

on each weekend day

Minutes ☐☐

b. On average, for how many minutes did you cycle?

on each weekday

Minutes ☐☐

on each weekend day

Minutes ☐☐

Q82. Other physical activities in the **PAST FOUR WEEKS**

Please indicate the number of **occasions** and **total** time spent on each of the activities listed. *Write in other types of activity not listed, as applicable.*

a. SPORTS AND GAMES

Football (including coaching etc.)

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Golf

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Swimming

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Other activities, e.g. aerobics, ballroom dancing, keep fit, jogging, tennis
Other activity one (*please specify*)

1. ________________

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Other activity two (*please specify*)

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

GARDENING

Weeding, hoeing, pruning (not mowing)

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Manual lawn mowing

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Other gardening e.g. digging, planting, clearing ground, etc (*please specify*)

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

c. HOUSEWORK

Carry heavy shopping

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Cooking

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Hanging out washing

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Other housework e.g. dusting, ironing, hoovering

Other housework, activity one (*please specify*)

1.

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
0	1	2	3	4	5	6

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
0	1	2	3	4	5	6

Other housework, activity two *(please specify)*

1. [ ]

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

d. DO-IT-YOURSELF

Manual car washing

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

Painting/decorating

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

Other DIY e.g. household repairs, woodwork, bricklaying *(please specify)*

1. [ ]

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

e. ADDITIONAL/OTHER *(please specify)*

1. [ ]

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

2. [ ]

Occasions in the past 4 weeks *(Please tick one)*

None	1-2	3-4	5-10	11-15	16-20	21+
☐	☐	☐	☐	☐	☐	☐

Total hours in the past 4 weeks *(Please tick one)*

None	½	1-1½	2-3	4-5	6-10	11+
☐	☐	☐	☐	☐	☐	☐

Q83. How many times a week do you engage in vigorous physical activity enough to make you out of breath, and for how long in total? *(Please specify the activity)*

[ ]

Occasions per week *(Please tick one)*

None	1	2	3	4	5	6+
☐	☐	☐	☐	☐	☐	☐

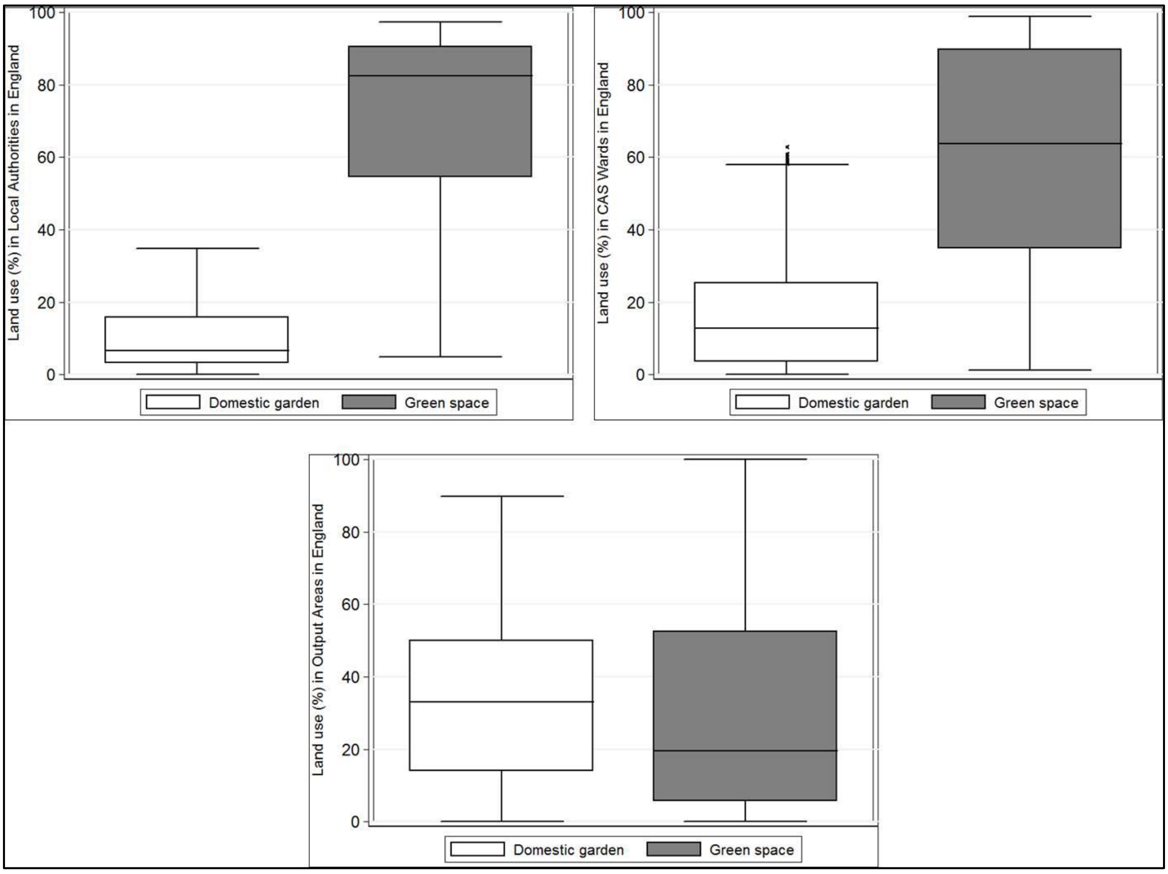
Total hours per week *(Please tick one)*

None	½	1	1½	2	2½	3+
☐	☐	☐	☐	☐	☐	☐

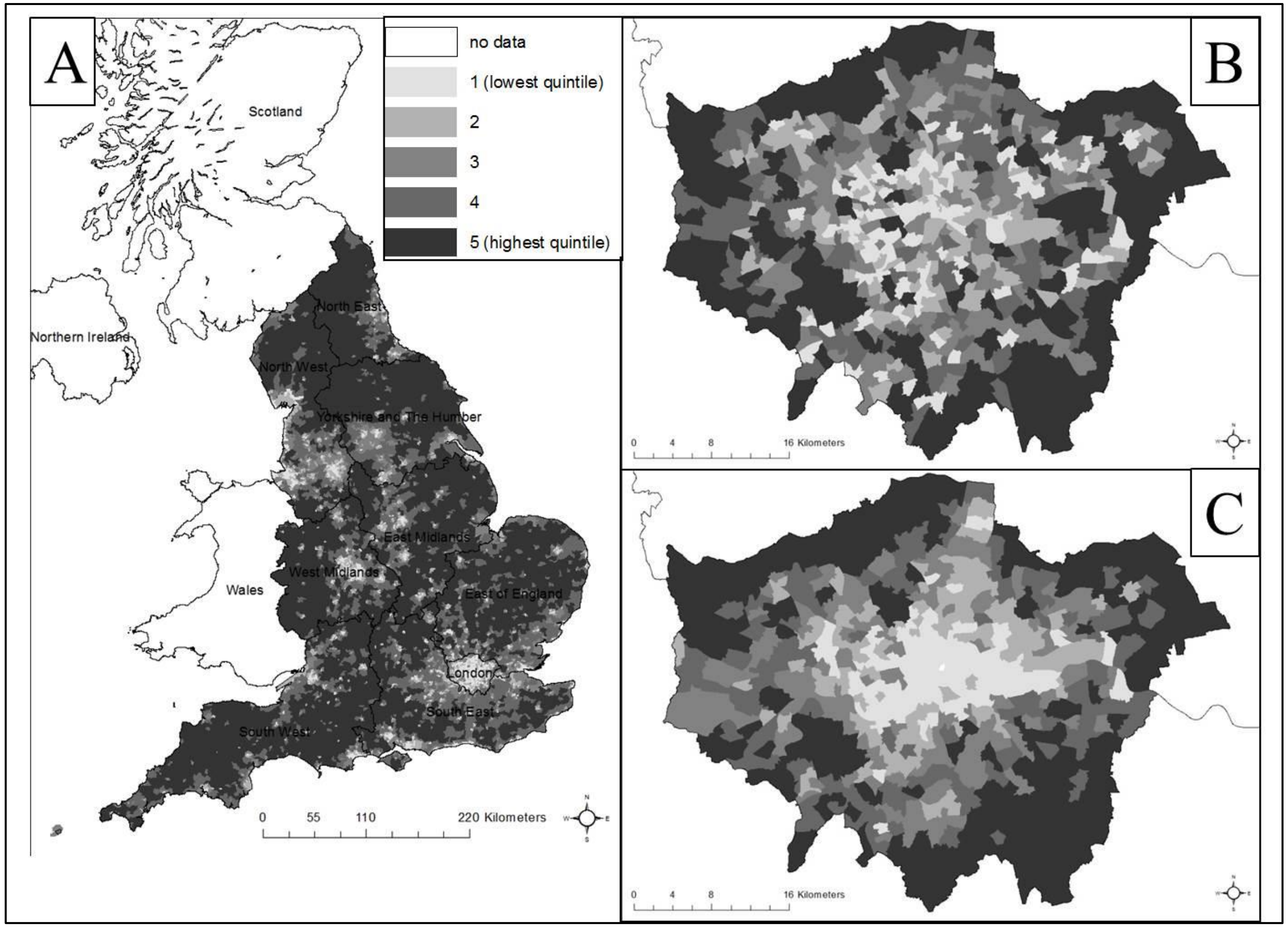
Q84. How would you describe your usual walking pace? *Please tick one box only.*

- Slow pace (i.e. less than 3 mph) ☐
- Steady average pace ☐
- Brisk pace ☐
- Fast pace (i.e. over 4 mph) ☐

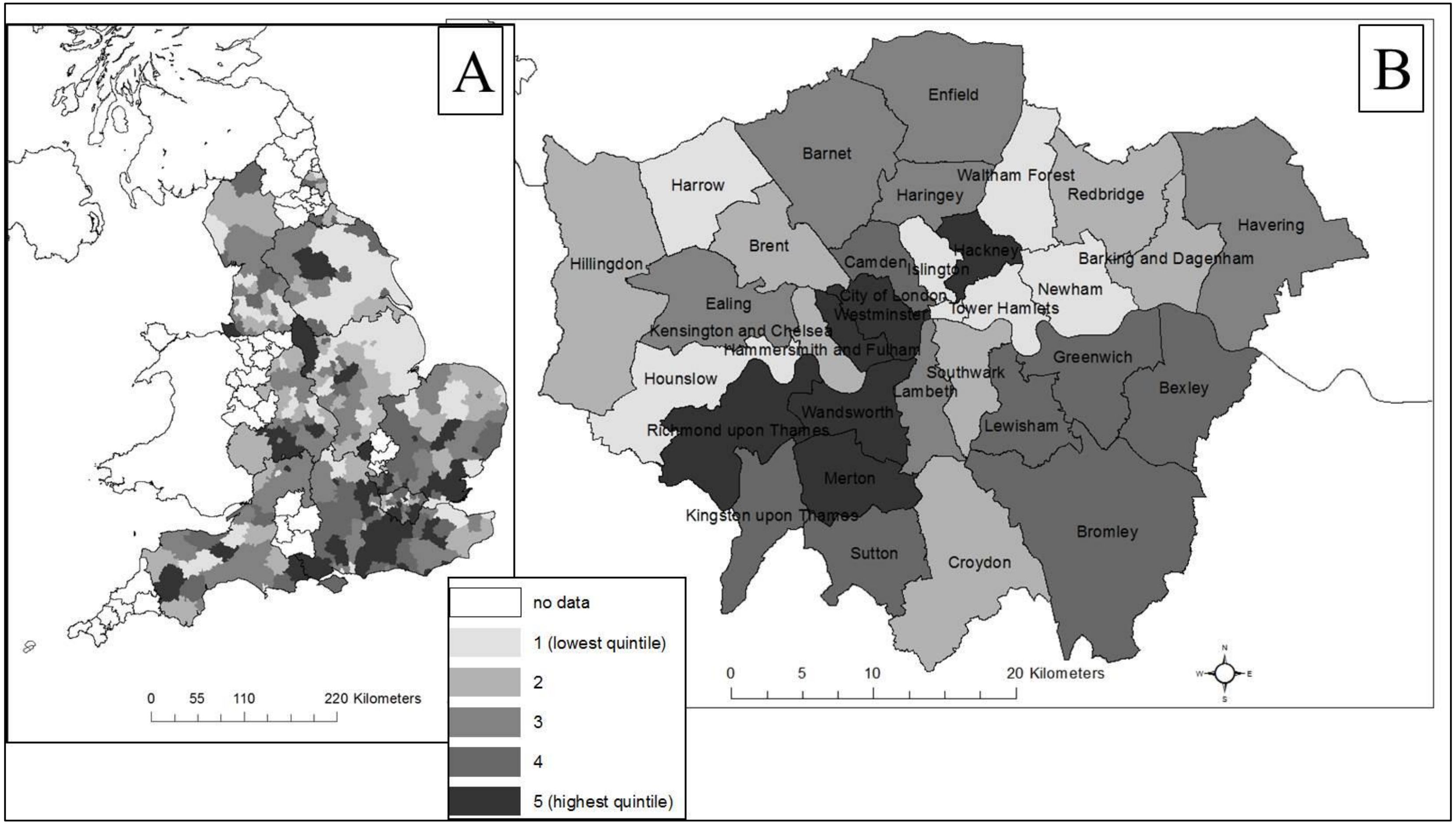
Appendix 3.8 Box plots illustrating variation in percentage domestic garden and greenspace as recorded in GLUD within the administrative units of LAs (top left), CAS wards (top right) and OAs (bottom) across England.



Appendix 3.9 Quintiles of percentage of land that is (A) greenspace in CAS wards across England; (B) greenspace in wards derived in and specifically for London; and (C) either greenspace or domestic garden in CAS wards derived in and specifically for London.



Appendix 3.10 Quintiles derived from English (A) and London (B) local authorities of percentage of Place Survey respondents reporting satisfaction with parks and open spaces in their local authority. Darker shades indicate higher levels of satisfaction.



4 Study 1: Associations between non-physical environmental factors and physical activity, and GHQ-12 depression outcomes

4.1 Introduction

4.1.1 Chapter overview

This chapter serves to clarify associations between non-physical environmental factors and physical activity, and GHQ-12 depression outcomes, enabling better modelling and understanding of associations of physical environmental factors and these outcomes.

Appropriate modelling of statistical association between an exposure and the hypothesised outcome involves adjustment for confounding factors which may influence the association. Confounders are factors which are associated with both the exposure and the outcome but do not lie on the causal pathway [195]. Therefore, if a variable is not associated with the outcome it cannot be a confounder, even if associated with the exposure. Several factors were identified in Chapter 2 as putative confounders of relationships between the neighbourhood physical environment exposure variables and the GHQ-12 depression outcome, and the physical activity outcomes, and data was collected for many of these factors in the study samples. However, it was necessary to establish whether these putative confounders were indeed associated with the outcomes of interest in the study samples, as suggested by the literature, before adjustment was made for them in the final statistical models because unnecessary adjustment would reduce statistical efficiency and potentially mask significant associations of the neighbourhood physical environment exposures with the outcomes. Studies 1A and 1B in this chapter addressed this issue with respect to the GHQ-12 depression and physical activity outcomes respectively. The results of analysis in Studies 1A and 1B informed the construction of statistical models of associations between physical environmental exposures and outcomes, with adjustment made for those factors which remained significantly associated with the outcome in the presence of the other variables.

Like Studies 1A and 1B, Study 1C was undertaken for clarification but had a distinct purpose. Study 1C was an examination of relationships between physical activity and depression, with statistical associations explored with respect to the influence of confounding factors identified in Studies 1A and 1 B, and physical activity modelled as a non-physical environmental exposure. This examination was necessary in this thesis in order to determine the potential for the hypothesised mediatory role for physical activity in relationships between physical environments and depression. Whilst Study 1C's cross-sectional design precluded causal inferences, based on the reviewed literature, a significant independent association would likely be bi-directional, with people more likely to be depressed as a result of lower physical activity levels as well as doing less physical activity as a consequence of being depressed.

4.1.2 Research questions

Study 1 addresses the following research questions:

- Which factors identified in the literature review as putative confounders of relationships between the neighbourhood physical environment exposure variables and depression are associated with a score of four or more on the 12-item General Health Questionnaire (GHQ-12 depression) in the

Whitehall II Study Phase 7 (WIIP7) and the 2008 Health Survey for England (HSE2008) study samples? (Study 1A)

- Which factors identified in the literature review as putative confounders of relationships between the neighbourhood physical environment exposure variables and physical activity are associated with meeting the World Health Organisation recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E) in the WIIP7 study sample? (Study 1B)
- Is physical activity associated with GHQ-12 depression in the WIIP7study sample? (Study 1C)
- Are associations between physical activity associated and GHQ-12 depression independent of individual-level sociodemographic factors and of area deprivation? (Study 1C)

4.1.3 Objective for Study 1A

The objectives are as follows:

- to determine the statistical significances of associations of individual level sociodemographic factors identified in the literature review as putative correlates of depression – sex, age, economic activity, income, occupational social class, marital status and ethnicity – with GHQ-12 depression in the WIIP7 and the HSE2008 study samples
- to determine the statistical significance of the association of area level deprivation (identified in the literature review as a putative correlate of depression) with GHQ-12 depression, in the WIIP7 and the HSE2008 study samples

4.1.4 Hypotheses for Study 1A

Individual level sociodemographic factors identified in the literature review as putative correlates of depression (sex, age, economic activity, income, occupational social class, marital status and ethnicity) are significantly associated with GHQ-12 depression in the WIIP7 and the HSE2008 study samples

Area level deprivation, identified in the literature review as a putative correlate of depression, is significantly associated with GHQ-12 depression in the WIIP7 and the HSE2008 study samples

4.1.5 Objective for Study 1B

The objectives are as follows:

- to determine the statistical significances of associations of individual level sociodemographic factors identified in the literature review as putative correlates of physical activity – sex, age, economic activity, car availability, marital status and ethnicity – with WHOPA_E in the WIIP7 study sample

- to determine the statistical significance of the association of area level deprivation (identified in the literature review as a putative correlate of physical activity) with WHOPA_E in the WIIP7 study sample

4.1.6 Hypotheses for Study 1B

Individual level sociodemographic factors identified in the literature review as putative correlates of physical activity (sex, age, economic activity, car availability, marital status and ethnicity) are significantly associated with WHOPA_E in the WIIP7 study sample

Areas level deprivation, identified in the literature review as a putative correlate of physical activity is significantly associated with WHOPA_E in the WIIP7 study sample

4.1.7 Objectives for Study 1C

The objectives are as follows:

- to determine whether odds of GHQ-12 depression are lower among those meeting the World Health Organisation recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E)
- to determine whether odds of GHQ-12 depression are lower among those meeting the World Health Organisation recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I)
- to determine whether odds of GHQ-12 depression are lower among those in the top tertile of the WIIP7 sample for time spent walking per week (TTW)
- to determine whether the associations are affected by individual-level sociodemographic factors and by area deprivation

4.1.8 Hypotheses for Study 1C

GHQ-12 depression is negatively associated with the physical activity outcomes of WHOPA_E, WHOPA_I and TTW.

The strength of associations is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and for area deprivation.

4.2 Methods

4.2.1 Study samples

Participants in Study 1A comprised the WIIP7 study sample (see 3.2.1.2) and, in separate analyses, the HSE2008 study sample (see 3.2.2.2). Participants in Study 1B and 1C comprised only the WIIP7 study sample. The WIIP7 study sample for studies 1A, 1B and 1C was drawn from the seventh wave, conducted in 2004/5, of the Whitehall II study, a longitudinal study of civil servants to examine the social determinants of health [182]. In contrast, the HSE2008 study sample for Study 1A was from the Health Survey for England conducted in 2008 which drew a nationally representative general population sample of adults living in households in England through multi-stage stratified probability sampling [183]. In the WIIP7 study sample of 6885 individuals, the mean age was 61.22 ± 6.00 years, ranging from 50.47 to 74.08 years, 29.77% were women. Among the 14,221 individuals in HSE2008, nationally representative of adults in England, the mean age was 48.94 ± 18.62 years, ranging from 16 to 97, and 55.48% were women.

4.2.2 Variables

4.2.2.1 Exposure variables

Based on a review of the literature, the following factors were considered as possible individual level confounders in the relationship between GHQ-12 depression and the environmental exposures: sex, age, economic activity, income, occupational social class, marital status and ethnicity. Also, a contextual factor, area level deprivation, was considered as a possible confounder of the relationship between environmental exposures and GHQ-12 depression because it has been associated with mental health outcomes in a Whitehall II study sample, with those living in poorer CAS ward-defined neighbourhoods more likely to have poor mental health [196].

The individual level sociodemographic exposure variables for the WIIP7 study sample in Study 1A were as follows: sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), income (<£20000 (ref)/ £20000 to <£40000/ £40000 to <£60000/ >=£60000), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white). For the HSE2008 study sample in Study 1A, the individual level sociodemographic exposure variables were are follows: sex (Male (ref)/ Female), age (16y-34y (ref)/ 35y-54y/ 55y+), economic activity (Employed (ref)/ Unemployed/ Retired/ Other), income (<=£14918 (ref)/ >£14918-£31871/>£31871), occupational social class (Managerial and professional (ref)/ Intermediate/ Routine and manual/ Other) and marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed). Although an occupational social classification systems variable existed in the WIIP7 dataset it was not comparable to the National Statistics Socio-economic Classification-derived variable in the HSE2008 dataset as all WIIP7 participants were office-based workers at recruitment into the study. Therefore, occupational social class was excluded as an exposure variable in analyses for the WIIP7 study sample. Also, ethnicity was absent as a variable in the dataset constructed for the HSE2008 study sample so it was excluded as an exposure for the HSE2008 study sample. The contextual factor, area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level [197], was an additional exposure variable for both the WIIP7 and HSE2008 study samples in Study 1A. This variable was already present in the HSE2008 dataset but it was necessary to attach it via participants' LSOA codes in the WIIP7 dataset. As an England only based measure, adjustment for area deprivation in the WIIP7 sample was restricted to the 6677 (97.0%) participants resident in England. IMD2004 captures area-level characteristics in terms of income, employment, health, education, barriers to housing and services, living environment, and crime.

For Study 1B the individual level sociodemographic factors considered, based on reviewed literature, as possible confounders in the relationship between physical activity and the environmental exposures were sex, age, economic activity, car availability, marital status and ethnicity. Specifically the exposure variables for the WIIP7 study sample were as follows: sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), car availability (Car available (ref)/No car available), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white). Area level deprivation was also considered as a possible confounder of the relationship between environmental exposures and the physical activity outcomes. In a study conducted in New Zealand, Annear *et al* found self-reported leisure-time physical activity was associated with area deprivation in the neighbourhoods, identified as a point on a map delineated with a 1000 metre radius, of older individuals [198], independent of individual level factors. In contrast another study, carried out in the United Kingdom, in which physical activity was objectively measured among older participants, and in which area deprivation was measured within the Lower Super Output Area (LSOA) of their medical practice, found that associations were eliminated after adjustment for individual level factors [199]. Again, operationalizing neighbourhoods as LSOA, Watts *et al* employed multilevel modelling to examine relationships between neighbourhood level and individual level characteristics and physical activity [200]. Their study, conducted in London, showed that physical activity was associated with distance to nearest greenspace and their modelling suggested that other measured neighbourhood-level characteristics did not account for the residual variance in the outcome between neighbourhoods. However, it was necessary to consider area deprivation as a contextual confounder in order to make the present study comparable with that of Mytton *et al*, who found associations between physical activity among participants of the Health Survey for England 2008 and greenspace, measured within LSOAs, that were independent of both individual level factors and area deprivation [201]. Therefore, an additional exposure variable for the WIIP7 study sample in Study 1B was England-based quintiles of the IMD2004 at LSOA level which had been attached via participants' LSOA codes for study 1A.

The exposure variables for Study 1C for the WIIP7 study sample were as follows: meeting the World Health Organisation recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E); meeting the World Health Organisation recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I); and being in the top tertile for time spent walking per week (TTW). The derivation of these variables is detailed in Section 3.3.3.2.2. The WHO recommended physical activity level was defined as follows:

- at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or at least 75 minutes of vigorous intensity aerobic physical activity throughout the week, or an equivalent combination of moderate and vigorous-intensity activity

4.2.2.2 Outcome variables

For Studies 1A and 1C, the outcome variable was GHQ-12 depression, defined as a score of four or more on the General Health Questionnaire (GHQ-12) [57] as detailed in 3.3.3.1.3.

Although three physical activity outcomes were specified in this thesis – WHOPA_E, WHOPA_I and TTW – for consistency, the selection of the correlates was informed by evaluation of statistical associations with only one of the outcomes, WHOPA_E. Initial analyses revealed variation between the physical activity outcomes in the strength and directions of associations with putative confounders and it was deemed necessary to adjust for a consistent core set of correlates. Therefore, for Study 1B the only outcome variable specified was WHOPA_E. However, univariate associations, and multivariate associations, of WHOPA_I with individual level factors in the WIIP7 study sample are given in Appendix 4.1 and Appendix 4.2, respectively. Also, univariate associations, and multivariate associations, of TTW with individual level factors in the WIIP7 study sample are given in Appendix 4.3 and Appendix 4.4, respectively.

4.2.2.3 Correlates

The results of analysis in Studies 1A and 1B informed the construction of statistical models in Study 1C. Individual level correlates for which adjustment was made in statistical models in Study 1C for the WIIP7 study sample were sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white). Adjustment was also made for the contextual factor, area deprivation, as England-based quintiles of the IMD2004 at LSOA level.

4.2.3 Statistical analysis of associations between physical activity and GHQ-12 depression

In Study 1A and 1B, contingency tables for each putative confounder were constructed and categories were appropriately collapsed where cells with low counts were identified. Bivariate logistic regression was then performed to detect significant associations between the putative confounders and each of the outcomes. Subsequently, multivariate regression was performed wherein variables identified as being significantly associated with the outcome in univariate analysis were regressed against the outcome variable to show the extent to which relationships were independent of one another. In both Studies 1A and 1B, multivariate regression was performed only for the WIIP7 study sample and only for the sociodemographic factors. This was because it was considered necessary for simplicity and consistency to identify only a core set of correlates rather that unique ones for each study sample for every association modelled between physical environmental exposures and outcomes in subsequent studies.

In Study 1C modelling of statistical associations in examination of possible relationships physical activity and the GHQ-12 depression employed multivariate logistic regression for reasons outlined in 3.3.3.1.3. In the first instance, bivariate logistic regression models were specified which included only the GHQ-12 outcome variable and the physical activity variables as exposures to determine the presence and strength of significant associations. Subsequently, for models in which significant associations were detected, adjustment was made for correlates – firstly individual level sociodemographic factors and then additionally area deprivation – which could confound the relationship. This showed the extent to which the relationships between the physical activity exposures and the GHQ-12 depression outcome were independent of other factors.

In Studies 1A, 1B and 1C results were computed as odds ratios alongside their 95% confidence intervals to indicate with 95% confidence the range within which the true value of the odds ratio lay.

4.3 Results

4.3.1 Associations of socio-demographic factors and area-level deprivation with GHQ-12 depression in WIIP7 and HSE2008 (Study 1A)

4.3.1.1 Univariate associations of GHQ-12 depression with individual level sociodemographic factors

In the WIIP7 sample, univariate analysis revealed all exposure variables except income to be significantly associated with GHQ-12 depression (Table 4.1). Women were significantly more likely to have the GHQ-12 depression outcome than men (OR=1.39, p<0.001). Older age groups were significantly increasingly less likely to be depressed than the reference age group of 50 years to less than 56 years. Relative to those remaining employed in the Civil Service, those working outside of it (OR=0.63, p<0.001) and those who were retired (OR=0.55, p<0.001) were significantly less likely to be depressed, whilst the long-term sick economically inactive were over threefold more likely to be depressed. Relative to married individuals, single people were more likely to be depressed (OR=1.39, p<0.001) as were divorced people (OR=1.59, p<0.001). Non-white people in WIIP7 were also more likely to be depressed (OR=1.52, p<0.001).

Univariate analysis in the HSE2008 study sample showed sex, economic activity, income and marital status to be significantly associated with GHQ-12 depression (Table 4.2). Women had significantly higher odds of GHQ-12 depression than men (OR=1.42, p<0.001). Those in older ages groups had slightly higher odds of GHQ-12 depression than the reference age group of 16 to 34-year olds although only significantly so in the 35 to 54 year age group (OR=1.13, p<0.05). Relative to those in employment, retired people had significantly higher odds of GHQ-12 depression (OR=1.46, p<0.001). However, those who were unemployed had twice as high odds of depression and those otherwise economically inactive were over threefold higher odds. Those in the highest income bracket had significantly lower odds than those in the lowest (OR=0.41, p<0.001). HSE2008 participants whose occupations were classified as lower grade than the reference category of managerial and professional, had significantly higher odds of GHQ-12 depression: the odds ratio for intermediate occupations was 1.24 (p<0.01) whilst that for routine and manual occupations was 1.55 (p<0.001). All unmarried people had significantly higher odds in the HSE2008 study sample, particularly those classified as divorced (OR=2.45, p<0.001).

Table 4.1 Univariate associations of GHQ12 depression with individual level sociodemographic factors in the WIIP7 study sample

Individual level socio-demographic factor		N	%	No	Yes	OR	CI	p
Sex	Male (ref)	4,679	70.6	4057	622	1		
	Female	1,947	29.4	1605	342	1.39	1.20-1.61	<0.001
Age group	50y to <56y (ref)	1,624	24.5	1306	318	1		
	>=56y to <60y	1,555	23.5	1310	245	0.77	0.64-0.92	0.005
	>=60y to <66y	1,692	25.5	1486	206	0.57	0.47-0.69	<0.001
	>=66y to 75y	1,755	26.5	1560	195	0.51	0.42-0.62	<0.001
Economic activity	Remaining in Civil Service (ref)	1,919	29.0	1563	356	1		
	Working outside Civil Service	1,367	20.6	1196	171	0.63	0.52-0.76	<0.001
	Not working - retired	2,965	44.8	2635	330	0.55	0.47-0.65	<0.001
	Out of work	267	4.0	208	59	1.25	0.91-1.70	0.167
	Not working-lt sick	108	1.6	60	48	3.51	2.36-5.22	<0.001
Income	<£20000 (ref)	1,211	23.2	1030	181	1		
	£20000 - <£40000	1,989	38.2	1710	279	0.93	0.76-1.14	0.472
	£40000 - <£60000	1,212	23.3	1027	185	1.03	0.82-1.28	0.827
	>=£60000	800	15.4	701	99	0.80	0.62-1.05	0.104
Marital status	Married (ref)	4,779	72.3	4144	635	1		
	Cohabit	275	4.2	230	45	1.28	0.921.78	0.147
	Single	1,011	15.3	834	177	1.39	1.151.66	<0.001
	Divorced	469	7.1	377	92	1.59	1.252.03	<0.001
	Widowed	77	1.2	62	15	1.58	0.892.79	0.116
Ethnicity	White (ref)	6,099	92.1	5,240	859	1		
	Non-white	527	8.0	422	105	1.52	1.211.90	<0.001

Table 4.2 Univariate associations of GHQ12 depression with individual level sociodemographic factors in the HSE2008 study sample.

Individual level socio-demographic factor		N	%	No	Yes	OR	CI	p
Sex	Male (ref)	6,331	44.5	5,608	723	1		
	Female	7,890	55.5	6,666	1,224	1.42	1.29-1.57	<0.001
Age group	16y-34y (ref)	3,634	25.6	3,159	475	1		
	35y-54y	4,922	34.6	4,205	717	1.13	1.00-1.28	0.048
	55y+	5,665	39.8	4,910	755	1.02	0.90-1.16	0.722
Economic activity	Employed (ref)	7,838	55.2	7,062	776	1		
	Unemployed	634	4.5	519	115	2.02	1.63-2.50	<0.001
	Retired	3,564	25.1	3,070	494	1.46	1.30-1.65	<0.001
	Other	2,171	15.3	1,610	561	3.17	2.81-3.58	<0.001
Income	<=£14918 (ref)	3,477	30.2	2,786	691	1		
	>£14918-£31871	3,906	33.9	3,408	498	0.59	0.52-0.67	<0.001
	>£31871	4,128	35.9	3,745	383	0.41	0.36-0.47	<0.001
Occupational social class	Managerial and professional (ref)	4,696	33.0	4,187	509	1		
	Intermediate	3,113	21.9	2,704	409	1.24	1.08-1.43	0.002
	Routine and manual	5,640	39.7	4,744	896	1.55	1.38-1.75	<0.001
	Other	762	5.4	631	131	1.71	1.39-2.11	<0.001
Marital status	Married (ref)	7,529	52.9	6,686	843	1		
	Cohabit	1,602	11.3	1,374	228	1.32	1.12-1.54	0.001
	Single	2,690	18.9	2,298	392	1.35	1.19-1.54	<0.001
	Divorced	1,285	9.0	982	303	2.45	2.11-2.84	<0.001
	Widowed	1,115	7.8	934	181	1.54	1.29-1.83	<0.001

4.3.1.2 Multivariate association of GHQ-12 depression with individual level sociodemographic factors in the WIIP7 study sample

In multivariate regression of the factors found to be significantly associated with GHQ-12 depression in the WIIP7 study sample, all remained significantly associated with this outcome (Table 4.3). Females had higher odds of GHQ-12 depression than males. Odds of GHQ-12 depression were also higher for those who were not married and for those whose ethnicity was non-white. However, older individuals had lower odds of GHQ-12 depression. Higher odds of GHQ-12 depression were found for those out of work, or not working due to long-term sickness, than for those remaining in the Civil Service. Conversely, odds of GHQ-12 depression for WIIP7 participants working outside the Civil Service, and for those who were retired, were lower than for those still in Civil Service employment.

Table 4.3 Multivariate association of GHQ-12 depression with individual level sociodemographic factors in the WIIP7 study sample

Individual level socio-demographic factor		OR	CI	p
Sex	Male (ref)	1		
	Female	1.26	1.08-1.47	0.003
Age group	50y to <56y (ref)	1		
	>=56y to <60y	0.81	0.67-0.97	0.026
	>=60y to <66y	0.69	0.55-0.87	0.002
	>=66y to 75y	0.67	0.51-0.87	0.003
Economic activity	Remaining in Civil Service (ref)	1		
	Working outside Civil Service	0.74	0.60-0.91	0.005
	Not working - retired	0.69	0.55-0.87	0.002
	Out of work	1.37	0.99-1.90	0.056
	Not working-lt sick	3.54	2.36-5.32	<0.001
Marital status	Married (ref)	1		
	Cohabit	1.11	0.79-1.56	0.532
	Single	1.25	1.03-1.51	0.024
	Divorced	1.44	1.12-1.86	0.005
	Widowed	1.56	0.86-2.80	0.141
Ethnicity	White (ref)	1		
	Non-white	1.48	1.17-1.88	0.001
_cons		0.17	0.13-0.21	<0.001

4.3.1.3 Univariate associations of GHQ-12 depression with area level deprivation

In both the WIIP7 and HSE2008 study samples, univariate analysis revealed higher deprivation to be significantly positively associated with GHQ-12 depression. As the WIIP7 study sample was not nationally representative, participants were not evenly distributed across the quintiles (Table 4.4). However, a significant associations with GHQ-12 depression were still detected, with WIIP7 participants in the most deprived LSOAs having higher odds of GHQ-12 depression than those in the least (OR=1.59, p=0.001). In the HSE2008 study sample, participants in the most deprived LSOAs had over twice the odd of GHQ-12 depression as those in the least (OR=2.08, p<0.001) (Table 4.5).

Table 4.4 Univariate associations of GHQ-12 depression with area level deprivation in the WIIP7 study sample.

Area level deprivation	N	%	No	Yes	OR	CI	p
Quintile 1 (least deprived) (ref)	2,621	40.8	2,278	343	1		
Quintile 2	1,465	22.8	1,268	197	1.03	0.85-1.25	0.744
Quintile 3	1,143	17.8	967	176	1.21	0.99-1.47	0.059

Area level deprivation	N	%	No	Yes	OR	CI	p
Quintile 4	824	12.8	682	142	1.38	1.12-1.71	0.003
Quintile 5 (most deprived)	372	5.8	300	72	1.59	1.20-2.11	0.001

Table 4.5 Univariate associations of GHQ-12 depression with area level deprivation in the HSE2008 study sample.

Area level deprivation	N	%	No	Yes	OR	CI	p
Quintile 1 (least deprived) (ref)	3,143	22.1	2,815	328	1		
Quintile 2	2,904	20.42	2,568	336	1.12	0.96-1.32	0.159
Quintile 3	2,862	20.13	2,479	383	1.33	1.13-1.55	<0.001
Quintile 4	2,752	19.35	2,351	401	1.46	1.25-1.71	<0.001
Quintile 5 (most deprived)	2,560	18	2,061	499	2.08	1.79-2.42	<0.001

4.3.2 Associations of socio-demographic factors and area-level deprivation with physical activity outcomes in WIIP7 (Study 1B)

4.3.2.1 Univariate associations of WHOPA_E with individual level factors

Univariate analysis in the WIIP7 study sample revealed all exposure variables were significantly associated with WHOPA_E (Table 4.6). Women had significantly lower odds of achieving WHOPA_E (OR=0.77, p<0.001). Older age groups were significantly increasingly more likely to achieve WHOPA_E than the reference age group of 50 years to less than 56 years. Relative to those remaining employed in the Civil Service, individuals who were economically inactive through retirement had over twice the odds of the WHO recommended physical activity level exclusive of walking as a contributory activity (WHOPA_E) (OR=2.01, p<0.001). Those economically inactive through unemployment were also more likely to achieve the WHOPA_E outcome (OR=1.98, p<0.001). Relative to those employed as civil servants, individuals who had left the Civil Service for employment elsewhere had significantly higher odds of achieving WHOPA_E (OR=1.47, p<0.001). Those in the WIIP7 study sample with no access to a car had significantly lower odds of achieving the WHOPA_E outcome (OR=0.54, p<0.001) although exploratory statistical analysis revealed they had significantly higher odds of being in the top tertile for time spent walking per week (TTW) (OR=1.76, p<0.001).

Table 4.6 Univariate associations of WHOPA_E with individual level factors in the WIIP7 study sample

Individual level socio-demographic factor		N	%	No	Yes	OR	CI	p
Sex	Male (ref)	4,722	70.54	3,328	1394	1		
	Female	1,972	29.46	1,493	479	0.77	0.68-0.86	<0.001
Age group	50y to <56y (ref)	1,638	24.47	1,279	359	1		
	>=56y to <60y	1,569	23.44	1,162	407	1.25	1.06-1.47	0.008
	>=60y to <66y	1,705	25.47	1,177	528	1.60	1.37-1.87	<0.001
	>=66y to 75y	1,782	26.62	1,203	579	1.71	1.47-2.00	<0.001
Economic activity	Remaining in Civil Service (ref)	1,939	29.0	1550	389	1		
	Working outside Civil Service	1,387	20.7	1013	374	1.47	1.25-1.73	<0.001
	Not working - retired	2,994	44.7	1990	1004	2.01	1.76-2.30	<0.001
	Out of work	268	4.0	179	89	1.98	1.50-2.61	<0.001
	Not working-lt sick	106	1.6	89	17	0.76	0.45-1.29	0.313
Car availability	Car available (ref)	5,779	86.8	4072	1707	1		
	No car available	881	13.2	719	162	0.54	0.45-0.64	<0.001
Marital status	Married (ref)	4,832	72.4	3364	1468	1		
	Cohabit	273	4.1	205	68	0.76	0.57-1.01	0.056
	Single	1,021	15.3	829	192	0.53	0.45-0.63	<0.001
	Divorced	475	7.1	350	125	0.82	0.66-1.01	0.065
	Widowed	78	1.2	65	13	0.46	0.25-0.83	0.011
Ethnicity	White (ref)	6,156	92.0	4,378	1,778	1		
	Non-white	538	8.0	443	95	0.53	0.42-0.66	<0.001

4.3.2.2 Multivariate associations of WHOPA_E with individual level factors

Multivariate regression of the factors found to be significantly associated with WHOPA_E in the WIIP7 study sample showed all to remain significantly associated with the outcome (Table 4.7). Odds of WHOPA_E were higher for those working outside the Civil Service, retired individuals and those out of work than for individuals still in Civil Service employment. However, individuals who were not working due to long-term sickness had lower odds of WHOPA_E than those remaining in the Civil Service. Lower odds of WHOPA_E were also found for females, those without car access, non-married individuals and those whose ethnicity was non-white. Significant association of WHOPA_E with age in the WIIP7 sample were not found in the multivariate regression.

Table 4.7 Multivariate association of WHOPA_E with individual level sociodemographic factors in the WIIP7 study sample

Individual level socio-demographic factor		OR	CI	p
Sex	Male (ref)	1		
	Female	0.86	0.76-0.98	0.026
Age group	50y to <56y (ref)	1		
	>=56y to <60y	1.06	0.90-1.26	0.485
	>=60y to <66y	1.06	0.87-1.28	0.575
	>=66y to 75y	1.03	0.83-1.26	0.815
Economic activity	Remaining in Civil Service (ref)	1		
	Working outside Civil Service	1.38	1.17-1.63	<0.001
	Not working - retired	2.08	1.82-2.39	<0.001
	Out of work	2.01	1.52-2.67	<0.001
	Not working-lt sick	0.87	0.51-1.49	0.616
Car availability	Car available (ref)	1		
	No car available	0.63	0.52-0.77	<0.001
Marital status	Married (ref)	1		
	Cohabit	0.82	0.62-1.09	0.175
	Single	0.60	0.50-0.71	<0.001
	Divorced	0.89	0.71-1.10	0.279
	Widowed	0.47	0.25-0.86	0.014
Ethnicity	White (ref)	1		
	Non-white	0.52	0.41-0.65	<0.001
cons		0.30	0.27-0.34	<0.001

4.3.2.3 Univariate associations of WHOPA_E with area level deprivation

In univariate analysis area deprivation was significantly associated with the WHOPA_E outcome in the WIIP7 study sample (Table 4.8). Those living in the most deprived LSOAs were less likely to achieve the WHOPA_E outcome than those in the least deprived areas (OR=0.46, p<0.001).

Table 4.8 Univariate associations of WHOPA_E with area level deprivation in the WIIP7 study sample.

Area level deprivation	N	%	No	Yes	OR	CI	p
Quintile 1 (least deprived) (ref)	2,661	41.0	1,825	836	1		
Quintile 2	1,472	22.7	1,019	453	0.97	0.85-1.11	0.670
Quintile 3	1,154	17.8	863	291	0.74	0.63-0.86	<0.001
Quintile 4	829	12.8	662	167	0.55	0.46-0.66	<0.001
Quintile 5 (most deprived)	376	5.8	310	66	0.46	0.35-0.61	<0.001

4.3.3 Associations between physical activity exposure and GHQ-12 depression in the WIIP7 study sample (Study 1C)

Results of statistical modelling of associations between the physical activity outcomes and GHQ-12 depression are presented without adjustment for correlates, with adjustment for individual-level sociodemographic factors only and with adjustment for individual-level sociodemographic factors and area deprivation in Table 4.9, Table 4.10 and Table 4.11, respectively. Before adjustment for individual level factors and area level deprivation the odds of the GHQ-12 depression outcome were significantly lower among those meeting the WHO recommended physical activity level excluding walking as a contributory physical activity (OR=0.60, $p<0.05$). Similarly, they were significantly lower among those meeting the WHO recommended physical activity level including walking if reported brisk (OR=0.61, $p<0.05$), and among those in the top tertile of the WIIP7 sample for time spent walking per week (OR=0.66, $p<0.05$). Even after adjustment for individual correlates known to be associated with depression, and after further adjustment for area deprivation, the odds remained significantly lower and adjustment only weakened the associations slightly.

Table 4.9 Associations between physical activity exposure and GHQ-12 depression in the WIIP7 study sample before adjustment for correlates.

Physical activity exposure	N	OR	CI	p
WHOPA_E[1]	6599	0.60	0.51-0.71	<0.001
WHOPA_I[2]	6425	0.61	0.53-0.70	<0.001
TTW[3]	6330	0.66	0.56-0.77	<0.001

[1]Meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity ;[2]Meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; [3]Being in the top tertile for time spent walking per week

Table 4.10 Associations between physical activity exposure and GHQ-12 depression in the WIIP7 study sample, after adjustment for individual factors.

Physical activity exposure	N	OR	CI	p
WHOPA_E[1]	6585	0.68	0.57-0.80	<0.001
WHOPA_I[2]	6411	0.66	0.57-0.77	<0.001
TTW[3]	6318	0.67	0.57-0.80	<0.001

[1]Meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity ;[2]Meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; [3]Being in the top tertile for time spent walking per week

Table 4.11 Associations between physical activity exposure and GHQ-12 depression in the WIIP7 study sample after adjustment for individual factors and area deprivation.

Physical activity exposure	N	OR	CI	p
WHOPA_E[1]	6384	0.67	0.56-0.80	<0.001
WHOPA_I[2]	6215	0.66	0.57-0.76	<0.001
TTW[3]	6121	0.67	0.57-0.79	<0.001

[1]Meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity ;[2]Meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; [3]Being in the top tertile for time spent walking per week

4.4 Discussion

The purpose of Study 1 was to enable better modelling and understanding of associations of physical environmental factors and physical activity, and GHQ-12 depression outcomes in subsequent studies within this thesis through clarification of associations between non-physical environmental factors and these outcomes.

Study 1A showed that many factors hypothesised to be putative correlates of GHQ-12 depression were significantly associated with the GHQ-12 outcome in the HSE2008 and WIIP7 study samples. Non-physical environment factors identified by multivariate regression as being significantly associated with the GHQ-12 depression in the WIIP7 study sample, independently of one another, were sex, age, economic activity, marital status and ethnicity. Therefore, in statistical modelling of associations between the neighbourhood physical environmental exposures and GHQ-12 depression in the WIIP7 study sample, it was deemed necessary to include sex, age, economic activity, marital status and ethnicity as correlates. It was preferable, for consistency and simplicity, to adjust for this core set of factors in statistical modelling of associations of physical environment exposure with the GHQ-12 depression outcome in the HSE2008 study sample as well. However, the set of factors for which statistical associations were tested, and which were significantly associated, in the WIIP7 study sample was not the same as that in the HSE2008 sample. The lack of a significant association between income and GHQ-12 depression in the WIIP7 study sample as found for the HSE2008 study sample may reflect the higher income bands in the WIIP7 study sample: it may be that above a threshold value, higher income does not reduce odds of GHQ-12 depression. Given the significant univariate associations of this variable, and also of occupational social class – an equivalent of which was absent in the WIIP7 dataset – with GHQ-12 depression in the HSE2008 study sample, these factors were considered important correlates for the HSE2008 study sample. Therefore, in the HSE2008 study sample additional adjustment was deemed appropriate for income and for occupational social class. It was not possible to adjust for ethnicity in the HSE2008 study sample due to the absence of this variable in the local authority linked dataset. In light of the significant association of area level deprivation with GHQ-12 depression in both the WIIP7 and HSE2008 study samples found in univariate analysis, it was considered necessary to include this area level factor alongside the individual level sociodemographic variables in statistical modelling of associations with physical environment exposures.

All individual level sociodemographic factors hypothesised to be putative correlates of physical activity – sex, age, economic activity, car availability, marital status and ethnicity – and the contextual factor of area level deprivation were found, in Study 1B, to be significantly associated with the physical activity outcome of meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity in the WIIP7 study sample. The significantly higher odds of this outcome for employed ex-civil servants relative to remaining Civil Service employees could reflect the predominantly sedentary nature of jobs particular to the Civil Service: it may be that those working elsewhere were in occupations that were more physically demanding and that might make them more physically active and thereby likely to meet physical activity recommendations. The findings in the present study suggesting that economic inactivity through retirement and unemployment are predictive of higher odds of meeting the recommended level of physical activity are inconsistent with the literature reviewed previously which suggested the reverse in general adult populations. With regard to the higher odds found for retirees, this disparity may be an attribute of the narrow age range of the WIIP7 study sample: the WIIP7 retirees were youthful relative to the median age of retirees in the general adult population and thus potentially more physically able to exploit free time for physical activity. With regard to the higher odds for unemployed individuals, the disparity may be explained by the fact that a higher proportion of the WIIP7 study sample than in the general adult population was married or cohabiting, possibly attributable to the high median age. As such they might not only have more time to be physically active than their employed WIIP7 counterparts but also be supported by a waged spouse or partner and thereby free of any perceived or actual economic barriers to participation in physical activity that might be more likely to exist in the general unemployed adult population with a lower rate of marriage and cohabitation.

The finding in Study 1B that those without car access had lower odds of meeting the WHO recommended physical activity level, excluding any walking as a contributory physical activity, was interesting in light of initial exploratory analysis which showed that they had higher odds of being in the top tertile for time spent walking per week. Access to a car is likely determined by income which as mentioned previously may itself provide the choice to participate in leisure-based moderate to vigorous physical activities that contributes to meeting physical activity recommendations, and the choice not to undertake walking as active transport. However, car ownership has been found to mediate the relationship between the physical environment, measured as residential density and as land use mix, and physical activity, independent of neighbourhood-level income [202]. Therefore, it was important to adjust for car access in statistical modelling of associations between the physical environment variables and physical activity. Adjustment was particularly important for modelling the association between walkability, the derivation of which is detailed in Chapter 7, and physical activity because this variable was partly derived from measures of residential dwelling density and land use mix.

Multivariate regression of the individual level socioeconomic factors found in univariate regression to be significantly associated with meeting the WHO recommended physical activity level, excluding any walking as a contributory physical activity, in the WIIP7 study sample showed all except age to remain significantly associated with the outcome. Although age did not remain significantly associated with the WHOPA_E in multivariate regression, adjustment for this factor in modelling associations with neighbourhood physical environmental exposures was deemed valuable given the importance of age as identified in the literature review. Therefore, in statistical models of associations between the neighbourhood physical environmental exposures and this outcome, and the other physical activity outcomes in the WIIP7 study sample, it was considered necessary that the following factors were included as correlates: sex, age, economic activity, car availability, marital status and ethnicity. Given the significant association of area level deprivation with meeting the WHO recommended physical activity level, excluding any walking as a contributory physical activity, found in univariate analysis, it was deemed appropriate that this factor was included as an area level correlate alongside these individual level sociodemographic variables in statistical models.

The hypotheses that the odds of GHQ-12 depression would be significantly lower among those meeting the WHO recommended physical activity level excluding walking as a contributory physical activity, among those meeting the WHO recommended physical activity level including walking if reported brisk, and among those in the top tertile of the WIIP7 sample for time spent walking per week held true. Also as predicted, significant associations remained, albeit reduced, after adjustment for individual-level sociodemographic factors and then further adjustment for area deprivation. Due to the cross-sectional design of this study it is impossible to determine the direction of the relationship between physical activity and depression. Physical health is known to support mental health and reduce risk of depression. As better physical health is a consequence of being sufficiently active for health as indicated by meeting the WHO recommended physical activity level, it follows that physical activity reduces an individual's depression risk. However, it has been found that whilst regular physical activity may help prevent future depressive symptoms, depressive symptoms may also prevent adults from engaging in regular physical activity, indicating that the relationship between physical activity and depression is reciprocal [203].

4.5 Summary

In summary, Study 1A and 1B clarified which of the non-physical environmental factors identified in the literature of as putative confounders of relationships between physical environmental exposures and depression, and between physical environmental exposures and physical activity, were significantly associated with the outcomes specified in this study. This information informed the specification of the sets of correlates for which adjustment was necessary in modelling associations in subsequent studies and these are detailed in Table 4.12. Study 1C, in summary, showed that as predicted, GHQ-12 depression was negatively associated with physical activity whether specified as meeting the WHO recommended physical activity level with or without walking as a contributory physical activity, or as being in the top tertile of the WIIP7 sample for time spent walking per week. These associations were independent of individual-level sociodemographic factors and area-level deprivation.

Table 4.12 Correlates for which adjustment was necessary for each study sample for modelling associations with physical environment exposures with each outcome.

Study sample	Outcome	Correlates
WIIP7	GHQ-12 depression	Sex (Male (ref)/ Female); Age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y); Economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick); Marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed); Ethnicity (White (ref)/ Non-white); Area deprivation (Quintile 1 (least deprived) (ref)/ Quintile 2/ Quintile 3/ Quintile 4/ Quintile 5 (most deprived))
HSE2008	GHQ-12 depression	Sex (Male (ref)/ Female); Age (16y-34y (ref)/ 35y-54y/ 55y+); Economic activity (Employed (ref)/ Unemployed/ Retired/ Other), income (<=£14918 (ref)/ >£14918-£31871/>£31871); Occupational social class (Managerial and professional (ref)/ Intermediate/ Routine and manual/ Other); Marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed); Area deprivation (Quintile 1 (least deprived) (ref)/ Quintile 2/ Quintile 3/ Quintile 4/ Quintile 5 (most deprived))
WIIP7	WHOPA_E[1]	Sex (Male (ref)/ Female); Age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y); Economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick); Car availability (Car available (ref)/No car available);Marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed); Ethnicity (White (ref)/ Non-white); Area deprivation (Quintile 1 (least deprived) (ref)/ Quintile 2/ Quintile 3/ Quintile 4/ Quintile 5 (most deprived))
WIIP7	WHOPA_I[2]	
WIIP7	TTW[3]	

[1]Meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity ;[2]Meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; [3]Being in the top tertile for time spent walking per week

4.6 Appendices

Appendix 4.1 Univariate associations of WHOPA_I with individual level factors in the WIIP7 study sample

Individual level socio-demographic factor		N	%	No	Yes	OR	CI	p
Sex	Male (ref)	4,624	71.01	2,038	2586	1		
	Female	1,888	28.99	1,141	747	0.52	0.46-0.58	<0.001
Age group	50y to <56y (ref)	1,611	24.74	786	825	1		
	>=56y to <60y	1,531	23.51	721	810	1.07	0.93-1.23	0.342
	>=60y to <66y	1,663	25.54	785	878	1.07	0.93-1.22	0.364
	>=66y to 75y	1,707	26.21	887	820	0.88	0.77-1.01	0.068
Economic activity	Remaining in Civil Service (ref)	1,897	29.13	959	938	1		
	Working outside Civil Service	1,356	20.82	622	734	1.21	1.05-1.39	0.008
	Not working - retired	2,894	44.44	1,405	1489	1.08	0.97-1.22	0.175
	Out of work	261	4.01	113	148	1.34	1.03-1.74	0.028
	Not working-lt sick	104	1.6	80	24	0.31	0.19-0.49	<0.001
Car availability	Car available (ref)	5,646	87.06	2651	2995	1		
	No car available	839	12.94	507	332	0.58	0.50-0.67	<0.001
Marital status	Married (ref)	4,703	72.39	2151	2552			
	Cohabit	271	4.17	132	139	0.89	0.69-1.13	0.340
	Single	994	15.3	596	398	0.56	0.49-0.65	<0.001
	Divorced	458	7.05	243	215	0.75	0.62-0.90	0.003
	Widowed	71	1.09	50	21	0.35	0.21-0.59	<0.001
Ethnicity	White (ref)	6,021	92.46	2,832	3,189	1		
	Non-white	491	7.54	347	144	0.37	0.30-0.45	<0.001

Appendix 4.2 Multivariate association of WHOPA_I with individual level sociodemographic factors in the WIIP7 study sample

Individual level socio-demographic factor		OR	CI	p
Sex	Male (ref)	1		
	Female	0.59	0.53-0.66	<0.001
Age group	50y to <56y (ref)			
	>=56y to <60y	0.99	0.85-1.15	0.899
	>=60y to <66y	0.88	0.74-1.05	0.148
	>=66y to 75y	0.69	0.57-0.83	<0.001
Economic activity	Remaining in Civil Service (ref)	1		
	Working outside Civil Service	1.19	1.02-1.39	0.027
	Not working - retired	1.42	1.20-1.68	<0.001

Individual level socio-demographic factor		OR	CI	p
	Out of work	1.64	1.24-2.16	<0.001
	Not working-lt sick	0.36	0.23-0.58	<0.001
Car availability	Car available (ref)	1		
	No car available	0.81	0.69-0.95	0.010
Marital status	Married (ref)	1		
	Cohabit	0.92	0.72-1.19	0.530
	Single	0.64	0.55-0.75	<0.001
	Divorced	0.96	0.78-1.18	0.701
	Widowed	0.55	0.32-0.93	0.026
Ethnicity	White (ref)	1		
	Non-white	0.41	0.33-0.50	<0.001
cons				

Appendix 4.3 Univariate associations of TTW with individual level factors in the WIIP7 study sample

Individual level socio-demographic factor		N	%	No	Yes	OR	CI	p
Sex	Male (ref)	4,610	71.93	3,174	1436	1		
	Female	1,799	28.07	1,203	596	1.10	0.97-1.23	0.126
Age group	50y to <56y (ref)	1,608	25.09	1,160	448	1		
	>=56y to <60y	1,513	23.61	1,048	465	1.15	0.98-1.34	0.078
	>=60y to <66y	1,625	25.35	1,103	522	1.23	1.05-1.42	0.008
	>=66y to 75y	1,663	25.95	1,066	597	1.45	1.25-1.68	<0.001
Economic activity	Remaining in Civil Service (ref)	1,886	29.43	1,330	556	1		
	Working outside Civil Service	1,351	21.08	987	364	0.88	0.75-1.03	0.115
	Not working - retired	2,825	44.08	1,827	998	1.31	1.15-1.48	<0.001
	Out of work	256	3.99	166	90	1.30	0.99-1.71	0.064
	Not working-lt sick	91	1.42	67	24	0.86	0.53-1.38	0.525
Car availability	Car available (ref)	5,563	87.13	3894	1669	1		
	No car available	822	12.87	469	353	1.76	1.51-2.04	<0.001
Marital status	Married (ref)	4,642	72.58	3218	1424	1		
	Cohabit	267	4.17	190	77	0.92	0.70-1.20	0.526
	Single	973	15.21	625	348	1.26	1.09-1.45	0.002
	Divorced	443	6.93	291	152	1.18	0.96-1.45	0.114
	Widowed	71	1.11	48	23	1.08	0.66-1.79	0.756
Ethnicity	White (ref)	5,946	92.78	4,019	1,927	1		
	Non-white	463	7.22	358	105	0.61	0.49-0.77	<0.001

Appendix 4.4 Multivariate association of TTW with individual level sociodemographic factors in the WIIP7 study sample

Individual level socio-demographic factor		OR	CI	p
Sex	Male (ref)	1		
	Female	1.00	0.88-1.13	0.956
Age group	50y to <56y (ref)	1		
	>=56y to <60y	1.15	0.98-1.35	0.087
	>=60y to <66y	1.20	1.00-1.45	0.055
	>=66y to 75y	1.32	1.08-1.62	0.007
Economic activity	Remaining in Civil Service (ref)	1		
	Working outside Civil Service	0.84	0.71-0.99	0.039
	Not working - retired	1.10	0.92-1.32	0.296
	Out of work	1.18	0.88-1.57	0.262
	Not working-lt sick	0.74	0.46-1.20	0.224
Car availability	Car available (ref)	1		
	No car available	1.72	1.46-2.03	<0.001
Marital status	Married (ref)	1		
	Cohabit	0.91	0.69-1.20	0.514
	Single	1.05	0.89-1.23	0.553
	Divorced	1.05	0.84-1.30	0.679
	Widowed	0.88	0.53-1.48	0.635
Ethnicity	White (ref)	1		
	Non-white	0.57	0.45-0.71	<0.001
cons				

5 Study 2: Associations between neighbourhood greenspace and (A) physical activity, and (B) GHQ-12 depression

5.1 Introduction

5.1.1 Chapter overview

This focus of this chapter is the investigation of associations of neighbourhood exposure to greenspace, and greenspace plus domestic garden, with physical activity in an occupational-specific sample of older adults and with depression in this sample as well as in a nationally representative survey of adults. Neighbourhood exposure to greenspace and to domestic garden is indicated by the proportion of these land uses within neighbourhoods operationalised at three administrative geographies, namely local authorities (LAs), Census Area Statistics (CAS) wards and output areas (OAs). Associations are explored with respect to the influence of sociodemographic factors and area deprivation.

5.1.2 Research questions

The research questions are as follows:

- Is proportion of neighbourhood greenspace, and greenspace plus domestic garden, associated with physical activity in the Whitehall II Phase 7 (WIIP7) [182] study sample, and with a score of four or more on the 12-item General Health Questionnaire (GHQ-12 depression) in the WIIP7 and the 2008 Health Survey for England (HSE2008) [183] study samples?
- How does operationalisation of neighbourhood affect associations?
- Are associations independent of individual-level sociodemographic factors and of area deprivation?

5.1.3 Objectives for Study 2A

The objectives for Study 2A are as follows:

- to determine whether odds of physical activity are higher for an individual in the WIIP7 study sample living within a neighbourhood, defined as an OA, a CAS ward or as a LA where proportion of greenspace, and greenspace plus domestic garden, are higher
- to determine whether the association is affected by individual-level sociodemographic factors and by area deprivation

5.1.4 Hypotheses for Study 2A

Physical activity outcomes are associated with the proportion of greenspace within administratively-defined neighbourhoods, with individuals living in neighbourhoods where proportion of greenspace is higher having higher odds of the physical activity outcomes.

The association between the physical activity outcomes and the proportion of greenspace plus domestic garden is stronger than the association between the physical activity outcomes and the proportion of greenspace only. In addition, associations are stronger for smaller administratively-defined neighbourhoods.

The strength of associations is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and for area deprivation.

5.1.5 Objectives for Study 2B

- to determine whether odds of depression are lower for an individual in the WIIP7 study sample living within a neighbourhood, defined as an OA, a CAS ward or as a LA where percentage of greenspace, and greenspace plus domestic garden, is higher
- to determine whether odds of depression are lower for an individual in the HSE2008 study sample living within a neighbourhood, defined as a LA where percentage of greenspace, and greenspace plus domestic garden, is higher
- to determine whether the association is affected by individual-level sociodemographic factors and by area deprivation

5.1.6 Hypotheses for Study 2B

GHQ-12 depression is associated with the proportion of greenspace within administratively-defined neighbourhoods, with individuals living in neighbourhoods where proportion of greenspace is higher having lower odds of GHQ-12 depression.

The association between GHQ-12 depression and the proportion of greenspace plus domestic garden is stronger than the association between GHQ-12 depression and the proportion of greenspace only. In addition, associations are stronger for smaller administratively-defined neighbourhoods.

The strength of associations is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and for area deprivation.

5.2 Methods

5.2.1 Study samples

Participants in Study 2A and 2B comprised the WIIP7 study sample (see 3.2.1.2) and, additionally in separate analyses of associations between greenspace and depression, participants in Study 2B comprised the HSE2008 study sample (see Section 3.2.2.2). The WIIP7 study sample was drawn from the seventh wave, conducted in 2004/5, of the Whitehall II study, a longitudinal study of civil servants to examine the social determinants of health [182]. In contrast, the HSE2008 study sample was from the Health Survey for England conducted in 2008 which drew a nationally representative general population sample of adults living in households in England through multi-stage stratified probability sampling. In the WIIP7 study population of 6885 individuals, the mean age was 61.2 ± 6.0 years, ranging from 50.5 to 74.1 years, 29.8% were women whilst, among the 14,221 individuals in HSE2008, nationally representative of adults in England, the mean age was 48.9 ± 18.6 years, ranging from 16 to 97, and 55.5% were women.

5.2.2 Variables

5.2.2.1 Exposure variables

The exposure variables were quintiles of the proportions of greenspace, and of greenspace plus domestic garden, were derived at each administrative level across England as detailed in Chapter 3 (Section 3.3.2.1). In addition they were specifically derived at each administrative level within London alone. Neighbourhoods were operationalized, in descending order of administrative geography hierarchy, as LAs, CAS wards and OAs.

5.2.2.2 Outcome variables

For Study 2A, three physical activity outcomes were specified; meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E); meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I); and being in the top tertile for time spent walking per week (TTW). Chapter 3 (Section 3.3.3.2) provides the rationale for the specification of these variables as such. The WHO recommended physical activity level [47] was defined as follows:

- at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or at least 75 minutes of vigorous intensity aerobic physical activity throughout the week, or an equivalent combination of moderate and vigorous-intensity activity

For Study 2B, GHQ-12 depression was the outcome and was defined as score of four or more on the GHQ-12, the rationale provided in Chapter 3 (Section 3.3.3.1).

5.2.2.3 Correlates

Individual level correlates entered into the regression models were sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), car availability (Car available (ref)/No car available), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white) in Study 2A. In Study 2B adjustment was made for the individual level factors of sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white) in the WIIP7 study sample, and for sex (Male (ref)/

Female), age (16y-34y (ref)/ 35y-54y/ 55y+), economic activity (Employed (ref)/ Unemployed/ Retired/ Other), income (<=£14918 (ref)/ >£14918-£31871/>£31871), occupational social class (Managerial and professional (ref)/ Intermediate/ Routine and manual/ Other) and marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) in the HSE2008 study sample. In both studies adjustment was also made for a contextual factor, area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level [197]. The justification for adjustment for these individual and area level factors is provided in Chapter 4 (Section 4.4).

5.2.3 Statistical analysis of associations between neighbourhood greenspace and physical activity and GHQ-12 depression outcomes

Associations between neighbourhood proportion of greenspace, and proportion of greenspace plus domestic garden, and the physical activity and GHQ-12 outcomes in the study samples were examined using multivariate logistic regression. This statistical modelling approach was chosen for reasons given in Chapter 3. First, bivariate logistic regression models were specified which included only the greenspace, or greenspace plus domestic garden, exposure variable and outcome to determine the presence and strength of significant associations. For models in which significant associations were detected, adjustment was then made for correlates – firstly individual level sociodemographic factors and then additionally area deprivation.– to show the extent to which the relationship between the exposure and outcome was independent of other factors. Three models were constructed; Model 1, the unadjusted model was univariate analysis between the exposure and the outcome; Model 2 was multivariate analysis adjusting for all individual factors significantly associated with the outcome in the study sample; and Model 3 adjusted for area deprivation, in addition to the individual factors. Owing to the large number of permutations of exposure and outcome variables, only the England-based quintiles of the greenspace exposure at each administrative geography were modelled in the adjusted analysis. Results were computed as odds ratios alongside their 95% confidence intervals to indicate with 95% confidence the range within which the true value of the odds ratio lay.

In each model the reference category of the exposure was Quintile 1, representing neighbourhoods the lowest proportion of greenspace, or greenspace plus domestic garden. Thus an odds ratio indicated the odds of the outcome in those exposed to, or living in a neighbourhood with a higher proportion of greenspace – in Quintile 2, 3, 4 or 5 – relative to those living in a neighbourhood with the lowest proportion. Tests for trend were performed to evaluate the association between the outcome, be it physical activity or GHQ-12 depression, and each greenspace variable with respect to the trend for a dose effect of the quintiles.

5.3 Results

5.3.1 Results for Study 2A: Associations between greenspace and physical activity outcomes in WIIP7

Empty models were constructed to examine all relationships and adjusted models were constructed to examine only a selection of the relationships identified in the empty models as significant. Three models were constructed for each physical activity outcome; Model 1, the unadjusted model was univariate analysis between the greenspace exposure and the outcome; Model 2 was multivariate analysis adjusting for all individual level sociodemographic factors significantly associated with the outcome in the WIIP7 study sample; and Model 3 adjusted for an area level variable, area deprivation as quintiles of deprivation based on the IMD2004 at LSOA level, in addition to the individual factors.

5.3.1.1 Associations between greenspace and WHOPA_E in WIIP7

The empty model results of the regression models constructed to examine the relationship between the exposure variables and meeting the WHO recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E) are presented in Table 5.1 for quintiles of greenspace as the exposure variable and in Table 5.2 for quintiles greenspace plus domestic garden. With respect to greenspace as the exposure variable, dose responses between the exposure and the outcome were evident when quintiles were derived for neighbourhoods defined as LAs and, especially, as CAS Wards from all in England: increasingly higher odds of WHOPA_E were associated with increasingly greater proportions of greenspace (Table 5.1). Relative to those living in the lowest quintile neighbourhoods, the odds of this outcome among individuals in the WIIP7 study sample living in the highest quintile was 1.64 ($p<0.001$) when neighbourhood was defined as the LA and 1.76 ($p<0.001$) as the CAS Ward. Operationalisation of neighbourhood as OA with quintiles derived from all England OAs yielded a dose response as well but it was less strong and only reached significance for the odds of WHOPA_E for the top quintile relative to the bottom (OR=1.61, $p<0.001$). When quintiles were derived for neighbourhoods operationalized specifically from administrative geographies of all in London, significant associations were only found for LAs. Relative to those living in London LAs with the lowest London-derived quintile of greenspace, individuals living in the second highest and the highest had odds of WHOPA_E of 1.49 ($p<0.05$) and 1.67 ($p<0.01$), respectively.

Significant dose responses between the exposure and the outcome were also evident when the exposure was defined as greenspace plus domestic garden, regardless of the administrative geography at which neighbourhood was operationalized or whether quintiles were derived from all administrative units in England, or selectively from only all those in London (Table 5.2). Associations tended to be stronger when neighbourhoods were defined as CAS Wards and as OAs than as LAs. Over twice the odds of WHOPA_E relative to those living in the lowest quintiles were found for WIIP7 study participants living as neighbourhoods defined as CAS Wards derived from all in London (OR=2.10, $p<0.001$), and also for WIIP7 study participants living as neighbourhoods defined as OAs derived from all in England (OR=2.34, $p<0.001$).

Table 5.1 Associations between quintiles of greenspace within administrative areas and WHOPA_E in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-48%)	2,806	43.22	2,157	649	1	REF	
		2(48-74%)	1,318	20.30	927	391	1.40	1.21-1.62	<0.001
		3(74-86%)	963	14.83	649	314	1.61	1.37-1.89	<0.001
		4(86-92%)	966	14.88	652	314	1.60	1.36-1.88	<0.001
		5(92-97%)	439	6.76	294	145	1.64	1.32-2.04	<0.001
Test for trend p < 0.001									
LA	London	1(5-21%)	366	12.54	300	66	1	REF	
		2(22-26%)	454	15.56	378	76	0.91	0.64-1.31	0.627
		3(27-34%)	526	18.03	404	122	1.37	0.98-1.92	0.064
		4(34-41%)	882	30.23	664	218	1.49	1.10-2.03	0.011
		5(41-59%)	690	23.65	505	185	1.67	1.21-2.28	0.002
Test for trend p < 0.001									
CAS Ward	England	1(1-30%)	2,499	38.49	1,909	590	1	REF	
		2(30-51%)	1,396	21.50	1,012	384	1.23	1.06-1.43	0.007
		3(51-77%)	1,320	20.33	927	393	1.37	1.18-1.59	<0.001
		4(77-92%)	817	12.58	533	284	1.72	1.45-2.05	<0.001
		5(92-99%)	460	7.09	298	162	1.76	1.42-2.18	<0.001
Test for trend p < 0.001									
CAS Ward	London	1(2-12%)	484	16.59	377	107	1	REF	
		2(12-19%)	527	18.06	426	101	0.84	0.62-1.13	0.248
		3(19-25%)	654	22.41	514	140	0.96	0.72-1.28	0.777
		4(25-38%)	572	19.60	435	137	1.11	0.83-1.48	0.479
		5(39-90%)	681	23.34	499	182	1.29	0.98-1.69	0.072
Test for trend p < 0.01									
OA	England	1(0-4%)	1,790	27.63	1,335	455	1	REF	
		2(4-13%)	1,293	19.96	958	335	1.03	0.87-1.21	0.759
		3(13-29%)	1,139	17.58	842	297	1.03	0.87-1.23	0.692
		4(29-64%)	1,096	16.92	783	313	1.17	0.99-1.39	0.064
		5(64-100%)	1,160	17.91	749	411	1.61	1.37-1.89	<0.001
Test for trend p < 0.001									
OA	London	1(0-1%)	687	23.63	521	166	1	REF	

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
		2(1-5%)	661	22.74	504	157	0.98	0.76-1.26	0.860
		3(5-12%)	570	19.61	449	121	0.85	0.65-1.10	0.217
		4(12-27%)	490	16.86	385	105	0.86	0.65-1.13	0.272
		5(27-100%)	499	17.17	383	116	0.95	0.72-1.25	0.714
Test for trend non-significant (p>0.05)									

Table 5.2 Associations between quintiles of greenspace plus domestic garden within administrative areas and WHOPA_E in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-69%)	2,407	37.08	1,851	556	1	REF	
		2(69-84%)	1,542	23.75	1,116	426	1.27	1.10-1.47	0.001
		3(84-92%)	863	13.29	580	283	1.62	1.37-1.93	<0.001
		4(92-95%)	1,039	16.00	705	334	1.58	1.34-1.85	<0.001
		5(95-98%)	641	9.87	427	214	1.67	1.38-2.02	<0.001
Test for trend p < 0.001									
LA	London	1(5-41%)	286	9.80	234	52	1	REF	
		2(42-52%)	540	18.51	441	99	1.01	0.70-1.46	0.957
		3(54-58%)	493	16.90	371	122	1.48	1.03-2.13	0.035
		4(61-68%)	618	21.18	465	153	1.48	1.04-2.10	0.029
		5(69-81%)	981	33.62	740	241	1.47	1.05-2.05	0.025
Test for trend p < 0.01									
CAS Ward	England	1(3-61%)	2,047	31.53	1,603	444	1	REF	
		2(61-74%)	1,593	24.54	1,143	450	1.42	1.22-1.65	<0.001
		3(74-87%)	1,401	21.58	989	412	1.50	1.29-1.76	<0.001
		4(87-95%)	906	13.96	591	315	1.92	1.62-2.29	<0.001
		5(95-99%)	545	8.39	353	192	1.96	1.60-2.41	<0.001
Test for trend p < 0.001									
CAS Ward	London	1(3-36%)	345	11.82	293	52	1	REF	
		2(36-49%)	416	14.26	324	92	1.60	1.10-2.33	0.014
		3(49-59%)	618	21.18	484	134	1.56	1.10-2.22	0.013

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
		4(59-67%)	671	23.00	518	153	1.66	1.18-2.35	0.004
		5(67-96%)	868	29.75	632	236	2.10	1.51-2.93	<0.001
Test for trend p < 0.001									
OA	England	1(0-46%)	1,006	17.02	825	181	1	REF	
		2(46-59%)	1,223	20.70	932	291	1.42	1.16-1.75	0.001
		3(59-68%)	1,135	19.21	806	329	1.86	1.51-2.29	<0.001
		4(68-82%)	1,489	25.20	1,036	453	1.99	1.64-2.42	<0.001
		5(82-100%)	1,056	17.87	698	358	2.34	1.90-2.87	<0.001
Test for trend p < 0.001									
OA	London	1(0-32%)	327	11.26	273	54	1	REF	
		2(32-44%)	415	14.29	342	73	1.08	0.73-1.59	0.699
		3(44-54%)	590	20.31	456	134	1.49	1.05-2.11	0.027
		4(54-64%)	690	23.75	519	171	1.67	1.19-2.34	0.003
		5(64-100%)	883	30.40	650	233	1.81	1.31-2.52	<0.001
Test for trend p < 0.001									

5.3.1.2 Associations between greenspace and WHOPA_I in WIIP7

Regression models were constructed to examine the relationship between the exposure variables and meeting the WHO recommended physical activity level including walking as a contributory component if reported as at least brisk (WHOPA_I). Results of the empty regression models which excluded all correlates are reported in Appendix 5.1 for quintiles of greenspace as the exposure variable and in Appendix 5.2 for quintiles greenspace plus domestic garden. Significant associations between the exposure, specified as the proportion of greenspace, and the outcome were found when neighbourhood was defined as LA with quintiles derived from all England LAs. Relative to the reference category of the lowest quintile, all higher quintiles were associated with higher odds of WHOPA_I. The highest odds were found for WIIP7 participants living in LAs which had the highest proportion of greenspace (OR=1.59, p<0.001) but a dose response was not apparent as for the outcome which excluded walking. Individuals living in LAs with the highest greenspace quintile derived from these administrative units in London only also had significantly higher odds (OR==1.52, p<0.001) than those in the reference category but associations for those living in LAs below the highest quintile were not significant. As found for neighbourhoods defined as LAs, neighbourhoods defined as CAS Wards yielded significant and positive associations between quintiles of greenspace and the outcome but only when the quintiles were derived from all England administrative units. Also, no dose response was observed: individuals living in CAS Wards with the second highest proportion of greenspace had the highest odds at 1.54 ($p<0.001$) of WHOPA_I relative to those living in CAS Wards with the lowest proportion of greenspace. At OA-level, the association between greenspace and the physical activity outcome only reached significance for the highest proportion of greenspace as a quintile derived from all England OAs.

Significant positive associations between greenspace plus domestic garden and WHOPA_I were found when neighbourhoods were operationalized at any of the administrative geographies, and when quintiles were derived either from the administrative units across the whole of England or only those across London. Whilst dose responses were not evident, individuals with the highest odds of meeting the physical activity outcome relative to those with the reference exposure tended to be in administrative units where the greenspace plus domestic garden proportion was in top quintile. The highest odds ratio of 2.16 ($p<0.001$) was found for individuals with the highest greenspace plus domestic garden proportion in their neighbourhood defined as OAs, the smallest administrative areas, and as calculated from all across England,.

5.3.1.3 Associations between greenspace and TTW in WIIP7

Empty regression models were constructed to examine associations between greenspace and the physical activity outcome of being in the top tertile of WIIP7 for time spent walking per week (TTW). These empty models, which excluded all correlates, are reported in Appendix 5.3 for quintiles of greenspace as the exposure variable and in Appendix 5.4 for quintiles greenspace plus domestic garden. Associations between greenspace and walking outcome were not concordant with those for the WHO physical activity recommendation outcomes and were mainly non-significant. Where significant associations were detected, the proportion of greenspace in the neighbourhood was negatively associated with the TTW outcome. Relative to those living in LAs in the lowest England-based quintile for greenspace, WIIP7 participants in these administrative units in the highest quintile were less likely to have the TTW outcome (OR=0.75, $p<0.05$). Also, when neighbourhoods were defined as CAS wards, those in the third highest London-based quintile for greenspace had significantly lower odds than those in the lowest (OR=0.75, $p<0.05$).

Negative associations between greenspace plus domestic garden and TTW were found, in contrast to the positive associations found with meeting the WHO physical activity recommended level. It was noted that more of the associations detected between greenspace plus domestic garden and the TTW outcome reached significance than for greenspace alone. Lower odds of TTW in the WIIP7 study sample were found for London-based quintiles of the exposure variable in neighbourhoods defined as LAs, for both England and London-based quintiles in CAS Ward neighbourhoods and for England-based quintiles in OA neighbourhoods. The lowest significant odds of TTW relative to individuals in the reference category quintile were for individuals living in the third highest quintile for London-based greenspace and domestic garden in CAS Ward neighbourhoods (OR=0.69, $p<0.05$). Dose responses to proportions of greenspace plus domestic garden were not apparent at any neighbourhood administrative level.

Adjustment for all individual factors significantly associated with the outcome in the WIIP7 study sample and an area-level factor (area deprivation as quintiles of deprivation based on the Index of Multiple Deprivation at LSOA level) was made in the models to further investigate the relationships of the England-based quintiles of greenspace, and for greenspace plus domestic garden, with the TTW outcome. Although the data is not shown, the analysis revealed the significant association between TTW and the England-based quintiles of greenspace in neighbourhoods defined as LAs remained significant, albeit attenuated, after adjustment for individual level factors and further adjustment for area deprivation. The significant associations between TTW and the England-based quintiles of greenspace plus domestic garden in neighbourhoods defined as CAS Wards and as OAs that were found in the empty models, however, were lost after adjustment for individual level factors.

5.3.1.4 Adjusted models for associations between greenspace and WHOPA_E in WIIP7

Additional adjusted models were constructed to examine relationships where the outcome was meeting the WHO recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E), and the greenspace exposure variables constructed as England-based quintiles at each administrative geography. Adjustment was made for all individual factors significantly associated with the outcome in the WIIP7 study sample and an area-level factor, area deprivation as quintiles of deprivation based on the Index of Multiple Deprivation at LSOA level. Model 1 was univariate analysis between the greenspace exposure and the outcome, Model 2 adjusted for individual factors, and Model 3 additionally adjusted for area deprivation.

For individuals living in neighbourhoods defined as LAs, significant associations between WHOPA_E and proportion of greenspace (Table 5.3), and between WHOPA_E and proportion of greenspace plus domestic garden (Table 5.4), were weaker than for neighbourhoods defined as smaller administrative geographies. However, all associations for LA neighbourhoods remained significant after adjustment for individual factors, with significance lost only after addition of area deprivation in Model 3. Operationalisation of neighbourhoods as CAS wards produced significant associations between WHOPA_E and proportion of greenspace (Table 5.5), and between WHOPA_E and proportion of greenspace plus domestic garden (Table 5.6), that weakened after adjustment for individual level factors and again after adjustment for area deprivation but that remained significant. Similarly, the associations between the outcome and greenspace (Table 5.7), and between the outcome and greenspace plus domestic garden (Table 5.8) weakened but remained significant in the partially and then the fully adjusted model when neighbourhoods were defined as OAs. The highest odds of WHOPA_E in the fully adjusted models were for individuals living in neighbourhoods defined as OAs and where the exposure was greenspace plus domestic garden: relative to those with the lowest proportion in quintile 1, the odds for an individual living in neighbourhoods with the highest proportion, in quintile 5, were 1.56 ($p<0.001$).

Table 5.3 Association between greenspace (GS) at LA level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
n		6492			6446			6446	
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.40	1.21-1.62	<0.001	1.20	1.02-1.40	0.030	1.10	0.93-1.29	0.260
3	1.61	1.37-1.89	<0.001	1.30	1.10-1.54	0.002	1.20	1.00-1.43	0.049
4	1.60	1.36-1.88	<0.001	1.26	1.07-1.50	0.007	1.15	0.97-1.37	0.112
5	1.64	1.32-2.04	<0.001	1.25	1.00-1.57	0.050	1.16	0.92-1.45	0.217
	Test for trend p < 0.001			Test for trend p = 0.002			Trend non-signf (p>0.05)		

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 5.4 Association between greenspace plus domestic garden (GS+DG) at LA level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS+DG		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
n		6492			6446			6446	
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.27	1.10-1.47	0.001	1.12	0.96-1.30	0.149	1.03	0.88-1.21	0.676
3	1.62	1.37-1.93	<0.001	1.32	1.10-1.58	0.002	1.20	0.99-1.44	0.057
4	1.58	1.34-1.85	<0.001	1.23	1.04-1.46	0.015	1.12	0.93-1.33	0.227
5	1.67	1.38-2.02	<0.001	1.27	1.05-1.55	0.016	1.15	0.94-1.41	0.177
	Test for trend p < 0.001			Test for trend p = 0.002			Trend non-signf (p>0.05)		

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 5.5 Association between greenspace (GS) at CAS ward level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
n		6492			6446			6446	
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.23	1.06-1.43	0.007	1.11	0.95-1.29	0.195	1.07	0.92-1.25	0.382
3	1.37	1.18-1.59	<0.001	1.13	0.97-1.32	0.122	1.05	0.90-1.23	0.541
4	1.72	1.45-2.05	<0.001	1.39	1.16-1.66	<0.001	1.29	1.07-1.55	0.006
5	1.76	1.42-2.18	<0.001	1.38	1.11-1.72	0.004	1.29	1.03-1.61	0.026
	Test for trend p < 0.001			Test for trend p < 0.001			Test for trend p < 0.01		

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity;
[b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 5.6 Association between greenspace plus domestic garden (GS+DG) at CAS ward level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS+DG	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
n	6492			6446			6446		
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.42	1.22-1.65	<0.001	1.19	1.01-1.38	0.035	1.08	0.92-1.28	0.343
3	1.50	1.29-1.76	<0.001	1.16	0.99-1.37	0.071	1.04	0.87-1.24	0.680
4	1.92	1.62-2.29	<0.001	1.48	1.24-1.78	<0.001	1.32	1.09-1.61	0.005
5	1.96	1.60-2.41	<0.001	1.44	1.16-1.78	0.001	1.28	1.02-1.60	0.031
	Test for trend p < 0.001			Test for trend p < 0.001			Test for trend p < 0.01		

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 5.7 Association between greenspace (GS) at OA level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
n	6478			6433			6433		
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.03	0.87-1.21	0.759	0.98	0.83-1.16	0.853	0.98	0.83-1.16	0.820
3	1.03	0.87-1.23	0.692	1.00	0.84-1.19	0.977	0.99	0.83-1.18	0.898
4	1.17	0.99-1.39	0.064	1.07	0.90-1.28	0.427	1.05	0.88-1.25	0.590
5	1.61	1.37-1.89	<0.001	1.33	1.12-1.57	0.001	1.27	1.07-1.50	0.005
	Test for trend p < 0.001			Test for trend p = 0.001			Test for trend p < 0.01		

Table 5.8 Association between greenspace plus domestic garden (GS+DG) at OA level and WHOPA_E in the WIIP7 study sample, before and after adjustment for individual factors and area level deprivation

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS+DG		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
n		5909			5864			5864	
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	1.42	1.16-1.75	0.001	1.26	1.01-1.55	0.039	1.20	0.96-1.49	0.107
3	1.86	1.51-2.29	<0.001	1.55	1.25-1.93	<0.001	1.44	1.15-1.81	0.002
4	1.99	1.64-2.42	<0.001	1.53	1.24-1.89	<0.001	1.40	1.12-1.76	0.003
5	2.34	1.90-2.87	<0.001	1.72	1.38-2.14	<0.001	1.56	1.23-1.98	<0.001
	Test for trend p < 0.001			Test for trend p < 0.001			Test for trend p < 0.001		

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity;
[b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

5.3.2 Results for Study 2B: Associations between greenspace and GHQ-12 depression

As for the physical activity outcomes, empty models were constructed to examine all relationships between greenspace and GHQ-12 depression, with adjusted models constructed to examine only a selection of the relationships identified in the empty models as significant. Three models were constructed; Model 1, the unadjusted model was univariate analysis between the greenspace exposure and GHQ-12 depression; Model 2, adjusting for individual level factors; and Model 3 additionally adjusted for the area deprivation variable.

5.3.2.1 Associations between greenspace and GHQ-12 depression in WIIP7

This part of the study examined possible relationships between the proportion of greenspace, and between the proportion of greenspace plus domestic garden, within WIIP7 participants' neighbourhoods and GHQ-12 depression. The results of the empty regression models constructed to examine the relationship between the exposure variables and GHQ-12 depression are presented in Table 5.9 for quintiles of greenspace as the exposure variable and in Table 5.10 for quintiles greenspace plus domestic garden. Greenspace as the exposure variable tended to be negatively associated with GHQ-12 depression in neighbourhoods defined at all three administrative geographies: an individual living in a neighbourhood defined as a LA, a CAS ward or as an OA, with a higher proportion of greenspace were less likely have the GHQ-12 outcome. Significant negative associations were found for both England and London-based LA quintiles, for England-based CAS ward quintiles and London-based OA quintiles. Specified as greenspace plus domestic garden, the exposure variable was negatively associated with GHQ-12 depression in the same manner as the exposure variable specified as greenspace alone. The lowest odds of GHQ-12 depression were identified for WIIP7 participants living in CAS ward neighbourhoods in the second highest London-based exposure quintile of greenspace plus domestic garden (OR=0.56, p<0.001).

Table 5.9 Associations between quintiles of greenspace (GS) within administrative areas and GHQ-12 depression in the WIIP7 study sample

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-48%)	2,777	43.22	2,344	433	1	REF	
		2(48-74%)	1,301	20.25	1,125	176	0.85	0.70-1.02	0.085
		3(74-86%)	951	14.8	808	143	0.96	0.78-1.18	0.682
		4(86-92%)	959	14.93	840	119	0.77	0.62-0.95	0.017
		5(92-97%)	437	6.8	378	59	0.84	0.63-1.13	0.260
Test for trend p < 0.05									
LA	London	1(5-21%)	367	12.71	300	67	1	REF	
		2(22-26%)	447	15.48	368	79	0.96	0.67-1.38	0.829
		3(27-34%)	520	18.01	451	69	0.69	0.47-0.99	0.043
		4(34-41%)	871	30.16	747	124	0.74	0.54-1.03	0.074
		5(41-59%)	683	23.65	585	98	0.75	0.53-1.05	0.098
Test for trend p < 0.05									
CAS Ward	England	1(1-30%)	2,475	38.52	2,105	370	1	REF	
		2(30-51%)	1,376	21.42	1,155	221	1.09	0.91-1.31	0.359
		3(51-77%)	1,305	20.31	1,125	180	0.91	0.75-1.10	0.338
		4(77-92%)	809	12.59	714	95	0.76	0.59-0.96	0.023
		5(92-99%)	460	7.16	396	64	0.92	0.69-1.22	0.565
Test for trend p < 0.05									
CAS Ward	London	1(2-12%)	480	16.62	401	79	1	REF	
		2(12-19%)	522	18.07	444	78	0.89	0.63-1.25	0.510
		3(19-25%)	652	22.58	550	102	0.94	0.68-1.30	0.712
		4(25-38%)	564	19.53	469	95	1.03	0.74-1.43	0.868
		5(39-90%)	670	23.2	587	83	0.72	0.51-1.00	0.051
Test for trend non-significant (p>0.05)									
OA	England	1(0-4%)	1,775	27.68	1,508	267	1	REF	

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
		2(4-13%)	1,273	19.85	1,105	168	0.86	0.70-1.06	0.151
		3(13-29%)	1,127	17.58	954	173	1.02	0.83-1.26	0.821
		4(29-64%)	1,083	16.89	916	167	1.03	0.83-1.27	0.785
		5(64-100%)	1,154	18	1,001	153	0.86	0.70-1.07	0.179
Test for trend non-significant (p>0.05)									
OA	London	1(0-1%)	682	23.7	562	120	1	REF	
		2(1-5%)	654	22.72	566	88	0.73	0.54-0.98	0.037
		3(5-12%)	563	19.56	491	72	0.69	0.50-0.94	0.020
		4(12-27%)	483	16.78	406	77	0.89	0.65-1.22	0.458
		5(27-100%)	496	17.23	417	79	0.89	0.65-1.21	0.451
Test for trend non-significant (p>0.05)									

Table 5.10 Associations between quintiles of greenspace plus domestic garden (GS+DG) within administrative areas and GHQ-12 depression in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-69%)	2,381	37.06	2,005	376	1	REF	
		2(69-84%)	1,527	23.77	1,316	211	0.85	0.71-1.03	0.092
		3(84-92%)	851	13.25	726	125	0.92	0.74-1.14	0.445
		4(92-95%)	1,029	16.02	891	138	0.83	0.67-1.02	0.075
		5(95-98%)	637	9.91	557	80	0.77	0.59-0.99	0.043
Test for trend p < 0.05									
LA	London	1(5-41%)	287	9.94	234	53	1	REF	
		2(42-52%)	530	18.35	435	95	0.96	0.66-1.40	0.848
		3(54-58%)	488	16.9	426	62	0.64	0.43-0.96	0.030
		4(61-68%)	610	21.12	513	97	0.83	0.58-1.21	0.337
		5(69-81%)	973	33.69	843	130	0.68	0.48-0.97	0.032
Test for trend p < 0.05									
CAS Ward	England	1(3-61%)	2,028	31.56	1,695	333	1	REF	
		2(61-74%)	1,571	24.45	1,350	221	0.83	0.69-1.00	0.053
		3(74-87%)	1,386	21.57	1,192	194	0.83	0.68-1.00	0.055
		4(87-95%)	896	13.95	782	114	0.74	0.59-0.93	0.011

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
		5(95-99%)	544	8.47	476	68	0.73	0.55-0.96	0.026
Test for trend p < 0.01									
CAS Ward	London	1(3-36%)	345	11.95	277	68	1	REF	
		2(36-49%)	407	14.09	339	68	0.82	0.56-1.19	0.287
		3(49-59%)	616	21.33	516	100	0.79	0.56-1.11	0.174
		4(59-67%)	666	23.06	586	80	0.56	0.39-0.79	0.001
		5(67-96%)	854	29.57	733	121	0.67	0.48-0.93	0.018
Test for trend p < 0.01									
OA	England	1(0-46%)	998	17.08	826	172	1	REF	
		2(46-59%)	1,208	20.68	1,027	181	0.85	0.67-1.06	0.151
		3(59-68%)	1,124	19.24	962	162	0.81	0.64-1.02	0.075
		4(68-82%)	1,463	25.04	1,267	196	0.74	0.59-0.93	0.009
		5(82-100%)	1,049	17.96	918	131	0.69	0.54-0.88	0.003
Test for trend p = 0.001									
OA	London	1(0-32%)	327	11.37	275	52	1	REF	
		2(32-44%)	408	14.19	337	71	1.11	0.75-1.65	0.588
		3(44-54%)	586	20.38	482	104	1.14	0.79-1.64	0.478
		4(54-64%)	687	23.89	592	95	0.85	0.59-1.23	0.381
		5(64-100%)	868	30.18	754	114	0.80	0.56-1.14	0.218
Test for trend p < 0.05									

5.3.2.2 Associations between greenspace and GHQ-12 depression in HSE2008

Possible relationships between the proportion of greenspace, and between the proportion of greenspace plus domestic garden, within HSE2008 participants' LA-level neighbourhoods and GHQ-12 depression were explored. The proportions were constructed as quintiles derived from areas of the particular land use across all LAs in England. The results of the empty regression models are presented in Table 5.11 for both quintiles of greenspace as the exposure variable and for quintiles greenspace plus domestic garden. Greenspace as the exposure variable, and greenspace plus domestic garden as the exposure variable were both negatively associated with GHQ-12 depression in an apparently dose dependent manner, with the highest quintiles associated with the lowest odds of depression. The lowest odds were identified for the highest greenspace plus domestic garden quintile, at 0.74 (p<0.001).

Table 5.11 Associations between England-based quintiles of greenspace (GS), and of greenspace plus domestic garden (GS+DG), within LAs and GHQ-12 depression in the HSE2008 study sample.

Land Use	Quintile	N	%	No	Yes	OR	CI	p
GS	1(5-48%)	3,708	26.07	3,138	570	1	REF	
	2(48-74%)	3,395	23.87	2,934	461	0.87	0.76-0.99	0.032
	3(74-86%)	2,807	19.74	2,418	389	0.89	0.77-1.02	0.088
	4(86-92%)	2,412	16.96	2,122	290	0.75	0.65-0.88	<0.001
	5(92-97%)	1,899	13.35	1,662	237	0.79	0.67-0.92	0.004
Test for trend p < 0.001								
GS+DG	1(5-69%)	3,620	25.46	3,058	562	1	REF	
	2(69-84%)	3,468	24.39	2,997	471	0.86	0.75-0.98	0.021
	3(84-92%)	2,849	20.03	2,452	397	0.88	0.77-1.01	0.074
	4(92-95%)	2,405	16.91	2,113	292	0.75	0.65-0.88	<0.001
	5(95-98%)	1,879	13.21	1,654	225	0.74	0.63-0.87	<0.001
Test for trend p < 0.001								

5.3.2.3 Adjusted models for associations between greenspace and depression in WIIP7 and HSE2008

Results of adjusted models constructed to examine relationships between GHQ-12 depression and the greenspace exposure variables constructed as England-based quintiles at the various administrative geographies are presented here. Adjustment was made for individual factors and area deprivation. Model 1 was univariate analysis between the greenspace exposure and the outcome, Model 2 adjusted for individual factors, and Model 3 adjusted for area deprivation in addition to the individual factors.

In the WIIP7 study sample, adjustment for individual level factors rendered all associations between greenspace and GHQ-12 depression, and between greenspace plus domestic garden and GHQ-12 depression, that had been significant in Model 1, the empty model, insignificant. However, in the HSE2008 study sample significant associations remained after adjustment for individual level factors. Relative to individuals living in quintile1 with the lowest exposure, HSE2008 living in neighbourhoods defined as LAs where proportion of greenspace was higher had lower odds of GHQ-12 depression: for those in quintile 4, the second highest, the odds ratio was 0.82 ($p<0.05$) (Table 5.12). After individual-level factor adjustment, HSE2008 participants living in LA-defined neighbourhoods where proportion of greenspace plus domestic garden was higher also had lower odds of GHQ-

12 depression, the odds ratios of 0.82 (p<0.05) and 0.83 (p<0.05) for those in quintiles 4 and 5 respectively (Table 5.13). These significant associations between greenspace and GHQ-12 depression, and between greenspace plus domestic garden and GHQ-12 depression were eliminated, however, upon addition of area deprivation as a contextual variable in Model 3.

Table 5.12 Association between greenspace (GS) at LA level and GHQ-12 depression, before and after adjustment for individual factors and area level deprivation in the HSE2008 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
n	14221			11498			11498		
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	0.87	0.76-0.99	0.032	0.90	0.77-1.04	0.161	0.93	0.80-1.09	0.375
3	0.89	0.77-1.02	0.088	0.93	0.79-1.09	0.374	1.00	0.85-1.18	0.973
4	0.75	0.65-0.88	<0.001	0.82	0.69-0.98	0.025	0.92	0.77-1.11	0.381
5	0.79	0.67-0.92	0.004	0.85	0.71-1.03	0.097	0.96	0.79-1.16	0.646
	Test for trend p < 0.001			Test for trend p < 0.05			Trend test non-signf (p>0.05)		
	[a]Adjustment for sex, age, economic activity, income, occupational social class and marital status; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

Table 5.13 Association between greenspace plus domestic garden (GS+DG) at LA level and GHQ-12 depression, before and after adjustment for individual factors and area level deprivation in the HSE2008 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
GS+DG	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
n	14221			11498			11498		
Quintile	OR	CI	p	OR	CI	p	OR	CI	p
1	1	REF		1	REF		1	REF	
2	0.86	0.75-0.98	0.021	0.86	0.74-1.01	0.062	0.90	0.77-1.05	0.162
3	0.88	0.77-1.01	0.074	0.93	0.79-1.09	0.385	1.01	0.85-1.19	0.931

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
4	0.75	0.65-0.88	<0.001	0.82	0.69-0.98	0.026	0.92	0.76-1.10	0.352
5	0.74	0.63-0.87	<0.001	0.83	0.69-1.00	0.046	0.94	0.77-1.14	0.530
	Test for trend p < 0.001			Test for trend p < 0.05			Trend test non-signf (p>0.05)		
	[a]Adjustment for sex, age, economic activity, income, occupational social class and marital status; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

5.3.3 Trend analysis

Overall patterns in the relationships between the greenspace exposures and the physical activity and GHQ-12 depression outcomes were also evaluated using statistical tests for trend. Positive associations were found between the greenspace exposures and the physical activity outcomes specified as WHOPA_E and as WHOPA_I. Tests for trend revealed significant dose effects of exposure to greenspace on WHOPA_E whether neighbourhoods were defined as LAs, CAS wards or OAs. Only when the quintiles were derived for London only OAs was the trend insignificant. Significant dose effects were also apparent, as indicated by trend analysis, when the exposure was greenspace plus domestic garden at all levels of neighbourhood delineation. Similarly, there were significant trends for dose effects of both the greenspace alone and the greenspace plus domestic garden exposure in the association with WHOPA_I. Only for the exposure of greenspace alone in neighbourhoods operationalised as CAS wards and OAs when quintiles were derived from London were trends for dose effects non-significant with this outcome.

Negative associations were found between the greenspace exposures and physical activity outcomes specified as TTW. However, whilst associations between the greenspace exposures and the WHO recommended physical activity level outcomes tended to show trends for significant dose effects regardless of neighbourhood delineation, those between these exposures and the TTW outcome were dependent on neighbourhood operationalisation. For both the greenspace exposure and the greenspace plus domestic garden exposure trend analysis revealed significant dose effects of their associations with the TTW outcome only when neighbourhoods were operationalised as the smaller administrative units of OAs.

As with the TTW outcome, the greenspace exposure variables tended to show negative associations with GHQ-12 depression in HSE2008. When the exposure was specified as greenspace alone, tests for trend revealed significant dose effects in the association with this outcome when neighbourhoods were operationalised as LAs and CAS wards but not as OAs. However, when specified as greenspace plus domestic gardens significant dose effects were found regardless of neighbourhood operationalisation. In HSE2008, the larger study sample significant dose effects were found for both greenspace exposure variables in LAs, the only neighbourhood operationalisation for this sample.

5.4 Discussion

This study sought to examine associations of proportion of neighbourhood greenspace with physical activity and with depression. There was a particular focus on how operationalisation of neighbourhood affected associations. In addition this study aimed to determine whether associations were independent of individual-level sociodemographic factors of area deprivation.

As hypothesised, individuals with greater exposure to neighbourhood greenspace, as indicated by proportion, tended to have higher odds of the physical activity outcomes of meeting the WHO recommended physical activity level including or excluding walking as a contributory component. Also, for neighbourhoods operationalized at a given administrative geography where associations between greenspace and physical activity were significant, and where the outcomes were specified as meeting the WHO recommended physical activity level including or excluding walking as a contributory component, association between greenspace plus domestic garden tended to be stronger. The proportion of greenspace and domestic garden within neighbourhoods was independently associated with meeting the WHO recommended physical activity level excluding walking because associations remained after adjustment for individual and area-level variables. Whilst causation cannot be inferred due to the cross-sectional design of this study, the significant associations found suggest that physical activity does occur in greenspace and that activities, such as gardening, may occur within individuals' own gardens, contributing to higher odds of meeting WHO recommended physical activity levels when domestic garden is added to the exposure. However, specification of the physical activity outcome as being in the top tertile for time spent walking per week produced results that were not concordant with those for physical activity levels: where significant associations were detected, the proportion of greenspace in the neighbourhood was negatively associated with the walking outcome. This counterintuitive finding, which does not support the hypothesis, may be a manifestation of the inaccuracy of individuals' self-reported walking or it may be that derivation of physical activity level variables from a broader range of physical activity domains and intensities gives a better indication of their potential utilization of any neighbourhood greenspace. Alternatively, it may be that higher proportions of greenspace and domestic garden are indicative of fewer pedestrian supportive infra-structures such as pavements and of reduced land use mix, a walkability-associated factor. Another explanation is that the older adults of whom the WIIP7 sample is representative are less likely to use neighbourhood greenspace for walking than younger adults. A moderating effect of age would account for the contrasting results of Lachowycz and Jones' study of 165,424 adults, 73% of whom were of working age, across England [204]. It showed that greenspace proportion in neighbourhoods operationalised as Middle Super Output Areas (MSOAs) was positively associated with walking for recreation.

The findings of an observational study by Mytton *et al* using Health Survey for England (HSE) data corroborate and shed light on, and are themselves illuminated by the findings of the present study [201]. Mytton *et al*'s study showed that those living in neighbourhoods defined as the administrative units of MSOAs, with the highest proportion of greenspace excluding gardens, were more likely to be sufficiently active than those living in the least green neighbourhoods, echoing the results of the present study with respect to the WHO recommended physical activity level outcome for the WIIP7 study sample. In addition, Mytton *et al* found that the types of physical activities done to a sufficient level that were associated with greenspace were not typical "greenspace recreation". Instead, they included domestic activities such as gardening. Across English LAs, administrative areas with a higher proportion of greenspace also tend to have a higher proportion of garden area, as a consequence of lower residential densities. The findings of the present study suggest that had Mytton *et al* examined associations between proportion of greenspace plus domestic garden and different domains of physical activity, they would have found even stronger positive associations with gardening. Operationalising neighbourhoods as a 1-km radius around a participant's home Maas *et al* found that people with more agricultural greenspace in their neighbourhood gardened more frequently and for longer [205]. The authors posited that this was attributable to individuals living in areas with more agricultural greenspace being more likely to own larger, more labour-demanding gardens. Whilst a simple conclusion upon finding a positive relationship between neighbourhood greenspace and physical activity is that people living in greener areas spend more time walking, running or cycling around in public parks, Maas *et al*'s study illustrates the importance of further investigation, including longitudinal studies: individuals in greener neighbourhoods may be more physically active but incorporation of a park into a neighbourhood as a public health intervention may not necessarily elicit the desired behaviour.

It was also hypothesised that associations between physical activity outcomes and greenspace, and between depression and greenspace, would be stronger for smaller administratively-defined neighbourhoods. There was a tendency for this to be true for the physical activity outcomes, with stronger associations with the exposure when neighbourhoods were defined as the smaller CAS ward areas, or as OAs, than when defined as LAs. This suggests that greenspace that is closer to home is more relevant to physical activity outcomes. It may be that an individual living in a large LA with a very high proportion of greenspace could live in a very built-up area with little proximal and, therefore, accessible greenspace offering potential physical activity opportunities.

Greenspace and domestic garden were also investigated in the WIIP7 and HSE2008 study samples in relation to depression, based on the hypothesis that greater proportions of these land uses in the neighbourhood would be associated with lower odds of depression by affording potential for mental health-protective physical activity. As expected significantly lower odds of depression were found for those living in neighbourhoods, defined as any of the administrative geographies in WIIP7 and defined as LAs in HSE2008. Associations tended to be stronger, and were detectable across all levels of administratively-defined neighbourhoods for the WIIP7 study sample, when the exposure was specified as greenspace plus domestic garden rather than as greenspace alone. All significant associations were lost however for the WIIP7 study sample, after adjustment for individual level factors, suggesting that variation in depression between areas of different proportions of greenspace was attributable to differences in attributes of the individuals between areas. Although attenuated, significantly reduced odds of depression remained for those in the HSE2008 study sample living in LA-defined neighbourhoods with higher proportions of greenspace, and with higher proportions of greenspace plus domestic garden. However, additional adjustment for CAS ward-level deprivation, as indicated by Index of Multiple Deprivation quintiles, eliminated these associations. Thus, whilst the variation in odds of depression between areas of different proportions of greenspace was not wholly attributable to differences in the composition of the HSE2008 study sample between them, the analyses suggest that multiple deprivation, perhaps negatively associated with greenspace proportion – with greater deprivation associated with less greenspace – accounted for the reduced odds of depression associated with higher proportions of greenspace. As the Index of Multiple Deprivation contains an indicator of living environment which includes air quality, inclusion of this area level correlate in the multivariate logistic analysis may have resulted in collinearity with the greenspace variables: higher proportions of greenspace are likely to be indicative of better air quality. Thus adjustment for area level deprivation exclusive of a living environment factor may have not resulted in a loss of significant associations between greenspace and depression.

5.5 Summary

In summary, these results suggest that proportions of green space and domestic garden within the neighbourhood are independently and positively significantly associated meeting the WHO recommended physical activity levels. They are, however, inversely associated with time spent walking, suggesting that the greenspace may have a differential effect on domains of physical activity. Greenspace and availability of domestic garden are also associated with GHQ-12 depression but not necessarily independently of individual and other area level factors. Also, delineation of neighbourhoods at different administrative geographies affects associations: it appears that modelling of green space exposure within smaller units is more relevant to individuals' physical activity and depression outcomes.

5.6 Appendices

Appendix 5.1 Associations between quintiles of greenspace within administrative areas and WHOPA_I in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-48%)	2,806	43.22	1,472	1,227	1	REF	
		2(48-74%)	1,318	20.30	589	707	1.44	1.26-1.64	<0.001
		3(74-86%)	963	14.83	415	527	1.52	1.31-1.77	<0.001
		4(86-92%)	966	14.88	433	517	1.43	1.23-1.66	<0.001
		5(92-97%)	439	6.76	184	244	1.59	1.29-1.95	<0.001
Test for trend p < 0.001									
LA	London	1(5-21%)	366	12.54	204	141	1	REF	
		2(22-26%)	454	15.56	264	165	0.90	0.68-1.21	0.496
		3(27-34%)	526	18.03	272	235	1.25	0.95-1.65	0.114
		4(34-41%)	882	30.23	453	395	1.26	0.98-1.63	0.072
		5(41-59%)	690	23.65	332	349	1.52	1.17-1.98	0.002
Test for trend p < 0.001									
CAS Ward	England	1(1-30%)	2,499	38.49	1,270	1,144	1	REF	
		2(30-51%)	1,396	21.50	700	652	1.03	0.90-1.18	0.623
		3(51-77%)	1,320	20.33	597	707	1.31	1.15-1.50	<0.001
		4(77-92%)	817	12.58	334	463	1.54	1.31-1.81	<0.001
		5(92-99%)	460	7.09	192	256	1.48	1.21-1.81	<0.001
Test for trend p < 0.001									
CAS Ward	London	1(2-12%)	484	16.59	248	225	1	REF	
		2(12-19%)	527	18.06	288	214	0.82	0.64-1.05	0.121
		3(19-25%)	654	22.41	334	292	0.96	0.76-1.22	0.761
		4(25-38%)	572	19.60	315	232	0.81	0.63-1.04	0.099
		5(39-90%)	681	23.34	340	322	1.04	0.82-1.32	0.722
Test for trend non-significant (p>0.05)									
OA	England	1(0-4%)	1,790	27.63	866	861	1	REF	
		2(4-13%)	1,293	19.96	647	607	0.94	0.82-1.09	0.434
		3(13-29%)	1,139	17.58	573	537	0.94	0.81-1.10	0.443
		4(29-64%)	1,096	16.92	530	543	1.03	0.88-1.20	0.699
		5(64-100%)	1,160	17.91	468	669	1.44	1.24-1.67	<0.001
Test for trend p < 0.001									

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
OA	London	1(0-1%)	687	23.63	349	311	1	REF	
		2(1-5%)	661	22.74	328	308	1.05	0.85-1.31	0.638
		3(5-12%)	570	19.61	315	231	0.82	0.65-1.03	0.095
		4(12-27%)	490	16.86	261	211	0.91	0.72-1.15	0.421
		5(27-100%)	499	17.17	265	220	0.93	0.74-1.18	0.555
Test for trend non-significant (p>0.05)									

Appendix 5.2 Associations between quintiles of greenspace plus domestic garden within administrative areas and WHOPA_I in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-69%)	2,407	37.08	1,267	1,048	1	REF	
		2(69-84%)	1,542	23.75	725	780	1.30	1.14-1.48	<0.001
		3(84-92%)	863	13.29	369	480	1.57	1.34-1.84	<0.001
		4(92-95%)	1,039	16.00	458	560	1.48	1.27-1.71	<0.001
		5(95-98%)	641	9.87	274	354	1.56	1.31-1.87	<0.001
Test for trend p < 0.001									
LA	London	1(5-41%)	286	9.80	163	107	1	REF	
		2(42-52%)	540	18.51	299	212	1.08	0.80-1.46	0.615
		3(54-58%)	493	16.90	268	202	1.15	0.85-1.56	0.374
		4(61-68%)	618	21.18	309	291	1.43	1.07-1.92	0.015
		5(69-81%)	981	33.62	486	473	1.48	1.13-1.95	0.005
Test for trend p < 0.001									
CAS Ward	England	1(3-61%)	2,047	31.53	1,088	876	1	REF	
		2(61-74%)	1,593	24.54	769	784	1.27	1.11-1.45	0.001
		3(74-87%)	1,401	21.58	638	741	1.44	1.26-1.66	<0.001
		4(87-95%)	906	13.96	370	517	1.74	1.48-2.04	<0.001
		5(95-99%)	545	8.39	228	304	1.66	1.36-2.01	<0.001
Test for trend p < 0.001									
CAS Ward	London	1(3-36%)	345	11.82	202	125	1	REF	
		2(36-49%)	416	14.26	213	181	1.37	1.02-1.85	0.037
		3(49-59%)	618	21.18	335	253	1.22	0.93-1.61	0.158
		4(59-67%)	671	23.00	361	297	1.33	1.01-1.74	0.039
		5(67-96%)	868	29.75	414	429	1.67	1.29-2.17	<0.001

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
Test for trend p < 0.001									
OA	England	1(0-46%)	1,006	17.02	576	381	1	REF	
		2(46-59%)	1,223	20.70	619	561	1.37	1.15-1.63	<0.001
		3(59-68%)	1,135	19.21	545	563	1.56	1.31-1.86	<0.001
		4(68-82%)	1,489	25.20	677	786	1.76	1.49-2.07	<0.001
		5(82-100%)	1,056	17.87	426	610	2.16	1.81-2.59	<0.001
Test for trend p < 0.001									
OA	London	1(0-32%)	327	11.26	190	122	1	REF	
		2(32-44%)	415	14.29	246	144	0.91	0.67-1.24	0.554
		3(44-54%)	590	20.31	302	263	1.36	1.02-1.80	0.034
		4(54-64%)	690	23.75	354	312	1.37	1.04-1.80	0.023
		5(64-100%)	883	30.40	425	440	1.61	1.24-2.10	<0.001
Test for trend p < 0.001									

Appendix 5.3 Associations between quintiles of greenspace within administrative areas and TTW in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-48%)	2,649	42.65	1,794	855	1	REF	
		2(48-74%)	1,272	20.48	882	390	0.93	0.80-1.07	0.309
		3(74-86%)	934	15.04	642	292	0.95	0.81-1.12	0.568
		4(86-92%)	932	15.01	610	322	1.11	0.95-1.30	0.204
		5(92-97%)	424	6.83	312	112	0.75	0.60-0.95	0.016
Test for trend non-significant (p>0.05)									
LA	London	1(5-21%)	340	12.34	218	122	1	REF	
		2(22-26%)	430	15.60	288	142	0.88	0.65-1.19	0.407
		3(27-34%)	507	18.40	356	151	0.76	0.57-1.02	0.063
		4(34-41%)	815	29.57	548	267	0.87	0.67-1.14	0.306
		5(41-59%)	664	24.09	455	209	0.82	0.62-1.08	0.160
Test for trend non-significant (p>0.05)									
CAS Ward	England	1(1-30%)	2,375	38.24	1,603	772	1	REF	
		2(30-51%)	1,333	21.46	912	421	0.96	0.83-1.11	0.564
		3(51-77%)	1,269	20.43	879	390	0.92	0.80-1.07	0.274

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
		4(77-92%)	788	12.69	536	252	0.98	0.82-1.16	0.785
		5(92-99%)	446	7.18	310	136	0.91	0.73-1.13	0.404
Test for trend non-significant (p>0.05)									
CAS Ward	London	1(2-12%)	458	16.62	295	163	1	REF	
		2(12-19%)	494	17.92	327	167	0.92	0.71-1.21	0.563
		3(19-25%)	623	22.61	440	183	0.75	0.58-0.97	0.031
		4(25-38%)	538	19.52	360	178	0.89	0.69-1.16	0.407
		5(39-90%)	643	23.33	443	200	0.82	0.63-1.05	0.119
Test for trend non-significant (p>0.05)									
OA	England	1(0-4%)	1,707	27.54	1,141	566	1	REF	
		2(4-13%)	1,226	19.78	830	396	0.96	0.82-1.12	0.626
		3(13-29%)	1,087	17.54	732	355	0.98	0.83-1.15	0.784
		4(29-64%)	1,048	16.91	736	312	0.85	0.72-1.01	0.064
		5(64-100%)	1,130	18.23	792	338	0.86	0.73-1.01	0.069
Test for trend p < 0.05									
OA	London	1(0-1%)	657	23.93	440	217	1	REF	
		2(1-5%)	625	22.77	412	213	1.05	0.83-1.32	0.690
		3(5-12%)	530	19.31	374	156	0.85	0.66-1.08	0.185
		4(12-27%)	457	16.65	307	150	0.99	0.77-1.28	0.943
		5(27-100%)	476	17.34	324	152	0.95	0.74-1.22	0.698
Test for trend non-significant (p>0.05)									

Appendix 5.4 Associations between quintiles of greenspace plus domestic garden within administrative areas and TTW in the WIIP7 study sample.

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
LA	England	1(5-69%)	2,281	36.73	1,547	734	1	REF	
		2(69-84%)	1,467	23.62	1,006	461	0.97	0.84-1.11	0.629
		3(84-92%)	837	13.48	585	252	0.91	0.76-1.08	0.270
		4(92-95%)	1,006	16.20	669	337	1.06	0.91-1.24	0.457
		5(95-98%)	620	9.98	433	187	0.91	0.75-1.10	0.339
Test for trend non-significant (p>0.05)									
LA	London	1(5-41%)	271	9.83	167	104	1	REF	
		2(42-52%)	506	18.36	353	153	0.70	0.51-0.95	0.022
		3(54-58%)	472	17.13	308	164	0.86	0.63-1.17	0.321
		4(61-68%)	579	21.01	403	176	0.70	0.52-0.95	0.021
		5(69-81%)	928	33.67	634	294	0.74	0.56-0.99	0.040
Test for trend non-significant (p>0.05)									
CAS Ward	England	1(3-61%)	1,936	31.17	1,294	642	1	REF	
		2(61-74%)	1,526	24.57	1,046	480	0.92	0.801.07	0.287
		3(74-87%)	1,347	21.69	944	403	0.86	0.74-1.00	0.050
		4(87-95%)	877	14.12	591	286	0.98	0.82-1.16	0.774
		5(95-99%)	525	8.45	365	160	0.88	0.72-1.09	0.245
Test for trend non-significant (p>0.05)									
CAS Ward	London	1(3-36%)	331	12.01	204	127	1	REF	
		2(36-49%)	392	14.22	253	139	0.88	0.65-1.20	0.419
		3(49-59%)	574	20.83	401	173	0.69	0.52-0.92	0.011
		4(59-67%)	634	23.00	442	192	0.70	0.53-0.92	0.011
		5(67-96%)	825	29.93	565	260	0.74	0.57-0.96	0.026
Test for trend p < 0.05									
OA	England	1(0-46%)	940	16.67	606	334	1	REF	
		2(46-59%)	1,161	20.59	775	386	0.90	0.75-1.08	0.273
		3(59-68%)	1,079	19.13	751	328	0.79	0.66-0.95	0.014
		4(68-82%)	1,439	25.52	976	463	0.86	0.72-1.02	0.090
		5(82-100%)	1,020	18.09	718	302	0.76	0.63-0.92	0.005
Test for trend p < 0.01									

Admin unit	Quintile deriv	Quintile	N	%	No	Yes	OR	CI	p
OA	London	1(0-32%)	308	11.23	198	110	1	REF	
		2(32-44%)	388	14.15	255	133	0.94	0.69-1.28	0.693
		3(44-54%)	560	20.42	374	186	0.90	0.67-1.20	0.457
		4(54-64%)	643	23.44	438	205	0.84	0.63-1.12	0.240
		5(64-100%)	844	30.77	592	252	0.77	0.58-1.01	0.058
Test for trend $p < 0.05$									

6 Study 3: Associations between neighbourhood environmental quality and (A) physical activity, and (B) GHQ-12 depression

6.1 Introduction

6.1.1 Chapter overview

This chapter focuses on associations of indicators of environmental quality with physical activity in an occupational-specific sample of older adults. It also focuses on associations of indicators of environmental quality with depression in this sample of older adults as well as in a nationally representative survey of adults. The environmental quality indicators considered are an objective measure – a multiple physical environment-specific index of deprivation – and overall resident satisfaction with parks and open spaces at local authority (LA) level as measured in a local government-administered England-wide survey, a perception measure. Also, the chapter explores the extent to which these area level exposure variables, and the greenspace variables – the focus of Chapter 5 – contribute to area level variation in physical activity and GHQ-12 depression outcomes.

6.1.2 Research questions

The research questions are as follows:

- Is objective environmental quality associated with physical activity outcomes in the Whitehall II phase 7 (WIIP7) [182] study sample, and with a score of four or more on the 12-item General Health Questionnaire (GHQ-12 depression) in the WIIP7 and the 2008 Health Survey for England (HSE2008) [183] study samples?
- How does scaling the exposure from Census Area Statistics (CAS) ward level [206], at which aggregate environmental data was collected, to LA level affect these associations?
- Is perceived environmental quality at LA level, associated with physical activity outcomes in WIIP7, and with the GHQ-12 depression outcome in WIIP7 and HSE2008?
- Are associations independent of individual-level sociodemographic factors and of area deprivation?
- Within each study population how much variation in physical activity and GHQ-12 depression outcomes between areas exists and to what extent is the variation attributable to the measured physical environmental exposures variables?

6.1.3 Objectives for Study 3A

The objectives for Study 3A are:

- to determine whether odds of the physical activity outcomes are lower for an individual in the WIIP7 sample living within neighbourhoods, defined as (a) a CAS ward and (b) a LA where objective environmental quality is worse than the median for the UK, for England and for London

- to determine whether odds of physical activity outcomes are lower for an individual in the WIIP7 sample living within a neighbourhood, defined as LA where perceived environmental quality is below the median percentage for England and for London
- to determine whether the associations are affected by individual-level sociodemographic factors and area deprivation
- to identify the amount of variation in physical activity outcomes between areas in each study population and to determine the contribution of each exposure variable to this variation

6.1.4 Hypotheses for Study 3A

Physical activity outcomes are associated with objective and perceived environmental quality, with individuals living in neighbourhoods where objective and perceived environmental quality is lower having lower odds of the physical activity outcomes.

Associations between objective environmental quality and the physical activity outcomes are stronger for CAS ward-defined neighbourhoods than LA-defined ones because the smaller units are more representative of the area to which an individual is exposed.

The strength of associations between environmental quality exposures and physical activity outcomes is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and area level deprivation.

Variation exists in physical activity outcomes between LAs, and to a greater extent between CAS wards, and each exposure variable contributes to this inter-area variation.

6.1.5 Objectives for Study 3B

The objectives for Study 3B are:

- to determine whether odds of GHQ-12 depression are higher for an individual in the WIIP7 sample living within a neighbourhood, defined as CAS ward and as LA where objective environmental quality is worse than the median score for the UK, for England and for London
- to determine whether odds of GHQ-12 depression are higher for an individual in the WIIP7 sample or in the HSE2008 sample living within a neighbourhood, defined as LA where perceived environmental quality is below the median percentage for England and for London
- to determine whether these associations are affected by individual-level sociodemographic factors and area deprivation
- to identify the amount of variation in depression between areas in each study population and to determine the contribution of each exposure variable to this variation

6.1.6 Hypotheses for Study 3B

GHQ-12 depression is associated with objective and perceived environmental quality, with individuals living in neighbourhoods where objective and perceived environmental quality is lower having higher odds of GHQ-12 depression.

For the WIIP7 sample associations between objective environmental quality and depression are stronger for CAS ward-defined neighbourhoods than local LA-defined ones because the smaller units are more representative of the area to which an individual is exposed.

The strength of associations between environmental quality exposures and GHQ-12 depression is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and area deprivation.

There is variation in GHQ-12 depression caseness between LAs, and to a greater extent between CAS wards, and each exposure variable contributes to this inter-area variation.

6.2 Methods

6.2.1 Study samples

Study 3A comprised the WIIP7 study sample (see 3.2.1.2) in the investigation of associations between environmental quality and physical activity, whereas Study 3B, investigating associations between environmental quality and depression comprised both the WIIP7 study sample and the HSE2008 study sample (see Section 3.2.2.2). The WIIP7 study sample was drawn from the seventh wave, conducted in 2004/5, of the Whitehall II study, a longitudinal study of civil servants to examine the social determinants of health [182]. The sample comprised 6885 individuals with a mean age of 61.2 years ± 6.0. The HSE2008 study population (N=14,221) from the HSE conducted in 2008 which drew a nationally representative general population sample of adults living in households in England through multi-stage stratified probability sampling [183]. At 48.9 years ± 18.6 the mean age of this sample was lower than that of the WIIP7 sample, an important consideration in comparisons of associations with outcomes in the two samples.

6.2.2 Variables

6.2.2.1 Exposure variables

Derivation of the exposure variables is detailed in Chapter 3 (Section 3.3.2). Briefly, the exposures were as follows:

- objective neighbourhood environmental quality (OEQ) as indicated by the Multiple Environmental Deprivation Index (MEDIx) [207]; and

- perceived neighbourhood environmental quality (PEQ) as indicated by local authority (LA)-level percentage satisfaction with parks and open spaces in the 2008 Place Survey [208].

OEQ exposure variables were derived into dichotomous variables with the exposed comprising individuals in the WIIP7 and HSE2008 study samples living in neighbourhoods of worse objectively measured environmental deprivation relative to the median level for all neighbourhoods – operationalised as administrative units of LAs or CAS ward – for all in the UK, for all in England only or for all in London only. Thus, the benchmark for poor quality for each exposure was independent of the study samples, and for each exposure the same benchmark was used for both study samples.

PEQ exposures, like OEQ exposures, were derived into dichotomous variables. The exposed, with respect to the PEQ variables, comprised individuals in the WIIP7 and HSE2008 study samples living in neighbourhoods of lower LA-level satisfaction relative to the median level for all neighbourhoods – operationalised as the administrative units of LAs – for all in England or for all in London only. Again, this meant that the benchmarks for poor quality were independent of the study samples and the same benchmarks were used for both study samples.

6.2.2.2 Outcome variables

Three physical activity outcomes were specified for Study 3A;

1. meeting the World Health Organisation recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E);
2. meeting the World Health Organisation recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I); and
3. being in the top tertile of WIIP7 for time spent walking per week (TTW).

Chapter 3 (Section 3.3.3.2) provides the rationale for the specification of these variables as such. The WHO recommended physical activity level [47] was defined as follows:

- at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or at least 75 minutes of vigorous intensity aerobic physical activity throughout the week, or an equivalent combination of moderate and vigorous-intensity activity

GHQ-12 depression was the outcome specified for Study 3B, and was defined as a score of four or more on the General Health Questionnaire (GHQ-12) [57] as detailed in Chapter 3 (Section 3.3.3.1).

6.2.2.3 Correlates

Individual level correlates entered into the regression models in Study 3A were sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), car availability (Car available (ref)/No car available), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white), and in Study 3B they were sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white) for WIIP7 and sex (Male (ref)/ Female), age (16y-34y (ref)/ 35y-54y/ 55y+), economic activity (Employed (ref)/ Unemployed/ Retired/ Other), income (<=£14918 (ref)/ >£14918-£31871/>£31871), occupational social class (Managerial and professional (ref)/ Intermediate/ Routine and manual/ Other) and marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) for HSE2008. Adjustment was also made for a contextual factor, area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level, capturing area-level characteristics in terms of income, employment, health, education, barriers to housing and services, living environment, and crime [197]. Owing to the geographical coverage of IMD2004, adjustment for area deprivation in the WIIP7 sample was restricted to the 6677 (97.0%) participants resident in England. Details of the identification and selection of these factors as correlates for which adjustment was necessary are provided in Chapter 4 (Section 4.4).

6.2.3 Statistical analysis of associations between neighbourhood environmental quality and physical activity and GHQ-12 depression outcomes

Multivariate logistic regression was the method selected to model statistical associations in examination of possible relationships with the neighbourhood environmental quality exposures and the outcomes for reasons outlined in Chapter 3 (Section 3.4). In the first instance, bivariate logistic regression models were specified which included only the outcome variable and exposure variable of interest to determine the presence and strength of significant associations. Subsequently, for models in which significant associations were detected, adjustment was made for correlates – firstly individual level sociodemographic factors and then additionally area deprivation – which could confound the relationship. This showed the extent to which the relationships between environmental quality exposures and outcomes were independent of other factors. Thus, three models were constructed; Model 1, the unadjusted model was univariate analysis between the exposure and the outcome; Model 2 was multivariate analysis adjusting for all individual factors significantly associated with the outcome in the study sample; and Model 3 adjusted for area deprivation, in addition to the individual factors. Results were computed as odds ratios alongside their 95% confidence intervals to indicate with 95% confidence the range within which the true value of the odds ratio lay.

In the case of the OEQ exposure, an odds ratio indicated the odds of the outcome in those exposed to, or living in, low OEQ neighbourhoods relative to those living in higher OEQ neighbourhoods. Similarly, for the PEQ exposure, an odds ratio indicated the odds of the outcome in those exposed to low PEQ neighbourhoods relative to those exposed to higher OEQ neighbourhoods. As explained in Chapter 3 in Section 3.4 it was necessary to compliment this single level logistic regression with a multilevel approach in order to increase the accuracy of quantification of the strength of significant associations between neighbourhood environmental exposures and outcomes. Multilevel modelling enabled the identification the level of unaccounted area level variation and thereby the determination of the relative contribution of the specified area level, or neighbourhood, exposure variables.

Each binary outcome variable was modelled with mixed – fixed and random – effects logistic regression. For GHQ-12 depression in the HSE2008 sample it was only possible to specify two-level models for the binary outcome that had the structure of individuals (Level 1) living in LAs (Level 2). For physical activity as the outcome in the WIIP7 population, however, three levels were included in the models: individuals (Level 1) were nested within CAS wards (Level 2) which were nested within LAs (Level 3). Previous studies suggested that there was considerable variation in GHQ-12 depression at the level of the household [209] but less at higher geographies [64]. In the WIIP7 sample the majority of individuals resided in a unique output area (OA), although there were 530 instances where two individuals shared an OA, 67 where three shared an OA and eight where four shared an OA. However, there were only 231 instances where two individuals shared a postcode and nine where three shared a postcode. This indicated that the majority of those who shared an OA were not living with, or a very proximal neighbour of, another participant in the study. Given that OA level in depression would not necessarily operate at household level, and that it could encompass a wider area, it was deemed valuable to investigate variation in GHQ-12 depression using four-level models comprising individuals (Level 1), nested within OAs (Level 2) nested within CAS wards (Level 3) which were nested within LAs (Level 4). For the physical activity outcomes in WIIP7, multilevel modelling was restricted to the WHOPA_E outcome as multilevel modelling for the other physical activity outcomes was beyond the scope of this thesis. For both the WIIP7 and HSE2008 samples area-level environmental exposure variables were limited to those that had been identified in single level logistic regression as being significantly and independently associated with the specified outcome, and therefore, for which it was useful to identify their contribution relative to unmeasured area level factors.

Whilst greenspace was the focus of the Study 2 in Chapter 5, multilevel models with greenspace as the exposure are reported in this chapter alongside multilevel models for the environmental quality exposures found to be significantly and independently associated with the outcomes. Several models were constructed for the physical activity and GHQ-12 depression outcomes in WIIP7 and for the GHQ-12 depression outcome in HSE2008. The first model, an empty or unconditional model without any exposure variables, was specified to decompose the amount of variance that existed at the area-level. The second contained individual-level variables and further models each contained a specific area-level variable in addition to the individual-level ones. The final model contained all individual and area-level variables. Random effects were measured as an intra-class correlation (ICC) which is the correlation between randomly chosen pairs of individuals in the same group. The ICC gives the variance partition coefficient (VPC), the proportion of total variance that is due to differences between groups, ranging from 0 (when there are no inter-group differences) to 1 (when there are no intra-group differences). An ICC of 0.2, for example, indicates that 20% of the variation is between groups and 80% within. Additionally, the LR test versus logistic regression test statistic was reported which compares the fit of the random effects multilevel model to that of the fixed effects single level logistic regression. If the p value is less than 0.05

the null hypothesis that there is no difference between the random effects and the fixed effects models can be rejected and a multilevel model used because it is a better fit. If it is more than 0.05, multilevel modelling is not necessary. However, the test statistic is conservative so it is recommended that the p value is halved, akin to one-tailed p-values, before a decision on whether or not to reject the hypothesis is taken [210].

6.3 Results

6.3.1 Results for Study 3A: Associations between environmental quality and physical activity outcomes in WIIP7

6.3.1.1 Objective environmental quality and physical activity in WIIP7

Meeting the WHO recommended physical activity level excluding walking as a contributory component (WHOPA_E) in the WIIP7 sample was associated with objective environmental quality (OEQ), as indicated by the (Multiple Environmental Deprivation (MEDIx) median score whether this was derived at LA level or CAS ward level, and whether it was UK, England or London-based, as shown in Table 6.1. Individuals living in LAs or CAS wards with worse OEQ had significantly lower odds of meeting the recommended level specified as WHOPA_E. The lowest odds ratio, at 0.61 ($p<0.001$), was detected in Model 1, before adjustment for correlates, when OEQ exposure variable was CAS ward-based and derived from all UK CAS wards. After adjustment for individual level factors, in Model 2, and additional adjustment for area deprivation, in Model 3, there was evidence that associations remained significant although slightly attenuated. For example, the odds ratio when the OEQ was derived for CAS wards and UK-based increased from 0.61 ($p<0.001$) in Model 1, to 0.73 ($p<0.001$) in Model 2 and to 0.77 ($p<0.001$) in Model 3. Whereas at LA-level with UK and England-based OEQ the significant associations remained in Model 3 with additional adjustment for area deprivation as evidenced by significant odds ratios of 0.86 ($p=0.012$) and 0.87 ($p=0.024$) respectively, the association between WHOPA_E and the London-based MEDIx median score was eliminated. At CAS ward level, however, all significant associations remained in Model 3, even with the London-based MEDIx median score. In line with the findings with the WHOPA_E outcome, the outcome which included walking if reported as brisk, WHOPA_I, was significantly associated with multiple environmental deprivation, whether the MEDIx median was derived at LA level or CAS ward level, and whether it was UK, England or London-based (Table 6.2). Individuals living in LAs or CAS wards where the MEDIx score was equal to or more than the median, indicative of poorer quality, had significantly lower odds of this outcome. The lowest odds of the WHOPA_I outcome were found for individuals in the WIIP7 study sample living in poorer quality neighbourhoods as indicated by the CAS ward-based OEQ variable derived from all UK CAS wards (OR=0.64; $p<0.001$) in Model 1. At LA-level, all associations remained significant although attenuated in Model 2 after adjustment for individual level factors. However, at CAS ward level odds ratios remained significant only when medians MEDIx scores were UK and England-based – at 0.78 ($p<0.001$) and 0.83 ($p=0.001$) respectively – but not when they were London-based. In Model 3, significant associations remained between WHOPA_I only for the CAS ward-level UK based MEDIx median (OR=0.83; p=0.001) and the CAS ward-level England based MEDIx median (OR=0.88; p=0.020). Significant associations between the objective measure of environmental quality as MEDIx score and physical activity were limited to outcomes specified as meeting WHO recommended levels. Even before adjustment for correlates, being in the top tertile for time spent walking per week (TTW) was not associated with this objective measure of environmental quality whether this was derived at LA level or CAS ward level, and whether it was UK, England or London-based (Table 6.3).

Table 6.1 Associations between objective environmental quality (OEQ) and WHOPA_E, before and after adjustment for individual level sociodemographic factors and area level deprivation in the WIIP7 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
OEQ	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
Admin area/ OEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
LA/ UK	0.66	0.60-0.74	<0.001	0.81	0.72-0.91	<0.001	0.86	0.76-0.97	0.012
LA/ England	0.68	0.61-0.76	<0.001	0.82	0.73-0.92	0.001	0.87	0.77-0.98	0.024
LA/ London	0.78	0.65-0.93	0.008	0.82	0.68-1.00	0.048	0.87	0.71-1.06	0.173
CASward/ UK	0.61	0.54-0.68	<0.001	0.73	0.65-0.82	<0.001	0.77	0.68-0.87	<0.001
CASward/ England	0.64	0.58-0.72	<0.001	0.78	0.69-0.88	<0.001	0.82	0.73-0.93	0.001
CASward/ London	0.73	0.60-0.88	<0.001	0.79	0.65-0.96	0.016	0.81	0.66-1.00	0.045
	[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

Table 6.2 Associations between objective environmental quality (OEQ) and WHOPA_I, before and after adjustment for individual level sociodemographic factors and area level deprivation in the WIIP7 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
OEQ	✓			✓			✓		
Individual factors				✓			✓		
Area deprivation							✓		
Admin area/ OEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
LA/ UK	0.69	0.62-0.76	<0.001	0.84	0.76-0.94	0.002	0.89	0.80-1.00	0.054
LA/ England	0.71	0.64-0.78	<0.001	0.87	0.78-0.97	0.013	0.92	0.82-1.02	0.142

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
LA/ London	0.79	0.67-0.92	0.003	0.83	0.70-0.97	0.023	0.87	0.73-1.04	0.122
CASward/ UK	0.64	0.58-0.71	<0.001	0.78	0.70-0.87	<0.001	0.83	0.74-0.93	0.001
CASward/ England	0.68	0.61-0.75	<0.001	0.83	0.75-0.92	0.001	0.88	0.78-0.98	0.020
CASward/ London	0.85	0.73-1.00	0.044	0.93	0.78-1.10	0.397	0.97	0.82-1.15	0.734

[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 6.3 Associations between objective environmental quality (OEQ) and TTW in the WIIP7 study sample.

Admin area/ OEQ derivation area	OR	CI	p
LA/ UK	0.93	0.84-1.04	0.188
LA/ England	0.98	0.88-1.10	0.764
LA/ London	1.13	0.95-1.33	0.162
CASward/ UK	0.97	0.92-1.03	0.285
CASward/ England	0.98	0.92-1.04	0.460
CASward/ London	1.04	0.94-1.16	0.435

6.3.1.2 Perceived environmental quality and physical activity in WIIP7

Significant associations between the subjective measure as perceived environmental quality (PEQ), as LA-level satisfaction with parks and open spaces, were not found for any of the three physical activity outcomes, regardless of whether the median score was England or London-based (Appendix 6.1).

6.3.2 Results for Study 3B: Associations between environmental quality and GHQ-12 depression in the WIIP7 and HSE2008 study samples

6.3.2.1 Objective environmental quality and GHQ-12 depression in WIIP7

In the WIIP7 study sample, living in LAs with a Multiple Environmental Deprivation Index (MEDIx) score equal to or greater than the median – poorer objective environmental quality (OEQ) – was associated with significantly higher odds of GHQ-12 depression (OR=1.39, p<0.01) but only when the median was London-based (Table 6.4). After adjustment for individual-level sociodemographic factors in Model 2 this significant association remained, albeit reduced (OR=1.31, p<0.05). It also remained significant in Model 3, additionally adjusting for LSOA-level IMD2004 (OR=1.32, p<0.05), supporting the presence of an association between OEQ as an indicator of physical environmental-specific deprivation and GHQ-12 depression independent of the generic area-level deprivation indicator, IMD2004. In contrast to OEQ at LA level, OEQ at CAS ward level was significantly associated with

GHQ-12 depression when the median was UK-based and England-based but not London-based. Individuals living in CAS wards with poorer OEQ than the UK-based median had significantly higher odds of having the GHQ-12 depression outcome (OR=1.18, p<0.05), as did those living in CAS wards with poorer OEQ than the England-based median (OR=1.15, p<0.05). However, after adjustment for individual level factors, these associations between CAS ward-level OEQ and GHQ-12 depression were no longer significant.

6.3.2.2 Objective environmental quality and GHQ-12 depression in HSE2008

LAs were the only neighbourhood operationalisations for HSE2008 and significant associations were identified between OEQ and GHQ-12 depression in this study sample, with individuals living in LA-defined neighbourhoods with poorer OEQ than the UK-based and England-based medians having higher odds of depression at 1.32 (p<0.001) and 1.26 (p<0.001), respectively (Table 6.5). Although attenuated these significantly higher odds of GHQ-12 depression remained in Model 2 after adjustment for individual level factors for both the UK-based medians (OR=1.19, p<0.01) and the England-based medians (OR=1.20, p<0.01). The addition of area deprivation as a correlate in Model 3, however, rendered associations not significant.

Table 6.4 Associations between objective environmental quality (OEQ) and GHQ-12 depression, before and after adjustment for individual level sociodemographic factors and area level deprivation in the WIIP7 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
OEQ		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
Admin area/ OEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
LA/ UK	1.13	0.98-1.30	0.091	n/a	n/a	n/a	n/a	n/a	n/a
LA/ England	1.10	0.95-1.27	0.198		n/a	n/a	n/a	n/a	n/a
LA/ London	1.39	1.13-1.72	0.002	1.30	1.05-1.61	0.015	1.32	1.05-1.66	0.017
CASward/ UK	1.18	1.03-1.35	0.019	0.99	0.85-1.15	0.881	n/a	n/a	n/a
CASward/ England	1.15	1.00-1.33	0.047	0.97	0.84-1.13	0.724	n/a	n/a	n/a
CASward/ London	1.18	0.95-1.45	0.133	n/a	n/a	n/a	n/a	n/a	n/a

[a]Adjustment for sex, age, economic activity, marital status and ethnicity;

[b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 6.5 Associations between objective environmental quality (OEQ) and GHQ-12 depression, before and after adjustment for individual level sociodemographic factors and area level deprivation in the HSE2008 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
OEQ		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
OEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
UK	1.32	1.19-1.46	<0.001	1.19	1.06-1.33	0.003	1.11	0.98-1.25	0.091
England	1.26	1.14-1.39	<0.001	1.20	1.07-1.34	0.002	1.12	0.99-1.26	0.072
London	1.18	0.90-1.56	0.240	n/a	n/a	n/a	n/a	n/a	n/a

[a]Adjustment for sex, age, economic activity, income, occupational social class and marital status; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

6.3.2.3 Perceived environmental quality and GHQ-12 depression in WIIP7

In the WIIP7 sample, living in LAs with lower than the median level of satisfaction with parks and open spaces – poorer perceived environmental quality (PEQ) – was not associated with the GHQ-12 depression outcome (Appendix 6.2).

6.3.2.4 Perceived environmental quality and GHQ-12 depression in HSE2008

In contrast to the findings in the WIIP7 study sample, significant associations between PEQ and the GHQ-12 depression outcome were found in the HSE2008 sample, with those living in LAs across England with lower than the median PEQ having higher odds of this outcome (OR=1.23; p<0.001), and those living in LAs across London with lower than the London-specific PEQ median also having higher odds (OR=1.41; p<0.05) (Table 6.6). The association found between England-based PEQ median and depression in the HSE2008 sample reduced but remained significant after adjustment for individual level sociodemographic factors (OR=1.13, p<0.05) in Model 2, whereas that between London-based satisfaction and GHQ-12 depression became insignificant. After additional adjustment for area deprivation in Model 3, the significant association with the England-based PEQ median that had remained in Model 2 was rendered not significant.

Table 6.6 Associations between perceived environmental quality (PEQ) at LA level and GHQ-12 depression, before and after adjustment for individual level sociodemographic factors and area level deprivation in the HSE2008 study sample.

	MODEL 1			MODEL 2[a]			MODEL 3[b]		
PEQ		✓			✓			✓	
Individual factors					✓			✓	
Area deprivation								✓	
PEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
England	1.23	1.11-1.36	0.001	1.13	1.00-1.26	0.042	1.07	0.96-1.21	0.228
London	1.41	1.06-1.87	0.017	1.20	0.85-1.69	0.310	n/a	n/a	n/a

[a]Adjustment for sex, age, economic activity, income, occupational social class and marital status; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

6.3.3 Area level variation in outcomes attributable to area level exposure variables

Multilevel modelling was performed to examine area level the variation in outcomes in the physical activity outcome of WHOPA_E and in the GHQ-12 depression outcome in the study samples and the extent to which this variation was attributable to the specified exposure variables. Reasons for the restriction of analysis of area level variation to only one of the physical activity outcomes are given in Section 6.2.3. Variation in the WHOPA_E and in the GHQ-12 depression outcomes in the WIIP7 study sample was examined between local authorities (LA), between CAS wards and between output areas (OAs). Between LA variation in the GHQ-12 depression outcome was examined in the HSE2008 study sample.

6.3.3.1 Area level variation in physical activity attributable to area level exposure variables in WIIP7

The results of multilevel modelling of the physical activity outcome, WHOPA_E, in the WIIP7 study sample are presented in Appendix 6.3. The total variation in the WHOPA_E outcome in the WIIP7 study sample that existed between LA areas was only 2.52% as indicated by the intra-class correlation (ICC) of 0.0252 in Model 1 but between CAS wards the variation was slightly higher at 3.86%. Adjustment for individual level factors, in Model 2, reduced between area variation to 0.87% and 2.34% in LAs and CAS wards, respectively, indicating that quite a large proportion of the small amount of variation in WHOPA_E between areas was attributable to variation in the characteristics of individuals between areas. The fact that the reduction in the ICC upon adjustment for individual factors was smaller for CAS wards is suggestive of a lesser role of individual level variation in WHOPA_E between these areas. The LR test versus logistic regression test statistics in both Models 1 and 2 were statistically significant indicating that multi- rather than single-level modelling was more appropriate. Addition of the measured exposure variables separately to subsequent models reduced between area variation in WHOPA_E further, and the greatest reduction was found in Model 4 when the UK-based OEQ variable was added: variation reduced from 0.87% to 0.37% between LAs and from 2.34% to 1.73% between CAS wards. This was greater than the reduction found either for the green space plus domestic garden (GS+DG) variable or the IMD2004 area deprivation indicator added separately, in Models 3 and 5 respectively. In Model 3 variation reduced from 0.87% to 0.61% between LAs and from 2.34% to 2.15% between CAS wards, and in Model 5 it reduced from 0.87% to 0.65% between LAs and from 2.34% to 2.27% between CAS wards. Addition of all variables simultaneously into the final models resulted in exactly or almost exactly the same levels of inter area variation as when the OEQ variable had been added alone: after adjustment for all measured environmental exposures the residual variations in WHOPA_E between LAs and between CAS wards were 039% and 1.73% respectively. Although perhaps negligible between LAs, there was variation between CAS wards areas for which the measured variables did not account. However, the LR test versus logistic regression test statistics were non-significant for all models with the exception of Models 1 and 2, suggesting that multilevel modelling, in taking account of random effects – non-measured environmental exposures – would not increase the accuracy of the estimates of the contributions of individual level and of area level factors as predictors of the outcome.

6.3.3.2 Area level variation in GHQ-12 depression attributable to area level exposure variables in HSE2008

Appendix 6.4 details the results of multilevel modelling of GHQ-12 depression in the HSE2008 study sample. In the HSE2008 study sample in which individuals were only identifiable to LA level, the total variation in the GHQ-12 depression outcome that existed between these areas was only 1.98% as indicated by the ICC of 0.0198 in Model 1. In Model 2, adjustment for individual level factors produced a marked reduction in the between LA variation in GHQ-12 depression to 0.64%, indicating a considerable proportion was attributable to variation in the characteristics of individuals between areas. Area variation in GHQ-12 depression was further reduced with addition of the measured exposure variables separately to subsequent models. Variation in GHQ-12 depression between LAs reduced from 0.64% to 0.34% in Model 3 with addition of the green space plus domestic garden (GS+DG) variable, to 0.39% in Model 4 with addition of the UK-based OEQ variable and to 0.43% in Model 5 with addition of the England-based PEQ variable However, by far the greatest reduction with addition of a single area level variable was found in Model 6. In Model 6, with addition of the IMD2004 area deprivation indicator, variation in GHQ-12 depression between LAs in the HSE2008 study sample reduced from 0.64% to 0.18%, suggesting that this factor played the largest role of the measured area level exposure variables in the between LA variation. Addition of all area-level exposure area variables at once, in Model 7, reduced the residual area variation to zero, indicating that all factors accounting for variation between LAs had been included. However, the LR test versus logistic regression test statistic was only significant for Model 1 which suggested that given the very small amount of variation between areas, single level modelling of GHQ-12 depression in this study sample would produce sufficiently accurate estimates of the relative contribution of both individual and area level factors as predictors of the outcome.

6.3.3.3 Area level variation in GHQ-12 depression attributable to area level exposure variables in WIIP7

The results of multilevel modelling of the GHQ-12 depression outcome in the WIIP7 study sample are presented in Appendix 6.5. In contrast to the between LA variation in GHQ-12 depression in the HSE2008 study sample of just under 2%, there was no variation between LAs in this outcome in the WIIP7 study sample as indicated by the ICC of <0.0001 in Model 1 before adjustment for individual level sociodemographic factors. However, in Model 1 4.36% of the variation in GHQ-12 depression was attributable to differences between CAS wards. This between CAS ward variation reduced only to 4.27% after accounting for individual level factors in Model 2, indicating that there was more homogeneity within CAS wards than within LAs in terms of the GHQ-12 depression outcome and that individual factors did not account for a substantive proportion of the variation between CAS wards. The ICC computed when the empty model nested individuals nested within OAs, within CAS wards, within LAs, was the same as when the empty model nested individuals only within CAS wards within LAs. This meant that the variation between OAs was entirely attributable to variation between the higher administrative geographies of CAS wards. Given this result, subsequent models were constructed to only contain three levels; individuals nested in CAS wards nested in LAs. In Model 3, with addition of the IMD2004 area deprivation indicator, variation in GHQ-12 depression between CAS wards in the WIIP7 study sample was not reduced, indicating that none of the variation was attributable to this measured environmental exposure variable. The finding that the LR test versus logistic regression test statistics, were non-significant for all models, indicated that single level modelling of the GHQ-12 outcome in this study sample was adequate and would not significantly overestimate the associations of individual level and measured environmental exposure variables by taking no account for random effects.

6.4 Discussion

The focus of this chapter was the examination of associations of both objective and perceived overall environmental quality with physical activity and GHQ-12 depression outcomes. The major aims were to investigate how operationalisation of neighbourhood affected associations and whether these associations were independent of individual-level sociodemographic factors and area deprivation. Also, there was an exploration of the extent to which the outcomes varied between areas and how much of this variation was attributable to the specified area level exposures.

Lower odds of meeting physical activity targets were hypothesised for those exposed to worse objective environmental quality (as indicated by an objective measure of multiple environmental deprivation, MEDIx). When the physical activity outcome was specified as meeting the WHO recommended physical activity level, with (WHOPA_I) or without (WHOPA_E) brisk or fast walking as a contributory component, this hypothesis proved correct; individuals living in worse OEQ neighbourhoods defined as either LAs or CAS wards had lower odds of being physical active even after adjusting for individual level factors and a generic measure of area deprivation. It may be that worse environmental quality deters people from undertaking the moderate to vigorous physical activity necessary to achieve recommended levels of physical activity. Odds tended to be lower where neighbourhoods were defined using smaller CAS wards rather than LAs suggesting that environmental quality within areas closer to individuals' homes had a greater influence on these physical activity outcomes and supporting the hypothesis that smaller units are more representative of the area to which an individual is exposed. When the physical activity outcome was being in the top tertile for time spent walking per week (TTW), no significant associations with OEQ were found. Whilst data on the purpose and location of WIIP7 participants' reported walking was not available, it can be assumed that almost all of the walking undertaken by those with the TTW outcome was undertaken outside because participants were asked to report walking outside the home or workplace. Thus, the finding that OEQ as indicated by MEDIx, an indicator of outdoor environmental quality, was associated with meeting recommended physical activity levels but not specifically with walking suggests that outdoor environmental quality may be associated with predisposition to relatively high intensity outdoor physical activities such as jogging, cycling and gardening. Alternatively, individuals living in areas of poorer environmental quality may be less inclined to leave home in order to expend energy.

The complete lack of associations identified between the physical activity outcomes and perceived environmental quality as indicated by a LA-level measure of local residents' satisfaction with parks and open spaces implies that local neighbourhood residents' satisfaction with greenspace may not reflect those of WIIP7 participants or does simply not impact on their likelihood of being physically active. Alternatively, it suggests that other factors were more important in determining use of such spaces for physical activity, that physical activity was not associated with such spaces in the WIIP7 sample, or that LAs did not constitute appropriate representations of neighbourhood in which environmental quality perceptions would affect physical activity. Other studies have found perception of overall environmental quality of neighbourhoods to be associated with transport-related physical activity [83][84]. In the present study transport-related and non-transport related physical activity could not be distinguished and analysed as separate outcomes. Had this been possible, significant associations between the perceived environment measure and transport-related physical activity may have emerged.

It was hypothesised that physical activity, specified as meeting the WHO recommended level would vary between areas but that most variation would be between individuals because individual factors are more important determinants of physical activity [7]. Some variation between areas, however, was expected because individuals with similar characteristics would be clustered within areas and also because areas would differ in their physical environmental attributes. Additionally it was expected that whilst variation would be detected between LAs and between CAS wards, a greater amount would be between the latter because a large proportion of any of the between area variation would be attributable to clustering of individuals, and their characteristics, and these would vary more between CAS wards than between LAs. Also, any environmental factors influencing physical activity outcomes would operate within smaller spatial units, as environmental factors would vary more between CAS wards than between LAs. As expected, with regard to meeting the WHO recommended physical activity level exclusive of brisk walking, variation existed between LAs, and to a greater extent between CAS wards but variation in physical activity between areas was low relative to that between individuals, supporting the hypothesis that individual level factors were more important than area-level factors in relation to outcomes. Addition of the OEQ exposure to multilevel models reduced residual between area variation in the physical activity outcome, supporting the notion that environmental quality was contributing to area-level variation in PA. Also, this reduction was greater than that produced for the green space plus domestic garden variable or for the IMD2004 deprivation variable, and addition of all variables simultaneously into the final models resulted in almost exactly the same levels of inter area variation as when the OEQ exposure had been added alone. This indicates that the OEQ exposure derived from MEDIx was capturing more of the elements of the environment that are important for physical activity compared to that of the green space plus domestic garden and of the area deprivation variables that were accountable for some of the inter area variation. Indeed MEDIx included a green space component, and it is possible that some of the dimensions of the IMD2004, such as unemployment, correlate with the multiple environmental deprivation measures captured by MEDIx.

In the multilevel models, the LR test versus logistic regression test statistics were significant only for the empty models and those adjusting for individual level factors for the WHOPA_E outcome in WIIP7 and for the GHQ-12 depression outcome in HSE2008, and for none of the models for the GHQ-12 depression outcome in WIIP7, which suggested that single level modelling was adequate. Nevertheless, the detection and identification of between area variation in health and health behaviour outcomes due to area-level factors has important public health implications: whilst changing an environment may impact an individual's health less than "changing" that individual, environmental modifications may be easier to make and can benefit whole populations.

Depression was hypothesised to be associated with both objective and subjective indicators of environmental quality because environmental quality is supportive of physical activity which is protective of mental health [211]. Although associations were found between the objective indicator of environmental quality in the WIIP7 and GHQ-12 depression, and as expected poorer quality was associated with higher odds of this outcome, at CAS ward level they were lost after adjustment for individual level factors, suggesting it was the individual characteristics of the WIIP7 individuals living in the poorer quality neighbourhoods were associated with a higher likelihood of GHQ-12 depression. At LA level, there was a significant association only between the London-based OEQ exposure and GHQ-12 depression which remained after adjusting for individual level factors. The fact that in London the operationalisation of neighbourhoods as LAs but not CAS wards produced a significant association after adjustment for individual level factors perhaps reflects the fact that there is a smaller difference in size between CAS wards and LAs in this Government Office Region, with LAs in London tending to be smaller than those in other regions. Power to detect significant associations within London LAs, with London-based OEQ exposures would be higher than in CAS wards as there would be more people per LA than per CAS ward. Whilst aggregated environmental measures are expected to be more relevant to health outcomes in smaller administratively defined areas which better represent the area to which an individual is exposed, detection of associations that are significant may prove difficult. Among the nationally representative HSE2008 sample associations between UK and England-based OEQ exposures at LA level remained significant after adjustment for individual level factors. It may be that associations were detectable for the LA level OEQ exposure variable derived from UK and England LAs within this study population but not for WIIP7 because the geographical spread of HSE2008 across England was more even. The results here for the HSE2008 study sample broadly align with those of Mitchell *et al* who found that after adjustment for individual level factors OEQ exposure as indicated by MEDIx, measured at CAS ward level, had a significant association with health [212].

For the WIIP7 sample, LA level residents' satisfaction with parks and open spaces, the indicator of perceived environmental quality (PEQ) was not associated with GHQ-12 depression. This suggests that the views of local neighbourhood residents' satisfaction with greenspace may not reflect those of WIIP7 participants, or simply satisfaction with greenspace does not impact on their mental health. Alternatively, other factors were more important in determining use of such spaces for physical activity or mental health-promoting physical activity was not associated with such spaces in the WIIP7 sample. In contrast to the lack of association in the WIIP7 sample, a significant association between LA level residents' satisfaction with parks and open spaces and GHQ-12 depression was found within the HSE2008 study sample; lower satisfaction was associated with higher odds of GHQ-12 depression after adjustment for individual level factors. It may be that for this population, with a lower mean age, parks and open spaces play a more important role in promoting mental health. Also, LA level residents' satisfaction with parks and open spaces may be more reflective of satisfaction with parks of the HSE2008 study sample as the 2008 Place Survey respondents were drawn from a general population in the same year as HSE2008, rendering 2008 Place Survey participants more similar to those of HSE2008 than to those of WIIP7. The finding that adjustment for general area deprivation, as indicated by IMD score, eliminated the significant association between satisfaction with parks and open spaces and GHQ-12 depression in the HSE2008 study sample was unsurprising as environmental satisfaction is likely to reflect overall deprivation. Also, the LSOA-level IMD measure attached to individuals in HSE2008 contains a living environment domain which contains an air pollution measure, again perhaps correlating with satisfaction with parks and open spaces. For older populations, more represented by the WIIP7, it may be that individual level factors such as level of physical functioning are far more important determinants of depression than parks and open spaces.

Using a cross-sectional, nationally representative survey of 8,979 adults aged 16 to 74 years, Weich *et al* investigated the prevalence of depression, as indicated by a GHQ-12 score of 3 or more, in Britain at individual, household and CAS ward levels [64]. They found that less than 1% of the total variance in depression occurred at CAS ward level. Adjustment for individual level factors rendered variation in their sample statistically non-significant at CAS ward level. However, at household level they found variance was 14.4% after adjusting for individual level factors. In the present study the majority of individuals in the WIIP7 sample of the present study who shared a common OA, the smallest administrative geography, did not cohabit as indicated by their distinct postcodes. Therefore any significant OA level variation in depression in the WIIP7 sample was unlikely to be entirely attributable to variation between households. However, the variation that was detected was specifically between CAS wards; multilevel modelling revealed virtually no variation between LAs and variation between OAs was attributable to variation between CAS wards. In the HSE2008 sample LA level variation in depression was detectable, in contrast to the lack of variation found at this level in the WIIP7 population. The small amount of variation at this level, two per cent, was largely attributable to individual factors. However, the remaining variation between LAs was attributable to the specified environmental quality indicators suggesting that environments measured at this scale are pertinent to depression outcomes. Given Weich *et al*'s findings and the findings in the present study for the WIIP7 sample, it is likely that much of the between area level variation detected is actually due to variation between smaller administrative areas that would be found were it possible to specify lower levels within the multilevel model for the HSE2008 sample. A recent study postcode-level study of over 30,000 residents aged 18 or over living in 777 neighbourhoods in south Wales found that poor self-reported general health was associated with lower residential neighbourhood quality (OR=1.36, $p<0.05$) as indicated by independent observer measurement using the Residential Environment Assessment Tool [209], suggesting smaller areas are more relevant to depression outcomes. Findings here suggest that associations between neighbourhoods and depression are dependent on neighbourhood definitions and population characteristics.

6.5 Summary

The results of this chapter can be summarised as showing that objectively measured environmental quality (OEQ), as indicated by MEDIx score, is negatively and independently associated with meeting WHO recommended physical activity levels (WHOPA_E and WHOPA_I) but not with being in the top tertile for time spent walking per week (TTW), suggestive of differential effects on domains of physical activity. As for the greenspace exposure, it appears that modelling of the OEQ exposure within smaller units is more relevant to individuals' physical activity outcomes. Also, the results show that perceived poorer environmental quality is not associated with any physical activity outcome, indicating that other factors are determinants. Variation in meeting the WHO recommended physical activity level existed between areas but was low relative to that between individuals, supporting the hypothesis that individual level factors were more important determinants. Residual between area variation in the physical activity outcome was reduced by addition of the OEQ exposure to multilevel models, supporting the notion that it was contributing to area-level variation. In contrast to the objective measure, lower perceived environmental quality, as indicated by LA level residents' satisfaction with parks and open spaces, is not associated with any of the physical activity outcomes. Objectively measured poorer environmental quality within neighbourhoods defined as LAs is positively and independently associated with depression in the nationally representative HSE2008 study sample. In the WIIP7 sample independent associations are restricted to certain operationalisations of neighbourhoods, in part a consequence of the uneven geographical spread of this study sample. Lower perceived environmental quality is associated with higher odds of GHQ-12 depression after adjustment for individual level factors in the HSE2008 but not the WIIP7 study sample. However, the association in the HSE2008 study sample is not independent of area-level deprivation, probably due to resident satisfaction being a reflection of this factor.

6.6 Appendices

Appendix 6.1 Associations between perceived environmental quality (PEQ) at LA level and the physical activity outcomes, WHOPA_E, WHOPA_I and TTW, in the WIIP7 study sample.

	WHOPA_E			WHOPA_I			TTW		
PEQ derivation area	OR	CI	p	OR	CI	p	OR	CI	p
England	0.99	0.87-1.11	0.820	0.93	0.83-1.04	0.195	0.96	0.85-1.09	0.546
London	0.99	0.83-1.18	0.890	0.88	0.76-1.02	0.089	0.90	0.76-1.05	0.182

Appendix 6.2 Associations between perceived environmental quality (PEQ) at LA level and GHQ-12 depression in the WIIP7 study sample.

PEQ derivation area	OR	CI	p
England	0.94	0.80-1.10	0.416
London	1.13	0.92-1.39	0.230

Appendix 6.3 ICCs for multilevel models of WHOPA_E in WIIP7.

Model: Factors included	LEVEL 3: LA			LEVEL 2: CAS ward nested in LAs				
	ICC	SE	CI	ICC	SE	CI	LR test versus logistic regression	p
1: None	0.0252	0.0073	0.0143-0.0442	0.0386	0.0149	0.0179-0.0812	32.61	<0.001
2: Individual	0.0087	0.0056	0.0025-0.0303	0.0234	0.0146	0.0068-0.0772	6.38	0.041
3: Individual & GS+DG	0.0061	0.0053	0.0011-0.0333	0.0215	0.0145	0.0056-0.0784	4.17	0.124
4: Individual & UK-based OEQ	0.0037	0.0047	0.0003-0.0439	0.0173	0.0144	0.0033-0.0848	2.38	0.304
5: Individual level factors & area deprivation	0.0065	0.0053	0.0013-0.0315	0.0227	0.0146	0.0064-0.0776	4.93	0.085
6: All factors	0.0039	0.0050	0.0003-0.0456	0.0173	0.0144	0.0034-0.0843	2.36	0.307

Appendix 6.4 ICCs for multilevel models of GHQ-12 depression in HSE2008 at local authority level.

Models: Factors included	ICC	SE	CI	LR test versus logistic regression	p
1: None	0.0198	0.0064	0.0100-0.0370	16.33	<0.001
2: Individual	0.0064	0.0068	0.0008-0.0501	1.01	0.158
3: Individual & GS+DG	0.0034	0.0067	0.0001-0.1363	0.28	0.297
4: Individual & UK-based OEQ	0.0039	0.0067	0.0001-0.1016	0.37	0.271
5: Individual & England-based PEQ	0.0043	0.0067	0.0002-0.0875	0.44	0.253
6: Individual level factors & area deprivation	0.0018	0.0066	0.0000-0.6811	0.08	0.389
7: All factors	<0.0001	<0.0001	n/a	0.00	1.000

Appendix 6.5 ICCs for multilevel models of GHQ-12 depression in WIIP7.

Model: Factors included	LEVEL 4: LA			LEVEL 3: CAS ward nested in LAs			LEVEL 2: OAs nested in CAS ward nested in LAs				
	ICC	SE	CI	ICC	SE	CI	ICC	SE	CI	LR test versus logistic regression	p
1: None	<0.0001	<0.0001	n/a	0.0436	0.0243	0.0144 -0.1249	0.0436	0.0243	0.0144 -0.1249	3.70	0.296
2: Individual	<0.0001	<0.0001	n/a	0.0427	0.0247	0.0134 -0.1273	n/a	n/a	n/a	3.44	0.179
3: Individual & area deprivation	<0.0001	<0.0001	n/a	0.0438	0.0251	0.0140 -0.1289	n/a	n/a	n/a	3.53	0.171

7 Study 4: Modelling the physical activity potential of London neighbourhoods

7.1 Introduction

7.1.1 Chapter overview

This chapter explains the growing interest in the relationship between the physical environment of cities and individuals' physical activity and discusses the rationale for using a walkability index to model the physical activity potential of London neighbourhoods. The particular study which provided the methodological framework for construction of the walkability index models in the present research is outlined with a technical description of model construction. Then a description of the development of the walkability index for the present study is detailed. Finally, descriptive statistics and maps depicting the walkability of London as measured by the walkability index models are presented and discussed with comparisons of the scores attributed to areas by the different models.

7.1.2 Background

Much of the research into urban form in relation to physical activity has examined cities of high income countries but more studies have focused on cities in North America and Australia than on those in the UK [213]. Thus, the urban form factors examined in the present study have been selected on the basis of evidence gathered in cities in predominately non-UK countries for association with physical activity. However, London's urban form, with respect to street connectivity, residential density and land use mix, is likely to differ in many ways from that of the cities in which most of the research has been conducted. This is primarily because it is significantly older and is constrained by a greenbelt, a land use policy to restrict urban growth and an intervention not commonly applied to cities in North America and Australia where there is an abundant supply of land. Over the nineteenth and twentieth centuries, driven partly by public health concerns, London underwent rapid change and expansion involving the dispersal of the population and the reorientation of a transport network from one primarily serving pedestrians to one whose main function was the accommodation of motorised transport. However, the resultant eradication of overcrowding-associated endemic infectious diseases and increased mobility may have come at a price by contributing to some of the public health crises emerging today: in London and other cities of high income countries, high blood pressure, obesity and overweight, and physical inactivity are the top three causes of death [214], the latter two contributing to the former. Since the 1950s when car ownership became widespread in the UK, the average distance travelled per person per day has risen from eight to almost fifty kilometres [215]. However, the law of constant travel time holds that the average number of daily trips per person and the time budget allocated to transport show stability [216]. Therefore, instead of enabling people to travel the same distance faster and thereby save time, greater mobility promotes travel over a greater distance for a longer period. Proximity as a function of time has remained constant but, as a function of distance, it has decreased as a result of greater mobility. Were people willing or able to use greater mobility to travel faster only, rather than further as well, they might be able to "save" time, a resource whose limitation is commonly reported to be a barrier to physical activity for individuals [217]. Greater mobility is advantageous if physical separation of people from other people or things is a desired outcome but is disadvantageous if it deters physical activity, a possibility given the current situation, with obesity-related diseases a major public health concern. Adams argues that whilst greater mobility, enabled by faster motorised transport, is regarded indicative of progress by fostering economic growth and physical freedom for individuals, excessive mobility – which he terms “hypermobility” – is harmful. He posits that regardless of the environmental costs associated with motorised transport, which technological advances will likely reduce, the ability and predisposition of individuals to traverse distances at much higher speeds that they could under their own “steam” – by foot or bicycle – reduces physical activity inherent in an individual's daily routines and thereby increases their likelihood of being obese. His theory is supported by the findings of Bassett *et al* [218] who examined the relationship between active travel - walking and cycling - and obesity rates in Europe, North America, and Australia. Their examination of national surveys conducted between 1994 and 2006 revealed that countries with the highest levels of active transportation, which tended to be in Europe, had the lowest obesity rates.

In addition to the walkability of urban form factors, the present study investigates the spatial scale at which physical environment is measured; in effect, how the neighbourhood is defined. Operational definitions of ‘neighbourhood’ and a discussion of their relative strengths and weaknesses, and the practicalities of using them are given in the literature review chapter.

7.1.3 Walkability Index model basis

Development of the walkability index models used in the present study was guided by a recent study which employed walkability models that differed in their land use mix component [129]. That study sought to investigate the impact of different representations of land use diversity on associations between neighbourhoods and particular types of walking behaviours. The subjects of the study, conducted by Christian *et al*, were the participants of the Residential Environments (RESIDE) project, a longitudinal analysis of the effects of neighbourhood on transport-related physical activity [219], for whom complete environmental data was available (n=1798). Participants of the RESIDE project were recruited between September 2003 and March 2005 inclusive as individuals scheduled to move into new housing developments in the Perth metropolitan area of Western Australia. The mean age of the sample was forty years and 40.5% were male. The outcome used in the study was self-reported walking within the neighbourhood and collected at baseline before participants moved into their new homes, ensuring a wide distribution of neighbourhoods across the study region. Neighbourhood was defined as a ten to fifteen-minute (1600m) walk from a participant's home, with individuals asked whether in a usual week they walked within their neighbourhood for recreation, health or fitness (classified as walking for recreation) or to get to or from places (classified as walking for transport) for >0mins (any time), >=60mins or >=150mins. The questionnaire used to collect the self-reported walking data contained a list of behaviour-specific destinations, with participants asked to tick destinations used for walking [220]. Destinations included for the behaviour specified as transport-related walking included work, public transport and shops, whilst those included for the behaviour specified as recreation-related walking included parks, the beach and walking trails. Responses were dichotomised as yes/no for walking for recreation, walking for transport and total walking.

The exposure variable constructed was a walkability score within each participant's neighbourhood defined as a fifteen minute walk from their home, a "walkable service area", and operationalized as a 1600m street network buffer. This was produced as a line based network buffer, using an extension to ArcGIS software, Network Analyst, with the network distance specified as 1600m and the line buffer distance as 100m. Street connectivity, residential density and land use mix within the neighbourhood were measured to indicate walkability using previously developed methods [113]. They calculated street connectivity as the density of three or more-way junctions in the neighbourhood and residential density as dwelling density of land designated as residential within the neighbourhood. Land use mix was calculated using a formula for entropy (Equation 7.1), a measure of "mixed-up-ness" [221], with greater "mixed-up-ness" indicative of greater walkability. For each of the walkability components, the mean score derived from all participants' neighbourhoods was calculated along with the standard deviation. For each participant's neighbourhood, this mean score was then subtracted from the participant's neighbourhood raw score and divided by the standard deviation to produce a standard z score, varying between -1 and +1. The standard score of the three walkability components were then summed to produce a final walkability score for each participant, indexing neighbourhood walkability for each participant relative to the average neighbourhood walkability in the sample. Participants were grouped into "Low", "Medium/Low", "Medium/High" and "High" walkable environments by quartiling their walkability index scores.

Christian *et al* conducted a literature review to identify land uses associated with specific walking behaviours, enabling them to construct models, varying in the land use mix component, predicted to fit with total walking, recreational walking and transport walking. Land uses were incrementally added to a basic LUM (LUM1) comprising those classified as "Residential", "Retail", "Office" and "Health, welfare and community". LUM2 was constructed to better explain walking for transport, additionally including non-sporting built recreational destinations, classified as "Entertainment, culture and recreation". To better explain both walking for transport and for recreation, LUM3 was specified which added to LUM2 the land uses of public open spaces, recreational sporting facilities and non-farmed natural land. This additional set of land uses was classified as "Public open space, sporting infrastructure, and primary and rural" to LUM3. Two further models were also built but these are not discussed here. A summary of the models constructed, including specific land uses categories, to explain specific walking behaviours in Christian *et al*'s study is given in Table 7.1. Christian *et al*'s land use data was derived from the Spatial Cadastral Database which classifies land uses at the level of cadastral parcels, areas of land defined by their ownership status for taxation purposes [222]. Because a cadastral parcel can contain more than one class of land use, Christian *et al* created a hierarchy of land use classifications to allocate land use to cadastral parcels on a mutually exclusive basis.

The researchers performed logistic regression analysis, adjusting for sex, age, education level, marital status and presence of children at home, to examine associations between each full model of walkability, which included all the three core dimensions of street connectivity, residential density and land use mix, and each specified walking behaviour outcome. They also examined associations of each component of walkability as z-scores with each outcome. They found a statistically significant trend for higher odds of walking in more walkable neighbourhoods, irrespective of the model specified. Walkability was more strongly associated with walking for transport than for recreation. Walking for transport for more than one hour per week was most strongly associated with Model 2, which contained the "Entertainment, culture and recreation" land uses alongside the core land uses of "Residential", "Retail", "Office" and "Health, welfare and community". Walking for recreation for any amount of time was most strongly associated with Model 3, which in addition to the land uses in Model 2 contained those classified as "Public open space, sporting infrastructure, and primary and rural". The researchers concluded that the lack of significant associations found between walkability and spending over 150 minutes on recreational walking per week were due to a lack of statistical power because only a small proportion of the study sample fell into this outcome category. The researchers concluded that land uses that include theoretically relevant transport-related destinations only, but not public open space, were better at capturing associations between walkability and transport walking, consistent with the findings of others [118].

Christian *et al*'s walkability models were constructed to differentially explain physical activity specifically as transport-related walking and as recreation-related walking, defined as walking to recreational destinations rather than recreational physical activity – be it walking or otherwise – per se. Whilst they found that walking for recreation for any amount of time was most strongly associated with Model 3, the contribution of walking to reach a recreational destination relative to that of walking as recreational physical activity per se is unclear. However, the explanatory power of Model 3 with respect to recreational-related transport theoretically extends to recreational physical activities, including those of moderate and vigorous intensities contributing to the WHO physical activity recommendations, because Model 3's land uses include recreational physical activity-supportive ones.

Equation 7.1 Formula used to calculate land use mix by Christian *et al* and also in the present study

H = land use mix score, i = the land use, pi = the proportion of the area covered by the land use against the sum of the area of the land uses of interest, n = the number of land use categories.

$$H = -1 \sum_{i=1}^{n} pi * \ln(pi)/\ln(n)$$

Table 7.1 Models constructed to explain specific walking behaviours in Christian *et al*'s study.

LUM model: Specific walking behaviour modelled:	1: Basic model	2: Transport-related	3: Transport and recreation related
"Residential", "Retail", "Office" and "Health, welfare and community" land use	✓	✓	✓
"Entertainment, culture and recreation" land use		✓	✓
"Public open space, sporting infrastructure, and primary and rural" land use			✓
Residential dwelling density	✓	✓	✓
Street connectivity	✓	✓	✓

7.1.4 Research questions

The research questions this study aimed to address were:

- What is the spatial variation in walkability with Output Areas as spatial units of enumeration across London?
- What is the effect of changing the particular land uses included within the land use mix dimension of the walkability index on this spatial variation?
- How does the operationalisation of a neighbourhood as a spatial unit of enumeration of walkability – an OA, a CAS ward, an LA or a grid cell – influence its walkability score?
- Do walkability scores attributed to OAs match subjectively assigned scores based on visual inspection?
- What are the practical challenges in constructing a walkability index for London?

7.1.5 Hypotheses for Study 4

There will be a gradual radial decay in walkability of London with Output Areas as spatial units of enumeration from the centre to the periphery when a basic land use mix, which excludes recreational facilities and green space, is specified in the model.

The decay will reflect the reduction in residential density and junction density, two of the core walkability dimensions. However, sequential addition of recreational facilities and then green space to the land use mix will reduce the uniformity of the radial decay in walkability because the area of these additional land uses does not reduce uniformly from the centre alongside residential dwelling density: there is much variation in green space area within Inner London.

Thus there is potential for greater variation in land use mix, and thereby walkability, within Inner London when additional land uses are included in the models.

The walkability scores attributed to OAs will match subjectively-assigned scores based on visual field evaluation.

Differential operationalisation of a neighbourhood of a given postcode point as a spatial unit of enumeration of walkability will influence its walkability score due to the effect of the “first law of geography”, a phenomenon summarised by Tobler as “everything is related to everything else, but near things are more related than distant things” [223]. Because proximal places are more similar in terms of environmental attributes, correlation of walkability scores of neighbourhoods that are more similar in scale will be higher. Conversely, the walkability scores of neighbourhoods that are more dissimilar in scale will have lower correlation as the larger unit effectively “averages out” the quantitatively assessed attributes of the physical environments of the smaller units that lie within it. Also correlation between the walkability scores of OAs and those of neighbourhoods delineated as buffers – which are derived as the mean of the scores of constituent OAs – will be low because buffers tend to be substantively larger than OAs.

7.1.6 Objectives

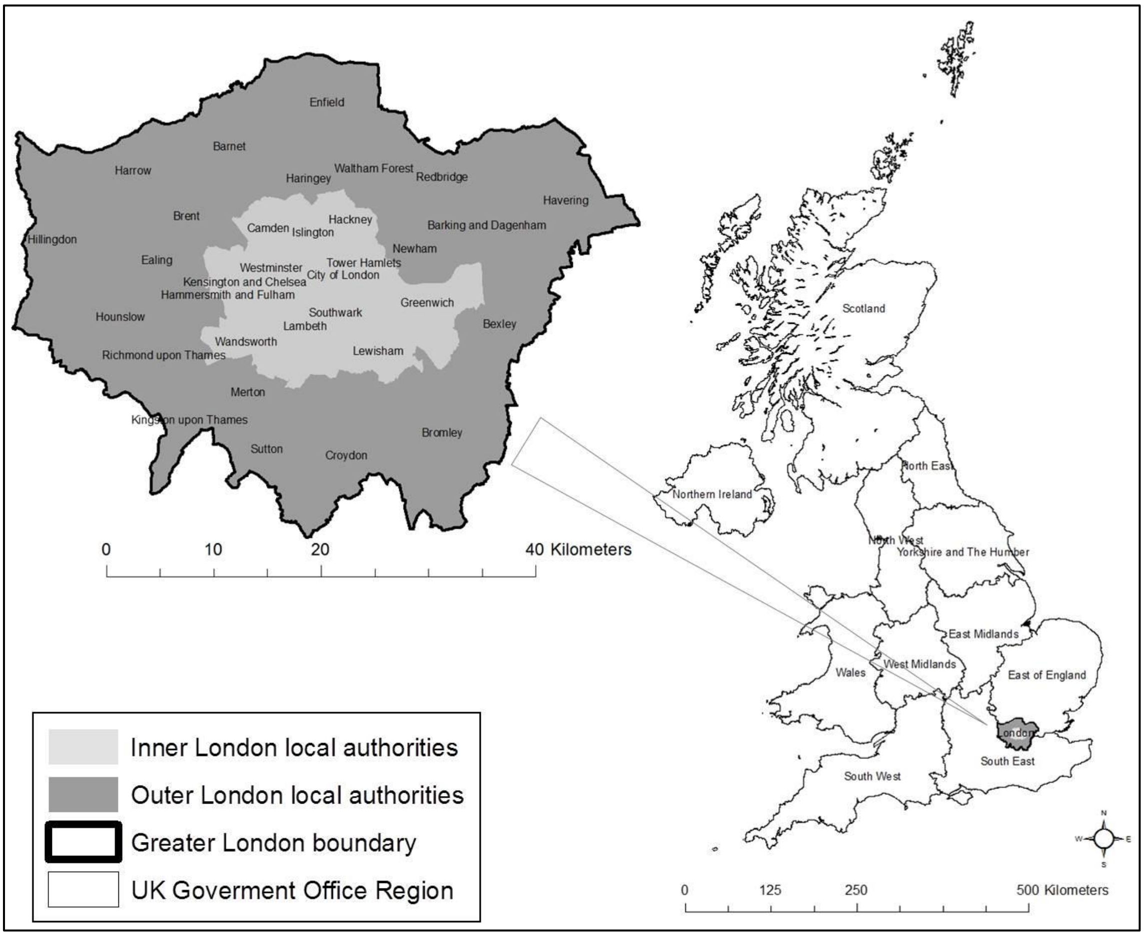
- To construct a walkability index, using Output Areas as spatial units of enumeration, comprising the three core dimensions of residential dwelling density, junction density and land use mix
- To construct two additional models of this index differing in the land uses included in the land use mix component
- To validate the walkability scores at OA level by visual evaluation of a sample of OAs attributed each score
- To compute walkability scores for the higher administrative geographies of CAS wards and LAs as spatial units of enumeration
- To construct gridded cell networks of 2000m by 2000m and of 200m by 200m covering London using GIS software, and to compute walkability scores for them as spatial units of enumeration
- To delineate neighbourhoods as 1600m circular buffers s and as 1600m polygon based network buffers around each LWIIP7 participant’s postcode point using GIS software, and to compute walkability scores from the walkability scores of the spatial units of enumeration they overlap.
- To produce maps enabling the visual comparison of the walkability of London for each different model with a distinct set of land uses included in its land use dimension, and to determine the correlation of scores between models.
- To produce maps enabling the visual comparison of the walkability of London measured within different enumeration units
- To attribute each LWIIP7 participant with a walkability scores of the spatial unit of enumeration in which his or her postcode falls, and to determine the correlation between the walkability scores of WIIP7 participants’ neighbourhoods operationalized as administrative areas, GIS software-computed buffers and gridded cells.

7.2 Methods

7.2.1 Walkability index spatial coverage and units

The purpose of developing a walkability index was to test the tool's ability to predict physical activity behaviours by examining associations between walkability and physical activity outcomes in an occupational-specific sample of older adults drawn from the seventh wave, conducted in 2004/5, of the Whitehall II study, a longitudinal study of civil servants to examine the social determinants of health [182]. The area selected for modelling the physical activity potential of neighbourhoods was Greater London, England because it contained the greatest proportion of the participants of the seventh wave of the Whitehall II study (WIIP7), a subset henceforth referred to as LWIIP7. Greater London is an administrative area of 1,572 km^2 comprising Inner London, namely the City of London and twelve Inner London boroughs covering an area of around 319 km^2, and twenty boroughs constituting Outer London (Map 7.1). At the 2001 Census the population of Greater London was 7,172,036, of which 38.57% was resident in Inner London. The metropolitan Green Belt, a statutory policy and land use designation to limit urban growth, roughly delineates the relatively built-up Greater London from countryside of surrounding counties.

Map 7.1 Greater London was the area in which physical activity potential of neighbourhoods was modelled.



The construction of the walkability index followed the method used in Christian *et al*'s study [129] but a major revision was made with regard to the spatial units of enumeration of the walkability index; rather than develop an index derived solely from the neighbourhoods of the study participants, defined as residential postcode-centred 1600m buffers, the present study produced a walkability index for the whole of Greater London in the spatially contiguous units of administrative areas and grid cells. This approach was taken for several reasons. Firstly, the participant-centred approach would indicate where study members lived, creating disclosure problems. Secondly, it would be difficult to demonstrate that the index is independent of the participants' data because the walkability range would only be derived from the participants' neighbourhoods. As the participants' median wealth was higher than that of Londoners they may have been under-represented in certain walkability types of neighbourhoods. Therefore, the range could be quite narrow and it would be nonsensical to label an individual's neighbourhood highly walkable, and to attribute someone's physical activity to that, when relative to other areas of Greater London it was not highly walkable. A walkability index for the whole of London in spatially contiguous enumeration units would represent all levels of walkability, and encompass greater heterogeneity, enabling a participant's neighbourhood walkabilities to be assessed relative to all other areas of London. Thirdly, a walkability index developed only from LWIIP7 neighbourhoods would make it inapplicable to other study populations. Lastly, although more data processing was involved in the revised method, it was simpler than the participant-centred one. The smallest administrative units, OAs, were used as the spatial units of enumeration to minimise intra-area variation and maximise inter-area variation. In addition, for comparison, larger administrative units and gridded cells were also employed as the enumeration units. A walkability index score was then attributed to each individual based on the spatial unit of enumeration in which their residential postcode fell. Walkability scores were also derived for each participant's postcode-centred buffer neighbourhood as a function of the walkability scores of the intersecting or overlapping enumeration units.

7.2.1.1 Construction and application of walkability index models

7.2.1.1.1 Geographical data management and sources

It was necessary to draw on a range of geographical data sources and to manage data within both statistical and geographical software systems in order to construct the walkability index and to apply it to neighbourhoods for individuals in the study sample. This section outlines the software used, explains the terminology used to describe geographical data storage and provides a rationale for the data storage used. It also details the geographical data sources, applications and preparation.

7.2.1.1.1.1 Managing the geographical data

The Geographical Information System (GIS) software used to store and manage the geographical data for the present study was ArcGIS for Desktop Advanced version 10.1 (ArcGIS10.1), a product licenced by the geodatabase management applications supplier Esri [181]. ArcGIS10.1 enables data building, modelling, analysis and map display, and consists of several integrated applications. ArcCatalog allows browsing of datasets and files and previewing of data on a map, and management of metadata for spatial datasets whilst ArcMap is used for viewing, creating and editing geospatial data. Geoprocessing, data conversion and analysis tools are stored in the ArcToolbox application. The software extensions added to ArcGIS10.1 for use in the present study were Productivity Suite, Spatial Analyst and Network Analyst. In addition, a geoprocessing tool, Line and Junction Connectivity, was downloaded under the Esri Toolshare scheme.

The way the geographical data was stored determined how it could be analysed so it is useful to briefly describe some GIS terminology. A feature is a type of vector data, typically a point, line or polygon, described by its coordinate system: a point is single xy coordinate, a line is a series of xy coordinates joined in a particular order and a polygon is a series of xy coordinates joined in a particular order with the start and end points coincidental such that an enclosed loop is formed. A feature class describes the particular geometry, as points, lines or polygons, which is common to a set of features. Vector data in a single feature class, such as a set of points, may be stored as a shapefile. Unlike storage as a shapefile, storage within a geodatabase accommodates data of different feature classes and allows relationships between them to be defined. This ability is important for various geographical applications such as network analysis to determine the distance along a road (a line) of a residential address (a point) to a park (a polygon). Within a geodatabase, vector data of different feature classes but with the same xy coordinate system can be stored in a feature dataset. Network analysis was necessary in the present study to delineate neighbourhoods as 1600m polygon based network buffers (PBNBs) so data was stored within geodatabases.

Attribute data, the tabular or textual data describing the characteristics of features such as a specific land use area within a neighbourhood, that was produced by geoprocessing in ArcMap such as area of a particular land use falling within an OA, was transferred from ArcMap to a statistical software package because the statistical processing capabilities of ArcMap were limited and inferior to those of software designed specifically for the task. The statistical software package used was StataIC (version 12, Stata Corp, College Station, Texas, USA) [180]. Tables from ArcMap were generally saved and imported to Stata as excel files. However, in some instances it was necessary to save and export the data in the form of text files able to accommodate the high number of rows unsupported by saving in Microsoft Excel format. For example, the table accompanying the topographical map providing land use data, had over eleven million rows, each representing a single land use polygon, all covering the entire London region. Following statistical analysis and computation of new variables, such as walkability scores, it was often useful to save tables and import them into ArcMap. Here geocodes within tables could be joined to geocodes in the tables of map features, and the newly introduced variables visualised on maps.

7.2.1.1.1.2 Sources of geographical data

The geographical data used in the present study was drawn from a wide range of sources. A summary of the requisite geographical data, and their sources and applications is given in Table 7.2 and details of each are given in the following sections.

Table 7.2 Summary of the requisite geographical data – their sources and their applications – in construction of the walkability index, neighbourhood delineation and walkability score attribution.

Geographical data (Source)	Walkability dimension			Neighbourhood delineation & walkability score attribution			
	Land use mix	Residential density	Street connectivity	1600m CB	1600m PBNB	Grid cell	Admin unit (OA, CAS ward & LA)
Digitised boundary data (UK Data Service Census Support)	✓	✓	✓	✓	✓	✓	✓
Topo Baselayer (UKMap)	✓	✓					
OA household counts (Casweb)		✓					
Roads layer (OSMM)			✓		✓		
Urban Paths layer (OSMM)			✓		✓		
R and UP combined network (OSMM)			✓				
Administrative unit codes (Geoconvert)							✓
Geocoded postcodes (Geoconvert)				✓	✓	✓	✓

7.2.1.1.1.2.1 Digitised boundary data

Digitised boundary data was crucial for construction of the walkability index, and for delineating study participants' neighbourhoods and calculating walkability scores within them. It was sourced online from UK Data Service Census Support, formerly UKBORDERS, which provides downloadable data and support for users of the 1971 - 2011 Censuses of Population [224]. A digitised boundary is an aerial representation of a geography of the census, such as an OA, that can be used in GIS and to which data such as counts or scores can be attributed and visualised.

7.2.1.1.1.2.2 UKMap

UKMap was fundamental to the construction of the walkability index as it constituted the spatial data necessary to compute land use mix and residential dwelling densities within OAs. The UKMap collection [225] is a mapping database product of Landmap, a service developed to provide UK academia with spatial data. For the present study, access to the UKMap Collection was kindly arranged by Haklay (*pers comm* [226]). Collection of the data involves taking 12.5cm resolution aerial photos which are then geometrically corrected using global positioning systems (GPS) and a detailed terrain model from the Cities Revealed LiDAR database. The photos are subsequently digitised into map features and tested for absolute positional accuracies compared to independent GPS points of 0.7m random mean squared error accuracy before map production. Using the digitised maps produced as a guide, a field survey team walks all areas where there is public access, collecting further qualitative information, including house numbers, street names and locality information, which is then compiled into the UKMap database. A comprehensive specification listing for all captured features is finally produced, including details of 280 land use codes and 8 different feature types.

UKMap contains integrated data layers, or sets of features, covering London at a topographic scale of 1:1000, including UKMap Topo Base and UKMap Topo Overlay. Features in UKMap Topo Base are depicted as polygons with no gaps or overlaps. Coded information is attributed to every feature in the base and overlay polygon which gives a 30 character Unique Classification Code (UCC) when aggregated. In ESRI format, compatible with ArcGIS, the UCC is supplied as separate constituent fields because the software does not support the use of a 30 character code. The constituent fields are a 1-character Geographic entity type (GET), a 9-character Geographic entity type sequential number (GTN), an 8-character Feature classification code (FCC), a 1-character Feature type code (FTC), a 6-character date of feature edit (DFE) and a 5-character Source reference code (SRC). The FCC is a hierarchical code based on the land component of the National Land Use Database (NLUD) version 4.4 [186]. The United Kingdom's NLUD establishes the national standard classification system for naming and defining land. Despite its name NLUD classifies groups of both land use and land cover, distinct but related terms. Land cover is "the observed (bio)physical cover on the earth's surface" whereas land use refers to "the arrangements, activities and inputs by people to produce, change or maintain a certain land cover type" [187]. Thus, land use essentially describes humans' interaction with land cover. UKMap adopts NLUD4.4 to classify the majority of the land uses it contains, naming their classification source as NLUD_4.4_Land_use. However, certain land "uses" for which there is only a land cover classification within NLUD4.4 are named as UKMap_features which combine the land use and land cover classification fields, giving precedence to land use codes as only a single category is possible. As a single feature can be represented in more than one layer its UCC allows cross-referencing between layers such as those of points, polygons and addresses.

The present study used data from UKMap Topo Base for calculating the land use mix and residential dwelling densities as components of the walkability index, rather than the topography layer of Ordnance Survey MasterMap (OSMM), which is the only comparable data source, because the data coding system was evaluated as more comprehensive, logical – cataloguing in UKMap is based on systematic coding, not free descriptive text as for OSMM – and fit for purpose. As a topographic map, UKMap Topo Base depicted the arrangement of artificial and natural features across areas. It was provided in ArcGIS-compatible ESRI format in ninety one 5km by 5km UKMap tiles covering the 1,583 square kilometre area of Greater London, referenced by a numbering system based on the GB National Grid coordinate system. The outermost tiles actually extended beyond the Greater London boundary, ensuring total coverage of the region. In order to enable land use areas to be calculated within the spatial units of OAs, some of which would span boundaries of UKMap tiles, the shapefiles were merged into a single continuous feature class covering London. Summary statistics for the areas (m^2) of UKMap Topo Base polygons are provided in Table 7.3 according to their Feature Classification Order level (FCC1), derived from the NLUD 4.4 classification. It is evident that of the 11,209,617 polygons of the UKMap topography layer, the highest proportion are residences and that their total area constitutes the largest proportion of the total area covered by the layer, although the median area of such polygons is among the lowest of all categories. It is noteworthy that polygons classified as “Agriculture & fisheries”, “Recreation & leisure” and “Unused land, water and buildings”, and of which subcategories were included in the land use mix component of the walkability index, have among the highest median areas. These findings had important implications for the assignation of land use polygons to OAs: it could be inaccurate to assign the entire area of a polygon from these land use classifications to an OA if it was identified to co-locate with an OA as it could exceed the OA area.

Table 7.3 Summary statistics for areas of all UKMap polygons depicting land uses in London based on their Feature Classification Order level (FCC1), derived from the NLUD 4.4 classification

FCC1	NLUD 4.4 category	N	Mean area (m^2)	SD (m^2)	Median (m^2)	Min (m^2)	Max (m^2)	Total area (km^2)
11	Agriculture & fisheries (AG)	54,979	4060	18,337	159	0.0006	827,178	223.00
12	Community & health services (CM)	127,400	183	550	68	0.0042	57,725	23.30
13	Defence (DF)	1136	512	5375	88	0.0020	169,155	0.58
14	Education places (ED)	149,489	370	3486	81	0.0089	775,152	55.30
15	Recreation & leisure (LE)	147,009	1680	9423	207	0.0021	612,641	247.00
16	Manufacturing (MA)	115,423	392	1439	87	0.0038	125,544	45.20
17	Mineral extraction (MI)	1021	4332	17,242	272	1.4749	271,501	4.42
18	Offices (OF)	84,115	174	457	63	0.0044	50,112	14.70
19	Residences (RS)	8,841,079	66	130	40	0.0008	44,487	582.00
20	Retail distribution & servicing (RT)	504,548	106	481	42	0.0008	209,167	53.60
21	Storage (ST)	69	251	475	48	2.0334	2577	0.02
22	Transport tracks & places (TR)	1,127,435	252	7434	68	0.0008	5,405,707	284.00
23	Unused land, water & buildings (UL)	21,091	1504	9346	133	0.0160	628,043	31.70
24	Utility services (UT)	32,128	904	11,186	98	0.0012	1,335,408	29.00
25	Wholesale distribution (WH)	827	345	1092	48	0.3051	20,561	0.29
99	No classification category	1868	446	2332	75	1.1395	82,191	0.83
All		11,209,617	142	3043	42	0.0006	5,405,707	1600.00

Map 7.2A, derived from UKMap's Level 1 FCC, shows that green space – loosely defined as "Agriculture & fisheries", Recreation and leisure", "Mineral extraction", "Unused land, water and buildings" and "Utility services" – predominates on the periphery of London, the physical manifestation of the green belt policy. However, sizeable areas of green space, particularly "Recreation and leisure" as parks, are distributed across both Inner and Outer London. Residences predominate across London, as indicated by dark blue areas, although it must be noted that this land use includes private gardens which would provide a greener bird's eye view. "Community and health services" and "Education" appear evenly dispersed across London but again it must be noted that these land classifications include areas of green space such as playing fields. Map 7.2B zooms in on Inner London to illustrate the higher concentration of shops and offices that exist in this region, with offices in particular heavily concentrated in very central London. The greater detail visible in Map 7.2B also illustrates how shops tend to track main roads, as opposed to minor roads which mainly infiltrate only residential land. Given the distribution of land uses illustrated in Map 7.2A and Map 7.2B, with greater diversity found within central London, land use mix within OAs was expected to be higher in the centre than peripheral areas.

7.2.1.1.1.2.3 Ordnance Survey MasterMap

Ordnance Survey MasterMap (OSMM) is a geographical database which digitally represents physical entities such as buildings and roads as topographic features. OSMM's layers include the Integrated Transport Network (ITN) layer [227], which has an additional Urban Paths (UP) Theme layer [228] and both were crucial in the calculation of the street connectivity for the walkability index, and also for network delineation of neighbourhoods as buffers for which the walkability scores were calculated. The OSMM ITN Layer was provided as Ordnance Survey Data Supply Digital Data under licence from Digimap Ordnance Survey Service at EDINA, the national academic data centre based at the University of Edinburgh [229] and the OSMM UP layer was provided directly by Ordnance Survey and received on disk.

The ITN layer represents the driveable roads – defined as ways to which systematic application of material has been applied to assist the passage of an ordinary motor vehicle such as a family car – of Great Britain. Whilst all public roads are included in the ITN layer, coverage of private roads is not complete. The UP layer represents the network of transport ways, urban paths, accessible to non-motor vehicle users such as pedestrians and cyclists including all man-made footpaths, subways, steps, footbridges and cycle paths. In the ITN and UP layers, roads and urban paths are topographically represented as links and their junctions as nodes. Whilst ITN and UP are independent networks, the OSMM developers have connected them through the creation of connecting links (CL), features that represent the logical connection between a path and a road. Each CL, which has no real world geometry and constitutes a zero length line, has two connecting nodes (CN) at each "end". One CN represents the position where the path meets the road whilst the other represents the place where the road meets the path. Thus, although at different ends of a CL these nodes are spatially coincidental: arguably, the link would be better represented as a single point. However, spatial analysis necessitates the ability to relate a particular road to a particular path so it is necessary to represent connections as distinct lines with distinct nodes; essential where a path joins to multiple roads, or a road to multiple paths.

It was necessary to build a combined ITN and UP layer network dataset in order to construct network buffers. Individual and combined ITN and UP layer network datasets were also required for the calculation of street connectivity as a component of the walkability. A transport network dataset enables the modelling of the use of roads and paths, such as the distance from A to B on roads accessible to pedestrians. It is built from the three core elements of the transport layer, namely edges (or links), junctions and turns, by specifying how they relate to one another and recording this attribute information against each feature. The OSMM data was converted to ESRI Personal Geodatabase format, compatible with ArcGIS10.1 using the ESRI ProductivitySuite3 OSMM Data Converter Tool, an extension to ArcGIS10.1. Subsequently the OSMM Data Preparation Tool was used to create a network dataset for the ITN layer and, separately, a network dataset for the UP layer as described elsewhere [230]. The Network Analyst extension in ArcGIS10.1 was then used to create a network dataset combining the individual ITN and the UP network datasets. Henceforth this combined network dataset is referred to as the integrated road and path (IRP) network dataset.

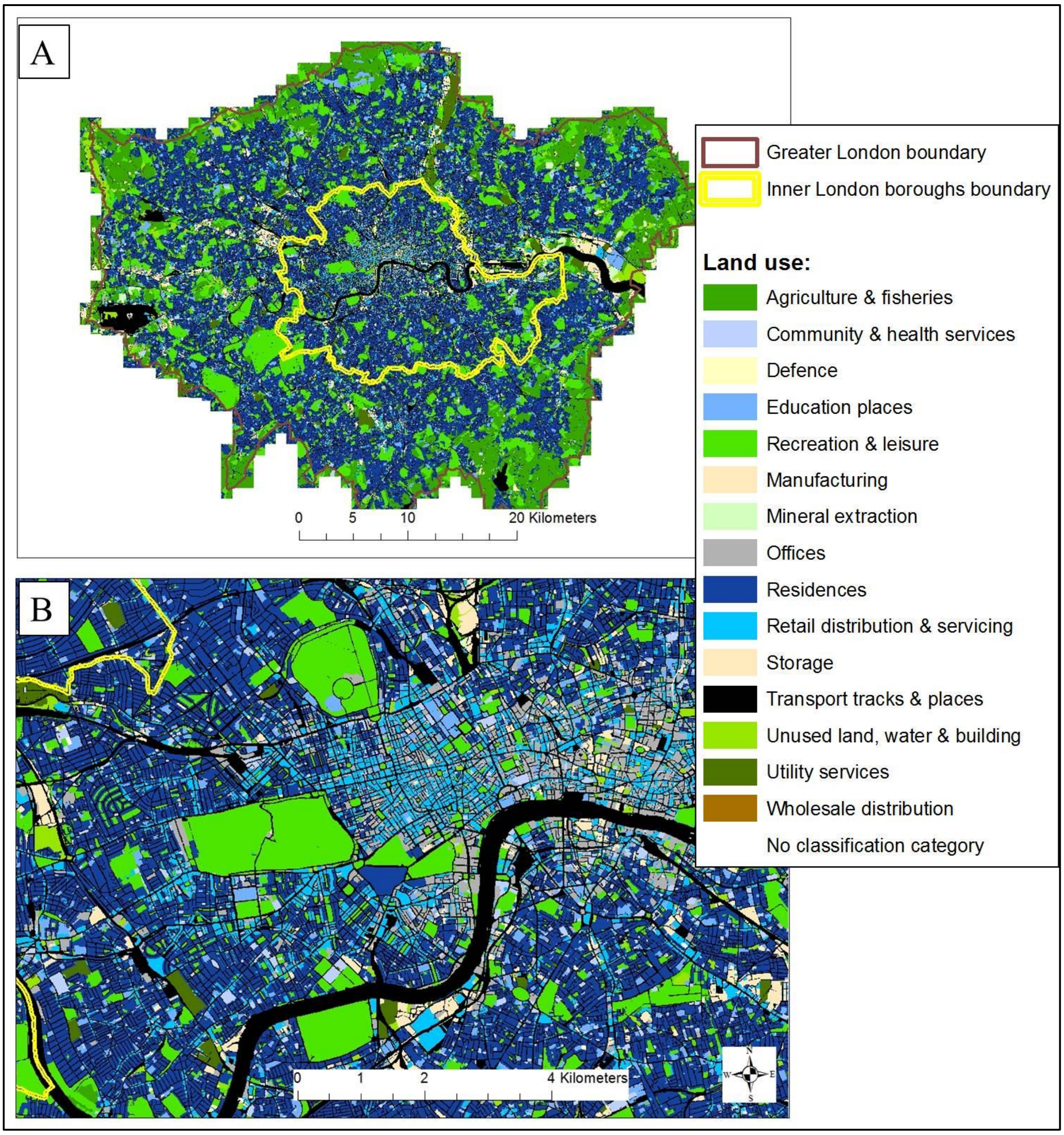
7.2.1.1.1.2.4 Casweb

Residential dwelling counts within each administrative area were required to derive the residential dwelling density component of the walkability index. These were obtained from Casweb, a web interface developed and supported by the Census Dissemination Unit from which information from the UK Census of Population can be extracted [231].

7.2.1.1.1.2.5 Geoconvert

Geoconvert, developed by the Census Dissemination Unit, is an online geography matching and conversion tool that converts data from one geography to another and provides metadata, such as deprivation scores, grid-reference coordinates and area classifications, for postcode and higher geographies [232]. Geoconvert was used to convert WIIP7 postcodes to easting and northing geocoordinates that could be geolocated and mapped in ArcMap, to construct buffers in which walkability scores could be computed, and also to match these geocoordinates to higher geographies to which walkability scores were attached.

Map 7.2 Land uses across London as indicated by UKMap's Level 1 Feature Classification Code.



7.2.1.1.2 Model specifications

Three walkability models were constructed to explain physical activity behaviours, broadly on the basis of Christian *et al*'s findings [129]. As per their method, all contained the fixed dimensions of residential dwelling density and intersection density alongside a LUM dimension which included a specific set of land uses unique to each model. Brief definitions and hypothesised relationships with walkability for each of these core dimensions are given in Table 7.4.

The model specifications differed slightly, however, because in the present study it was not possible to differentiate walking for the purpose of transport from that for recreation. Higher walkability as scored in all three models was expected to be associated with higher odds of both being in the top tertile for time spent walking per week, and for meeting the WHO recommended physical activity level in the WIIP7 sample.

The land use mix included in each model and the physical activity behaviours modelled are summarised in Table 7.5. Model 1, the basic model upon which subsequent models could be built to better explain particular physical activity behaviours, was identical to that of Christian *et al* [129] with respect to the broad land use categories included, namely "Residential", "Retail", "Office" and "Health, welfare and community". The precise land uses included are detailed in Table 7.6. Group residences were excluded from the "Residential" category in the land use mix dimension of walkability because the calculation of residential dwelling density dimension of walkability, which was distinct from the land use mix dimension, specifically required residential dwelling land use but not that of group residences. It was more convenient to create one residential land use dataset that could be used in the construction of both dimensions and, given that only a very small proportion of residential land uses were classified as group residences in terms of number and area, this was considered acceptable. Model 1 land uses were considered supportive of physical activity as personal business destinations potentially reached by foot or bike. Similar to Christian *et al*'s second model, Model 2 in the present study built on Model 1 to include "Entertainment, culture and recreation" land use categories. However, unlike their second model, Model 2 included sporting infrastructure. Model 2 was expected to fit better than Model 1 with both the walking and the WHO recommended physical activity level outcomes because it contained more destinations to which to walk, and wherein to be physically active, be it in the form of walking or another. Model 3 extended Model 2 to include a category named "Free recreational land". This was defined as predominantly natural land accessible to all, at no financial cost and potentially suitable for walking and cycling for transport, or for recreation, and for other informal recreational physical activities. It must be noted that whilst zoological and botanic gardens may constitute outdoor amenities accessible to all, they were not exclusive to Model 3; they were included in Model 2 because a fee to such places is usually charged. Inspection of the land use showed that the majority of land uses classified as "Unused land in natural or semi natural state" were "Woodland and scrub". Other specific types of land uses in this category were unlikely to afford recreational activity, such as "Peat, bog, freshwater marsh and swamp" but they were included as their number was small and programming their exclusion was deemed unnecessary. Including the greatest number of land use categories, as potential supports for physical activity Model 3 was expected to provide the best fit for all physical activity outcomes.

Table 7.4 Definitions of dimensions of neighbourhood hypothesised to confer walkability, adapted from Frank *et al* [113].

Neighbourhood dimensions	Definition	Hypothesised relationship with walkability
Net residential density	Number of residential units per unit area of residential land	Higher level creates critical mass of walkers, encouraging others to walk, and increases traffic congestion, deterring driving
Street connectivity	Number of junctions per unit area	Greater street connectivity increases directness of routes and provides alternative ones that encourage walking and cycling
Land use mix	Evenness of distribution of land area specific to a range of uses such as residential, retail, office and recreation.	Greater land uses mix confers greater ease of access to destinations by foot or bicycle

Table 7.5 Summary of land use mix included in each model and the physical activity behaviours modelled.

Model	Broad land use categories	Physical activity behaviour modelled
1	• Residential • Retail • Office • Health, welfare & community	Walking and all physical activity contributing to WHO recommended level
2	• Residential • Retail • Office • Health, welfare & community • Entertainment, culture & recreation	Walking and all physical activity contributing to WHO recommended level, with stronger associations than with Model 1
3	• Residential • Retail • Office • Health, welfare & community • Entertainment, culture & recreation • Free recreational land	Walking and all physical activity contributing to WHO recommended level with stronger associations than with Model 1 and 2

Table 7.6 Land uses included in land use mix models by broad category. (Data source: UKMap User Guide Version 4.3 [233])

Model	1	2	3
Residential • Dwellings; Non-dwelling structures - garages, sheds; Non-dwelling structures - swimming pools	✓	✓	✓
Retail • Retail distribution places; Retail Centre[a]	✓	✓	✓
Office • General offices; Central government administration office; Local government administration office	✓	✓	✓
Health; welfare & community • Health care places: Medical diagnosis and treatment centres; Dentist's surgery and consulting room; Doctor's surgery and consulting room • Medical auxiliary service centres: Ambulance stations; Blood transfusion centre; Family planning clinic; Forensic medicine centre; Hospitals; Ear, nose and throat hospital; Eye hospital; General hospital; Geriatric hospital; Isolation hospital; Maternity hospital; Mental hospital; Convalescent home • Medical research laboratory • Public convenience • Non-medical homes: Children's home; Handicapped and disabled people's home; Old people's home • Community protection services: Borstal institution; Prison; Civil Defence centre; Coastguard station; Fire station; Police station; Life boat station • Advertising places • Social meeting places: Church hall; Club meeting place; Community centre; Places of worship • Justice administration places; Law courts • Education places: Pre-primary schools; Primary schools; Middle school; Secondary schools; Sixth form college; Special school • Specialised, higher and further education centres: College of further education; Teacher training college; University teaching	✓	✓	✓

Model	1	2	3
establishment			
Entertainment, culture & recreation • Amusement places: Aquarium; Children’s playground ; Fun fair; Night club • Show places: Broadcasting, filming and sound recording studio; Cinema; Concert arena; Theatre; Leisure Centre • Libraries, museums and galleries: Lending library; Museum; Galleries; Art gallery • Land sport places: Association football ground; Cricket ground; Rugby football ground; Bowling green; Miniature golf course; Squash court; Tennis court; Ten pin bowling alley; Golf driving range; Golf course; Athletic ground; Gymnasium; Ice rink; Roller skating rink; Skiing and tobogganing run; Rock climbing; Target shooting places; Cycling circuit; Motor vehicle racing track; Dog racing track; Horse racing course; Horse training area; Hunting and shooting places • Water sport places: Swimming baths; Watercraft places; Boating facilities; Rod/recreational fishing place • Holiday camps: Camping site; Holiday camp site; Holiday caravan site; Youth hostel • Outdoor amenity places: Zoological garden3; Botanical garden4		✓	✓
Free recreational land • Grazing places: Permanent pasture; Rough grazing • Forestry places: Coniferous forest ; Coppice; Deciduous forest; Mixed forest; Tree nursery • Unused land in natural or semi natural state: Beach or sand dune; Cliff or natural outcrop ;Grass land; Heath and moorland; Peat, bog, freshwater marsh and swamp; Salt marsh (unused); Woodland and scrub • Nature reserve; Site of special scientific interest • Outdoor amenity places: Country park; Gardens (not private); Park ;Picnic site; Recreational open space ;Ancient monument ;Monument; Ruins			✓

[a]Each retail centre had a unique code and classification dependent on which floor or floors it was on, such as “Retail Centre - 2 floors - floor level -3”, but the specific land use names are not given here as they are too numerous.

7.2.1.1.3 Computation of core walkability component and final scores

Computation of core walkability component and final scores were based on the methods used by Leslie *et al* [234]. The main difference in the method employed by Leslie *et al* from that of Christian *et al*'s study, on which the present one is based [129], was in the use of measures of relative standing of walkability scores, the quantitative relationship a score for a particular spatial unit has with the walkability scores for all spatial units. Leslie *et al* took quantile measures, which related the scores to the median score of all spatial units, whereas Christian *et al* took z-scores, relating each score to the mean score. Core walkability component scores were computed for land use mix, street connectivity and residential dwelling density. Computer programmes were written to perform these procedures, which involved processing data using both ArcGIS and Stata software, and were recorded in Stata do files for reference.

7.2.1.1.3.1 Land use mix in OAs

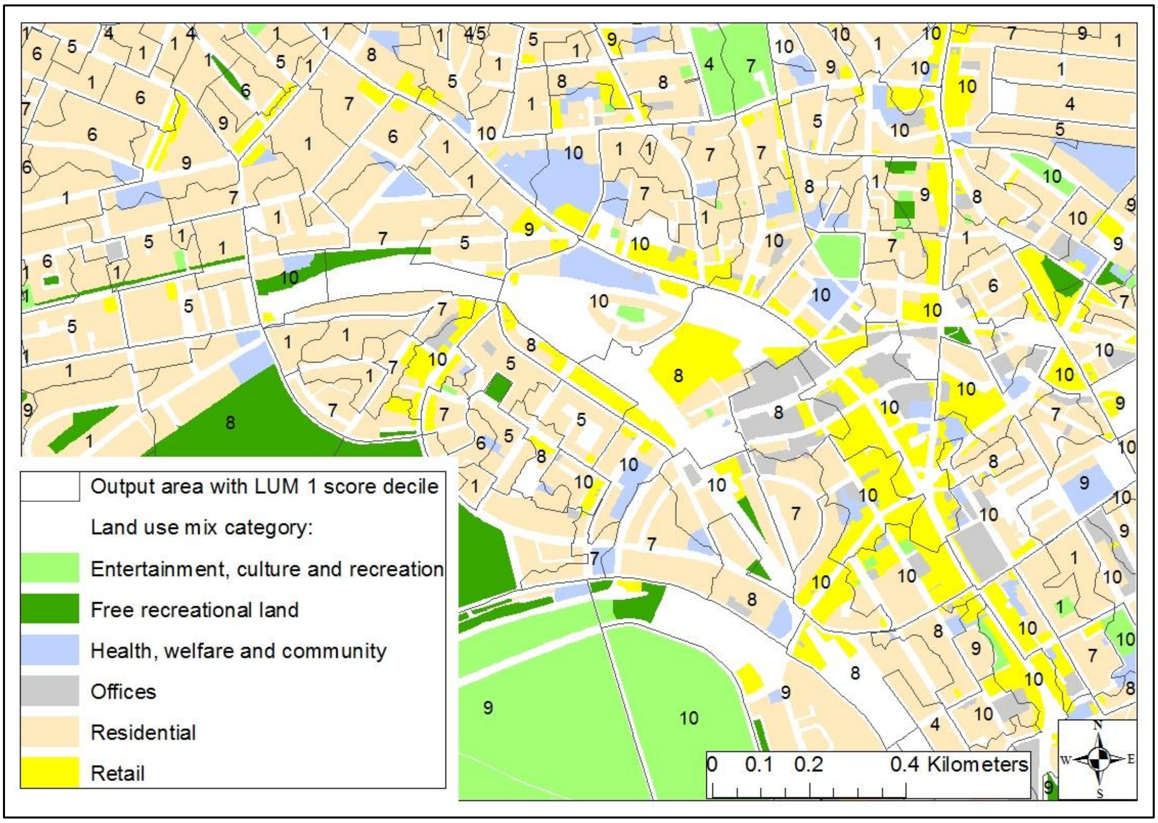
The first step in the derivation of land use mix scores for OAs was associating the OAs with co-located polygons depicting land use in the UKMap Topo baselayer. Rows were matched in the attribute table of the OA boundary layer to those in the UKMap topography layer, based on their relative spatial locations (where a UKMap polygon centroid fell within a OA polygon), and an additional feature layer was produced in which every UKMap polygon was assigned the attributes of the OA to which it was spatially matched.

Two methods of determining areas of land use within OAs were used, the method selected influenced by two factors, namely computational practicality and the area of the UKMap polygon relative to the area of the OA in which its centroid fell. Computationally, assigning an entire area of an UKMap polygon to an OA in which its centroid fell was simpler than determining the area of overlap between the UKMap polygon and one or more OAs, and assigning only the overlapping portions of the UKMap polygon to the OAs. In order to determine the most suitable method for attribution of areas of land use within OAs it was necessary to gauge the likelihood that a UKMap topography layer polygon in a given land use category would fall entirely within an OA. A UKMap polygon whose area was less than or equal to five per cent of that of the OA in which its centroid fell was considered very unlikely to cover more than one OA. If the likelihood was high, the entire area of a UKMap polygon was attributed to the OA in which its centroid fell – its "host" OA – but, if not, only the intersecting area was attributed to the OA. Statistical analysis revealed virtually all of the UKMap polygons (99.70%) had areas that were five per cent or less than the OAs in which their centroids fell.

The dataset produced by spatial association of OAs with land use polygons and attribution of land use areas to OAs was used to categorise land use mix and calculate its entropy, or "mixed-up-ness". To describe the land uses for inclusion in each land use mix models a new variable "category" was created and each polygon, or observation, assigned a label based on its specific eight digit feature classification code (FCC). In the UKMap database there were 856 eight digit FCCs, each divided into four pairs representing the order (Level 1), the group (Level 2), the land use class (Level 3), the land use subclass

(Level 4), and in Stata each level represented a single variable. Therefore, attribution of given land uses to a particular category required careful consideration of the combination of attributable codes. The label assigned indicated whether land use represented “Residential”, “Retail”, “Office”, “Health, welfare & community”, “Entertainment, culture & recreation” or “Free recreational land”. Additional variables were computed which gave the sum of a particular land use category for each OA. Three new variables were then generated, one for each land use mix (LUM) model, each containing the sum of land use areas of their constituent categories. LUM model 1 (LUM1) contained “Residential”, “Health, welfare and community”, “Retail” and “Offices”, LUM2 contained “Entertainment, culture and recreation” in addition to all those of LUM1, and LUM3 contained “Free recreational land” in addition to all those in LUM2. Additional variables were derived from this dataset and new datasets generated, enabling calculation of an entropy score variable for each LUM model according to Equation 7.1 as specified by Christian [129]. The LUM entropy scores were recoded into deciles with a score of 1 indicative of the lowest entropy and a score of 10 indicative of the highest. Map 7.3 shows various land use categories, represented as different colours, and the decile scores within OAs. These scores indicate the extent of the land use mix for LUM1 within each OA, higher scores indicative of greater land use “mixed-up-ness”. It must be noted that the LUM1 model excluded the categories “Entertainment, culture and recreation” and “Free recreational land” so their relative areas were not used in the calculation of the decile values shown.

Map 7.3 An illustration of entropy scores within OAs showing their spatial relationship with various land uses.



7.2.1.1.3.2 Street Connectivity in OAs

Street connectivity relates to the feasibility of walking from one point to another: the more connected the streets, the more direct the route [235]. In the present study the indicator of street connectivity was junction density within OAs. The requisite geographical data included the digitised boundaries for OAs, Ordnance Survey MasterMap (OSMM) Integrated Transport Network (ITN) layer (comprising only roads), the Urban Paths (UP) layer (comprising only urban paths) and the integrated road and path (IRP) network dataset. Points identified from road and urban paths network as representative of street connectivity were counted within each OA and junction density was calculated as the number of junctions in an OA divided by its area. These densities were then recoded into deciles with OAs scoring 1 having the lowest junction density and those scoring 10 the highest.

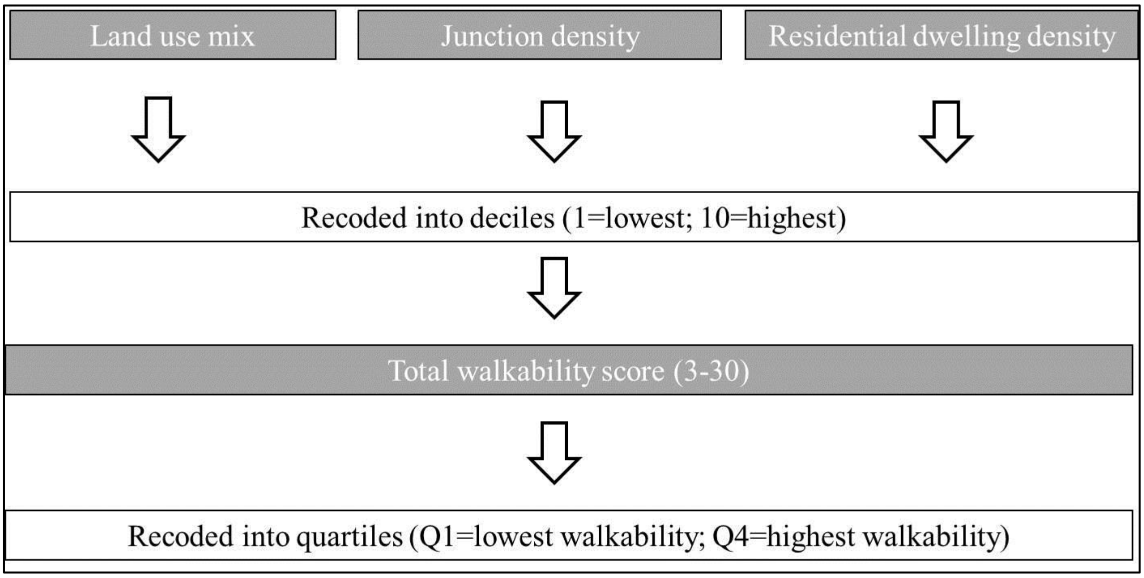
7.2.1.1.3.3 Residential dwelling density in OAs

Residential dwelling density was the simplest component of the walkability index in terms of its calculation. It has been proposed that high residential densities create more walkable environments by providing a critical mass of people who walk and are seen to be walking by others who are, in turn, encouraged by safety in numbers to walk as well [113]. Also, traffic congestion associated with higher residential densities may promote active travel above non-active travel [107]. The 2001 census definition of dwellings is accommodation comprising either a single household space or several household spaces sharing some facilities, whilst the definition of household is the person or people occupying the dwelling. The purpose of the residential density component of the walkability index was to capture the extent to which an area is populated by residents so the number of occupied household spaces was deemed a more accurate reflection of this quality than a count of the dwellings, the physical buildings containing households, which may include unoccupied ones. OA household counts, classified as all resident occupied household spaces were downloaded from Casweb [236]. Residential dwelling density for each OA was calculated as the number of occupied households divided by the area of land classified as residential in UKMap within the OA which had been calculated for the land use mix component of the walkability index. Residential dwelling densities were then recoded into deciles with OAs scoring 1 having the lowest residential dwelling density and those scoring 10 the highest.

7.2.1.1.4 Final computation of walkability scores

The overall walkability score for each model was calculated as the sum of the three core walkability component decile scores and the final walkability scores were then recoded into quartiles as illustrated in Figure 7.1. In addition to computation and analysis as full models, walkability was also calculated and analysed for each of the three core components as deciles and quartiles of the raw scores. Descriptive statistical analysis was performed in Stata to examine correlations between scores. The Stata dataset was then exported as an excel spreadsheet to ArcGIS and joined by the key variable OA code to the attribute table of an OA boundary layer in an ArcMap document. This enabled visualisation of the scores across London OAs for the full models and for each component.

Figure 7.1 Final computation of walkability scores.



7.2.1.2 Measurement of walkability using other spatial units of enumeration

7.2.1.2.1 Construction of grid covering London and measurement of walkability within its cells

Although intra-administrative area size variation of each kind of unit in which walkability was enumerated in the present study across London – OAs, CAS wards and LAs – was less than across the whole of the UK as discussed in Chapter 3 in Section 3.3.1.4, its existence meant that walkability scores assigned to each administrative unit would be subject to the scaling effect of the modifiable areal unit problem (MAUP). Considerable variation in administrative area shape, giving rise to the zoning effect of MAUP, was also evident on inspection of a map of the units across London. Effort was therefore made to mitigate the MAUP using ArcGIS software with the construction of a grid cell network superimposed over London, standardising the walkability enumeration units in which walkability was measured in terms of size and shape, and reducing the subjectivity of boundary selection. Two grid cell sizes were specified; 2000m by 2000m, giving an area of 4,000,000m^2 that approximated the median area of the “walkable” 1600m polygon base network buffer and twice the median CAS ward area, and 200m by 200m, giving an area of 40,000m^2 that approximated the median London OA size. The method outlined for OAs was used to compute grid cell walkability scores. However, in computation of the land use mix parameter, the entire area of a particular land use was assigned to the grid cell in which its centroid fell rather, regardless of its size relative to the grid cell. This was because the larger grid cells 2000m by 2000m grid cells were virtually all substantially larger than the UKMap polygons, and for the smaller 200m by 200m grid cells it was deemed too computationally demanding to calculate only the intersecting areas. Also, whereas the calculation of residential density within the administrative units of OAs was straightforward – residential dwelling counts were specifically collected and readily available for these units from the Census Dissemination Unit [231] – counts were not available for the “artificial” grid cells. Therefore, it was necessary to "count" residential dwellings within each cell by attributing the number of residential dwellings in an OA to the grid cell in which its centroid fell.

7.2.1.2.2 Measurement of walkability within CAS wards and LAs

Walkability scores were computed for CAS wards and LAs following the same method outlined for OAs. However, given that OAs were aggregated to, and nested entirely within these higher geographies, it was possible to attribute OA residential dwelling counts to them more accurately than to grid cells, for which residential dwelling counts were calculated as the sum of those of the hosted OAs whose centroids fell inside them. In addition to showing spatial variation in CAS ward walkability model scores in maps, the extent to which the models correlated was also shown. This was achieved by calculating the net difference in between the walkability quartile score of land use mix 1 (L1Q) and that of land use mix 2 (L2Q), and between L1Q and the walkability quartile score of land use mix 3 (L3Q) for each CAS ward. Net differences between the walkability quartile score of the full model 1 (M1Q) and that of the full model 2 (M2Q), and between M1Q and the walkability quartile score of the full model 3 (M3Q), were also calculated. The model difference scores were exported from STATA as an excel spreadsheet along with their CAS ward codes and the table added to an ArcMap document. In ArcMap the table was joined to a CAS ward shapefile via the administrative unit code and the values symbolised on a map.

7.2.1.3 Attribution of walkability scores to unique WIIP7 postcodes

Walkability scores were attributed to each unique LWIIP7 postcode simply based on the walkability index spatial unit of enumeration in which it fell. The postcode-attached walkability scores were then matched to LWIIP7 participants based on postcode of residence.

7.2.1.4 Construction of buffers around WIIP7 postcodes and measurement of walkability therein

7.2.1.4.1 Conversion of postcodes to geocoordinates

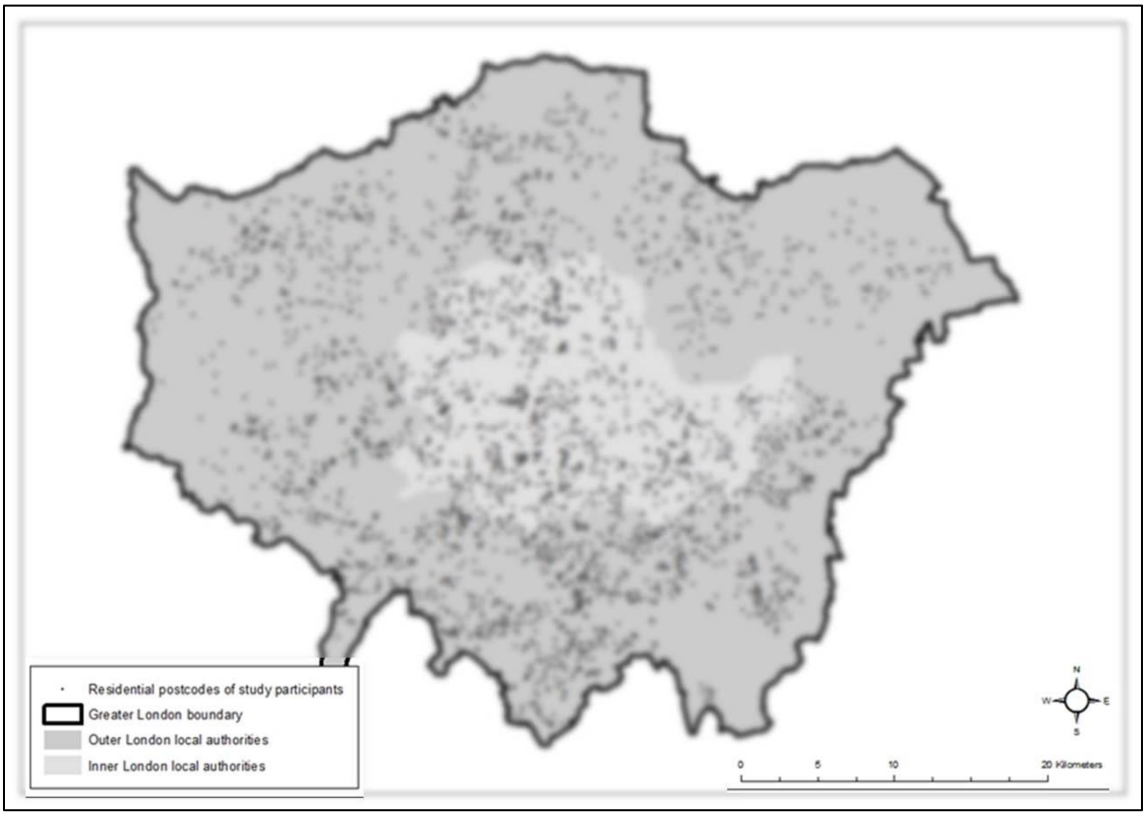
To construct buffers around LWIIP7 postcodes in which to assign walkability scores, it was first necessary to convert the postcodes to eastings and northing geocoordinates that could be geolocated and mapped in ArcMap. The tool also provided information about the positional quality – the accuracy of the assigned grid reference pair – of each postcode. Descriptive analysis and reference to the National Statistics Postcode Directory 2006 User Guide [237] revealed that 2889 of the 2995 unique input London postcodes had a Positional Quality Indicator (PQI) code of 1, indicating that they occurred within the building of the matched address closest to the postcode mean – the mean position of the delivery points at that postcode – and as such had the highest positional quality possible of one metre resolution. Only one postcode had a PQI code of 5, indicating that the coordinates were calculated by reference to surrounding postcode grid references, whilst five had a PQI code of 8, indicating that use of the postcode had terminated and that the last known ONS grid reference was used. Thus, the vast majority of postcodes were mapped to a very high degree of accuracy and only six were mapped with a slightly lower level of confidence, suggesting the postcodes were an excellent proxy for residential location. The postcode data, including the X and Y coordinate data for each postcode, was added as a table in a map document in ArcMap and used to create a layer of postcodes as points that could be displayed as a map (Map 7.4). Inspection of the postcodes and their geocoordinates in Stata revealed that only 2888 of the 2894 sets of geocoordinates associated with each postcodes were unique, indicating that some postcodes shared the same geocoordinate pair. Therefore, duplicate geocoordinates were removed from the points layer prior to construction of buffers.

7.2.1.4.2 Construction of circular and polygon-based network buffers

In ArcGIS the Buffer tool was used to create circular buffers (CBs) centred on each of the unique sets of geocoordinates points for LWIIP7 participants' postcodes as a proxy residential location. A linear unit of 1600m specified so that circular areas with a radius of 1600m were produced. The Network Analyst extension in ArcGIS was used to construct polygon based network buffers (PBNBs) centred on each of the unique sets of geocoordinates points using the IRP network dataset layer, comprising roads and paths in a combined network. It was noted that the ArcGIS software failed to construct PBNBs for geocoordinate pairs that were closer to a path than a road. Therefore, it was necessary to "snap" all points, representing geocoordinate pairs, to the nearest road prior to performing this operation. The impedance specified was 1600m, producing a buffer that encompassed all accessible roads and paths up to and including 1600 metres along the network from the point, representing a LWIIP7 participant's residence.

Given that a walkability score for each buffer was calculated as the mean of the walkability scores of the spatial enumeration units – LAs, CAS wards, OAs or grid cells – that it intersected or fell within, it was imperative that the buffers were entirely within the Greater London boundary: a walkability score for a buffer that encompassed an area outside London would not be accurately calculated as the mean would be derived from scores covering only a portion of the buffer. Therefore, such buffers were excluded based on their locations relative to the Greater London boundary. Analysis in ArcGIS revealed that 2633 of the 2888 unique CB buffers lay completely within the Greater London boundary (Map 7.5). A larger number of PBNB buffers, 2737, lay completely within the Greater London boundary, reflecting the fact that in all instances PBNBs had smaller areas, rendering them more likely to do so. Figure 7.2 depicts the process by which postcodes were selected for buffer construction.

Map 7.4 The 2888 unique geocoordinate pairs, representing residential locations, mapped for the 3020 LWIIP7 participants (image blurred to protect anonymity).



Map 7.5 2633 unique CBs fell completely within the Greater London boundary (A) whilst 255 CBs did not (B).

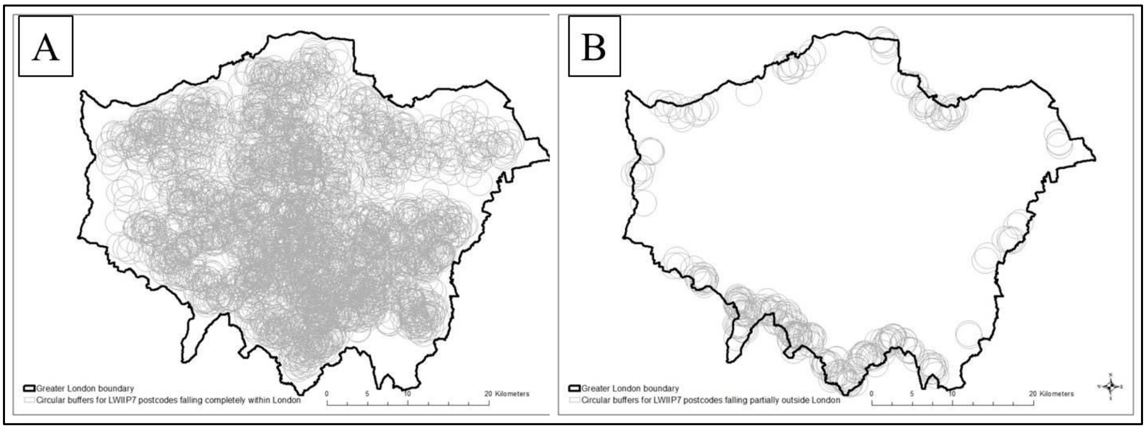
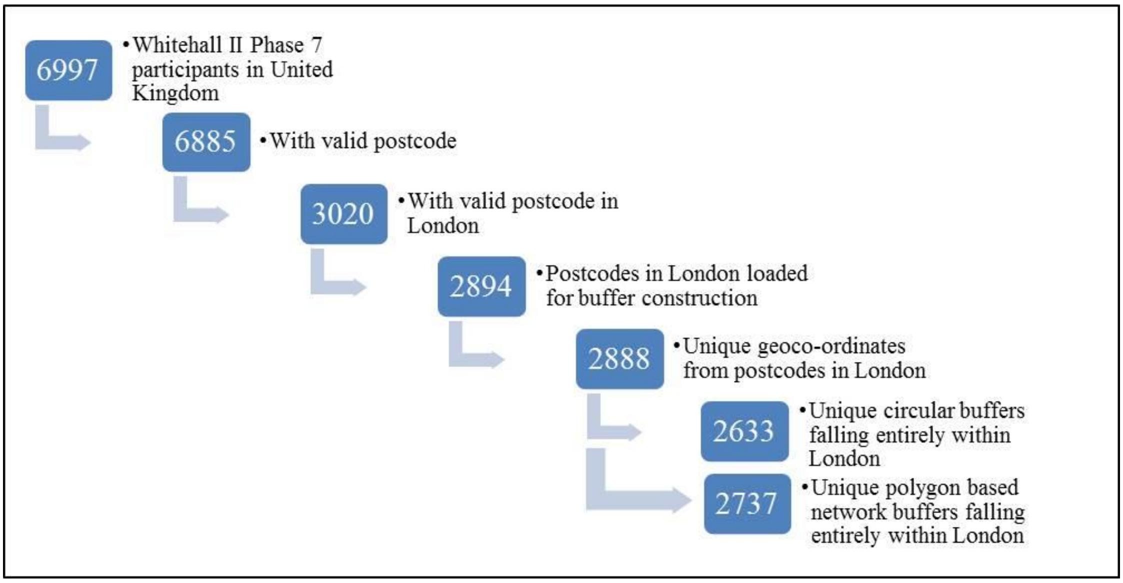


Figure 7.2 Process of selection of residential locations for buffer construction.



7.2.1.4.3 Attribution of walkability scores to buffers

Buffer areas tended to be larger than OAs, the smallest units used as spatial units of walkability enumeration. This meant that multiple OAs tended to fall within each buffer and it was possible to derive a walkability score for a given buffer from those of the OAs it "hosted". Two methods were considered for the derivation of walkability scores for buffers. The first method considered was to spatially join each buffer, using the Spatial Join tool in ArcGIS, to the OA polygons whose centroids fell within it. For each buffer, the tool would compute the number of OA polygons falling within it together with the mean of any numerical attributes including, critically, walkability scores. The disadvantage of this method was that it was not possible to weight the walkability score of each OA based on the proportion of the buffer area it covered. Thus, an OA polygon covering a larger proportion of the buffer and with the highest walkability quartile score of four, for example, would contribute equally to the buffer's mean score as one covering a small proportion and with the lowest walkability quartile score of one. In such a case, the walkability of the buffer could be underestimated. However when this method was tested, it was noted that the minimum number of OAs within a CB was thirty five, the maximum 494 and the median 174. Within a PBNB the minimum, maximum and median were two, 290 and 69, respectively. Thus, weighting scores would not have a substantive effect in most cases. The second approach considered was to intersect buffers with OA polygons, generating new polygons with attributes denoting the buffer area and area of the OA polygons intersecting the buffer, alongside the walkability score of that unit. The advantage of this method was that it rendered weighting of quartile scores possible but the usefulness of score weighting was limited, as explained, and the intersection procedure was more computationally demanding. Therefore, the former method was chosen for assignment of walkability quartile scores to buffers. The walkability scores joined were the quartile scores for the full models and for each component as detailed in Table 7.7. In STATA, the mean walkability scores that had been assigned in ArcGIS for each buffer were converted from decimals to whole numbers.

Table 7.7 Walkability model components and abbreviations.

		Components				
Abbrev	Model	Land use mix 1	Land use mix 2	Land use mix 3	Residential dwelling density	Junction density
M1Q	Full model 1 quartile	✓			✓	✓
M2Q	Full model 2 quartile		✓		✓	✓
M3Q	Full model 3 quartile			✓	✓	✓
L1Q	Land use mix 1 quartile	✓				
L2Q	Land use mix 2 quartile		✓			
L3Q	Land use mix 3 quartile			✓		
RQ	Residential dwelling density quartile				✓	
JQ	Junction density quartile					✓
L1D	Land use mix 1 decile	✓				
L2D	Land use mix 2 decile		✓			
L3D	Land use mix 3 decile			✓		
RD	Residential dwelling density decile				✓	
JD	Junction density decile					✓

7.2.1.5 Validation of walkability index

Five OAs were selected at random using Stata software to represent each of the four full model 1 walkability scores derived for the 24,140 London OAs. The selected dataset, comprising 20 OAs, was exported as an excel spreadsheet which was added as a layer to an ArcMap document. In the ArcMap document it was joined, via the unique OA codes to an OA shapefile. The joined OAs were selected and highlighted for identification on a map. A streetmap baselayer was added to identify the street names within each selected OAs such that they could be identified, viewed and their images saved using Street View technology available in Google Earth [238]. This technology displays panoramas of images taken from cars mounted with nine directional cameras which provide 360 degree views at a height of approximately 3 metres. There is some limited image coverage of places inaccessible by car, such as pedestrianised areas, for which images are taken by cameras mounted on tricycles and other vehicles.

7.2.2 Statistical analysis of neighbourhood areas and walkability scores

Statistical analyses were performed to describe the size of the areas, and numbers, of the constructed spatial units defining neighbourhoods Correlation analyses were done to investigate the similarity of walkability scores between models and also for a given model between different spatial units of enumeration.

7.3 Results

In this section descriptive statistics are first provided to summarise, in terms of their areas and numbers, the constructed spatial units defining neighbourhoods. With regard to the derivation of walkability scores for each neighbourhood spatial unit, the extent to which quantiling the raw scores of the three core walkability dimensions was successful is illustrated by charting the frequencies of each decile, failure to quantile indicated by a frequency above or below ten per cent. Walkability scores are provided for each core dimension and for the full models as deciles and as quartiles, the components and abbreviations of which are given in Table 7.7. Descriptive statistics are provided for all neighbourhood spatial unit walkability scores, including correlations between the various models for a given spatial unit to show, for example, the effect of changing the particular land uses included within the land use mix dimension of the walkability index. Also, correlations between the various spatial models for a given model are reported to illustrate how operationalisation of a neighbourhood of a given postcode point, as an administratively defined area, or as a GIS software-computed gridded cell or residential postcode-centred buffer influenced the walkability score attributed to it. In addition, spatial variation in scores of particular walkability models across London is illustrated as maps for selected spatial units. Finally, photographs of areas of London of various levels of objectively measured walkability are presented alongside their walkability score.

7.3.1 Summary of spatial units defining neighbourhoods

The areal sizes of the various delineations of neighbourhoods in which walkability was calculated, and the number of LWIIP7 participants to whom a neighbourhood could be assigned based on residential location, are summarised in Table 7.8. Comprising 39,125 units, 200m by 200m grid cells defined the most neighbourhoods. The uniform area of 40,000m^2 of each of these grid cells was smaller than the mean area of OAs, explaining their greater number. As the largest units, the uniform areal size exceeding that of even the largest OA, the 2000m by 2000m cells comprised the fewest units, at just 334. Computation of walkability scores within spatial units was dependent on them falling entirely within London because residential dwelling data, a component of the walkability models, was derived from OAs of London only even though geographical coverage of other data used in the calculations extended beyond the boundary. Thus, the neighbourhoods of all individuals in LWIIP7 identified from residential location could be assigned an OA neighbourhood, and as aggregates of OAs, a CAS ward and LA neighbourhood as well. However, there were eleven individuals among the LWIIP7 sample of 3020, represented by ten unique residential postcode geocoordinates whose neighbourhoods could not be defined as 200m by 200m cells, and 223 represented by 234 unique residential postcode geocoordinates whose neighbourhoods could not be defined as 2000m by 2000m cells. Visualisation of these unique residential locations revealed they all fell near the London boundary where cell coverage was incomplete. In the same manner buffers, constituting neighbourhoods delineated around each unique set of residential postcode geocoordinates, were also limited by the London boundary. As shown in Table 7.8 fewer CB neighbourhoods than PBNB ones could be constructed owing to their larger size and, therein their higher propensity to extend beyond the boundary.

Whereas there was no variation in the size of CBs, delineated simply as a Euclidean distance of 1600 metres, there was considerable variation in the areas of PBNBs, a reflection of variation in the density of the road and path network on which their delineation was based; the greater the network density, the greater the area of the PBNB. The relative sizes of neighbourhoods around a geocoded postcode point, as an example exogenous to the LWIIP7 sample, by spatial definition as PBNB, CB and OA is shown in Map 7.6. The area of the CB is seen to exceed that of the PBNB which, in turn exceeds that of the OA. In this case, the postcode point lies in an Inner London borough and the spatial extent of its PBNB, at 2,953,517m^2, is 37 per cent of that of the CB (data not shown). The median area of the 2737 PBNBs from unique geocoordinates from postcodes of the LWIIP7 sample was just under half that of each CB, and their areas ranged between seven and sixty six per cent of those of the CB area. Upon stratification of data pertaining to the buffers by location, statistical analysis revealed the impact of network density and of residential dwelling density on characteristics of buffers. Comparative examination of buffers in Inner and Outer London found that in Inner London the median area of PBNBs was 51% of the CB area but in Outer London it was only 45%, a reflection of the higher network density of Inner London: the greater the network density, the smaller the difference in circular and PBNB areas. The finding that the PBNB for the example postcode in Map 7.6 represents a lower proportion of the CB than of both these median proportions suggests that, despite being in an Inner London borough, it occurs in an area with a lower road and path network density than the majority of the unique geocoordinates from postcodes of the LWIIP7 sample. In addition to impacting PBNB area, location of buffers determined the number of OAs overlapped. The higher median numbers of OAs in CBs and PBNBs in Inner London, 243 and 131 respectively, than in Outer London, 123 and 28 respectively, reflected the higher residential density of Inner London which contains smaller OAs. Thus the location of PBNBs determined not only the number of OAs hosted but also the buffer area, whereas the location of CBs determined only the number of OAs. However, the effect of location-dependent variation in numbers of OAs hosted by buffers on walkability scores calculated as means of OA scores was likely to be low given that even in Outer London the median numbers of OAs hosted by both CBs and PBNBs was quite high.

Statistical examination of the spatial units shows that the delineation of the neighbourhoods in which walkability was calculated dictated the number of LWIIP7 participants to which each type of neighbourhood could be assigned based on residential location. The areal sizes of the spatial units of neighbourhood delineation also impacted on ability to derive walkability scores for each. Table 7.9 charts frequencies (per cent) of walkability component deciles by walkability enumeration unit, illustrating the extent to which quantiling the raw scores of the three core walkability dimensions was successful. Land use mix entropies could be neither quantiled into deciles for OAs nor for the 200m by 200m GCs as more than ten per cent of these spatial units had entropies of zero. An entropy of zero was the manifestation of a spatial unit containing none of the land uses specified in a particular land use model, or that of a spatial unit being entirely covered by one land use. The small average sizes of these spatial units relative to the other delineations of neighbourhoods accounts for high proportions scoring zero entropies: the smaller the areal size the greater the likelihood the spatial unit would contain a single specific land use and, if a spatial unit did contain a relevant land use, the higher the chance the land use area would exceed that of the unit. With regard to residential dwelling and junction density quantilation, failure only occurred within neighbourhoods defined as 200m by 200m GCs, the result of a high proportion – greater than ten per cent – of these spatial units containing no junctions or residential dwellings. The problem did not arise for OAs because, as administrative areas, these spatial units contained residential dwellings, and therefore road and path junctions, regardless of size. The non-perfect "tens" shown for LA component decile frequencies reflects the low number of these spatial units across London, at thirty three.

Table 7.8 Areal sizes of neighbourhoods in which walkability was calculated and number of LWIIP7 participants to whom a neighbourhood could be assigned based on residential location.

Neigh'd	N	N (%) assig'd to LWIIP 7	Min area (m^2)	Max area (m^2)	Mean area (m^2)	SD (m^2)	Median area (m^2)
OA	24,140	3020 (100.0)	128	12,300,000	66,061	34,870	229,621
CAS ward	633	3020 (100.0)	129,865	29,000,000	2,519,305	2,577,294	1,844,511
LA	33	3020 (100.0)	3,151,466	150,000,000	48,300,000	32,800,000	38,700,000
200m by 200m GC	39,125	3009 (99.6)	40,000	40,000	40,000	0	40,000
2000m by 2000m GC	334	2786 (92.3)	4,000,000	4,000,000	4,000,000	0	4,000,000
1600m PBNB	2737	2860 (94.7)	600,577	5,311,818	3,678,976	713,310	3,795,421
1600m CB	2633	2751 (91.1)	8,042,477	8,042,477	8,042,477	0	8,042,477

7.3.2 Correlations between walkability model scores by spatial units of enumeration

Correlations between the three land use mix model scores were positive and high for all spatial units, and were higher between models 1 and 2, and between models 2 and 3, than between models 1 and 3 for which the difference in the number of land uses included was the greatest. Results are given for the administratively-defined spatial units of OAs, CAS wards and LAs in Table 7.10 and for the GCs in Table 7.11. In Table 7.10 correlations are given for walkability scores as deciles for each land use mix only model (L1D, L2D and L3D), for the residential dwelling density only model (RD) and the junction density only model (JD), and for full walkability model scores as quartiles (M1Q, M2Q and M3Q). In

Table 7.11 correlations are given for walkability scores as deciles for each land use mix only model (L1D, L2D and L3D), and for the residential dwelling density only model (RD) and the junction density only model (JD). Only for 2000m by 2000m GCs are correlations given for full walkability model scores as quartiles (M1Q, M2Q and M3Q). Given that buffer scores were derived from OA scores, correlations between walkability model scores for these spatial units are not given. For all spatial units, correlations between each of the land use mix models and residential dwelling density tended to be positive but low with the exception of that between land use mix model 1 and residential dwelling density in LAs, with a correlation coefficient of 0.53 indicating moderate positive correlation. The same trend was seen for junction density. Residential dwelling density and junction density showed moderately to highly positive correlations for the larger spatial units of 2000m by 2000m GCs, CAS wards and LAs. Correlations between the full models, which differed only in their land use mix component, were very high for all spatial units with coefficients tending to be greater than 0.90. This was because the residential dwelling density and junction density components, the scores of which contributed to the same extent as that of the land use mix component, did not vary between models.

7.3.3 Correlations between land use mix dimension walkability scores of spatial units of enumeration by land use mix model

Correlations were calculated between walkability scores of the land use mix dimension of spatial units of enumeration by land use mix model (Table 7.12). They were calculated for each walkability score between OAs and CAS wards, between OAs and LAs, and between OAs and 2000m by 2000m GCs. In addition correlations were calculated for each walkability score between CAS wards and LAs, and between CAS wards and 2000m by 2000m GCs. Finally, they were calculated between LAs and 2000m by 2000m GCs. The correlations between scores for any two different spatial units were stronger for land use mix model 1, than for land use mix models 2 and 3 which included more types of land use. This was because the potential for variation in scores grew with the introduction of more land use types. Weak positive correlations were found between OAs and all other neighbourhoods, a likely consequence of the quantilation problem for land use mix entropy within OAs described previously. Stronger positive correlations were found between CAS wards and LAs, and between LAs and 2000m by 2000m GCs. The strongest positive correlations were between CAS wards and 2000m by 2000m GCs. This was possibly because they were approximately the same size.

7.3.4 Correlations between neighbourhood delineation land use mix dimension walkability scores by walkability model

Whereas administrative areas and (grid cells) GCs were used as the spatial units of enumeration for construction of the walkability index, the circular and PBNBs were specifically constructed as neighbourhoods around residential postcode and their walkability scores derived from those of OAs. A spatial unit of enumeration could also define the neighbourhood of a LWII7 participant whose residential postcode fell within it. Therefore, it was useful to measure the extent to which the scores of the various neighbourhood definitions correlated for the LWIIP7 sample. Correlation coefficients between scores for buffers and for the other neighbourhood delineations are given in Table 7.13 but correlations between different administrative units and grid cells are excluded as their values are similar to those presented in Table 7.12 and calculated from all units across London. Between a buffer and any other neighbourhood delineation, score correlations were stronger for L1Q, than for L2Q and L3Q. The correlations between the land use mix dimension walkability scores of CBs and PBNBs were moderately high but the correlation between each buffer type and OAs was weak, reflecting the much larger area of buffers relative to OAs, as illustrated in Map 7.6, and suggesting wide variation in the scores of the OAs from which each buffer score was derived, as the mean of OA scores. The correlations between scores of each buffer type and those of the other neighbourhood delineations were weak but stronger than those between each buffer type and OAs. This was due to the fact that the other neighbourhood delineations, such as LAs, were closer in size to those of the buffers, as illustrated in Map 7.7.

7.3.5 Spatial variation in walkability model scores across London, by spatial units of enumeration

Spatial variation in walkability is presented for OAs across London as L1Q, L2Q, L3Q, RQ, JQ, M1Q, M2Q and M3Q in Map 7.8, and as L1D, L2D, L3D, RD, JD for CAS wards, LAs, 200m by 200m grid cells and 2000m by 2000m grid cells in Map 7.9, Map 7.10, Map 7.11 and Map 7.12, respectively. In addition, spatial variation in net differences between L1Q and L2Q, between L1Q and L3Q, between M1Q and M2Q, and between M1Q and M3Q, for CAS wards are illustrated in Map 7.13 with white indicating there is no difference and darker shades indicating greater difference.

The OA level spatial variation in walkability scores shown in Map 7.8 appears similar for the three land use mix only models, L1Q, L2Q and L3Q, with higher walkability as indicated by these models in Inner London and around the periphery. As discussed in Section 7.3.1, however, the land use mix entropies calculated for OAs could not be quantiled into deciles, from which quartile scores were derived. Thus, there is an uneven distribution of quartiles and consequently a lack of variation between models. With respect to walkability modelled as residential dwelling density, OAs in Inner London tend to have higher walkability than Outer London ones. OAs are delineated to contain approximately the same number of people and residential dwelling density was calculated as the number of dwellings per area of land allocated as residential. Although theoretically the area of land allocated as residential could be the same regardless of OA size, it was likely that larger OAs would have a greater area of residential land. Indeed, lower residential dwelling density, indicative of lower walkability, was found in larger OAs of Outer London. In the core City of London region, the central business district, however, walkability with respect to this dimension is evidently low. Junction density within OAs is less concentrated within Inner London, instead showing an unsurprisingly similar patterning to that of road networks. The spatial distributions of OA walkability scores as indicated by the full models M1Q, M2Q and M3Q appear very similar as expected given the similarity in distributions of L1Q, L2Q and L3Q, the only dimensions by which the full models varied. All three full models for OAs show Inner London to be more walkable, with a radial decay in walkability into Outer London. The spatial variation in land use mix walkability scores for the higher administrative units of CAS wards, shown in Map 7.9, is model dependent, in contrast to that of OAs. Whilst walkability in terms of land use mix is high in Inner London for all 3 models, the outskirts of London become relatively more walkable as “Entertainment, culture and recreation” and “Free recreational land” land uses are introduced to L2D and L3D. This can be attributed to the greater areas of open spaces within these regions into which these categories of land use fall. CAS ward level variation in walkability as modelled by residential dwelling density and junction density is similar to that seen for OAs. The spatial variation in net differences between walkability model scores for CAS wards (Map 7.13) demonstrate the higher correlations between the full walkability models than between land use mix only models as shown in Table 7.10. The equal weightings contributed by residential dwelling density and junction density to the full scores reduce any land use mix score difference. Although, variation LA level walkability is found to be land use mix model dependent as for CAS wards, with respect to this dimension of walkability, patterning is less evident, a consequence of LAs being larger areal units of data aggregation (Map 7.10).

The land use mix decile quantilation problem illustrated in the OAs of Map 7.8 can also be seen in the 200m by 200m grid cells of Map 7.11, with large areas of white representing areas of zero land use mix entropy within these particularly small spatial units. However, given that these units were of uniform size and their areas did not deviate to the particularly small sizes of some OAs, there were discernible differences in the spatial variation in land use mix models across London similar to those found for CAS wards: the gradient of radial decay in walkability from Inner to Outer London reduced in L2D and further in L3D. As for LAs, the 2000m by 2000m grid cells constituted large spatial units of data aggregation which rendered spatial patterning of walkability unclear (Map 7.12).

7.3.6 Validation of walkability index

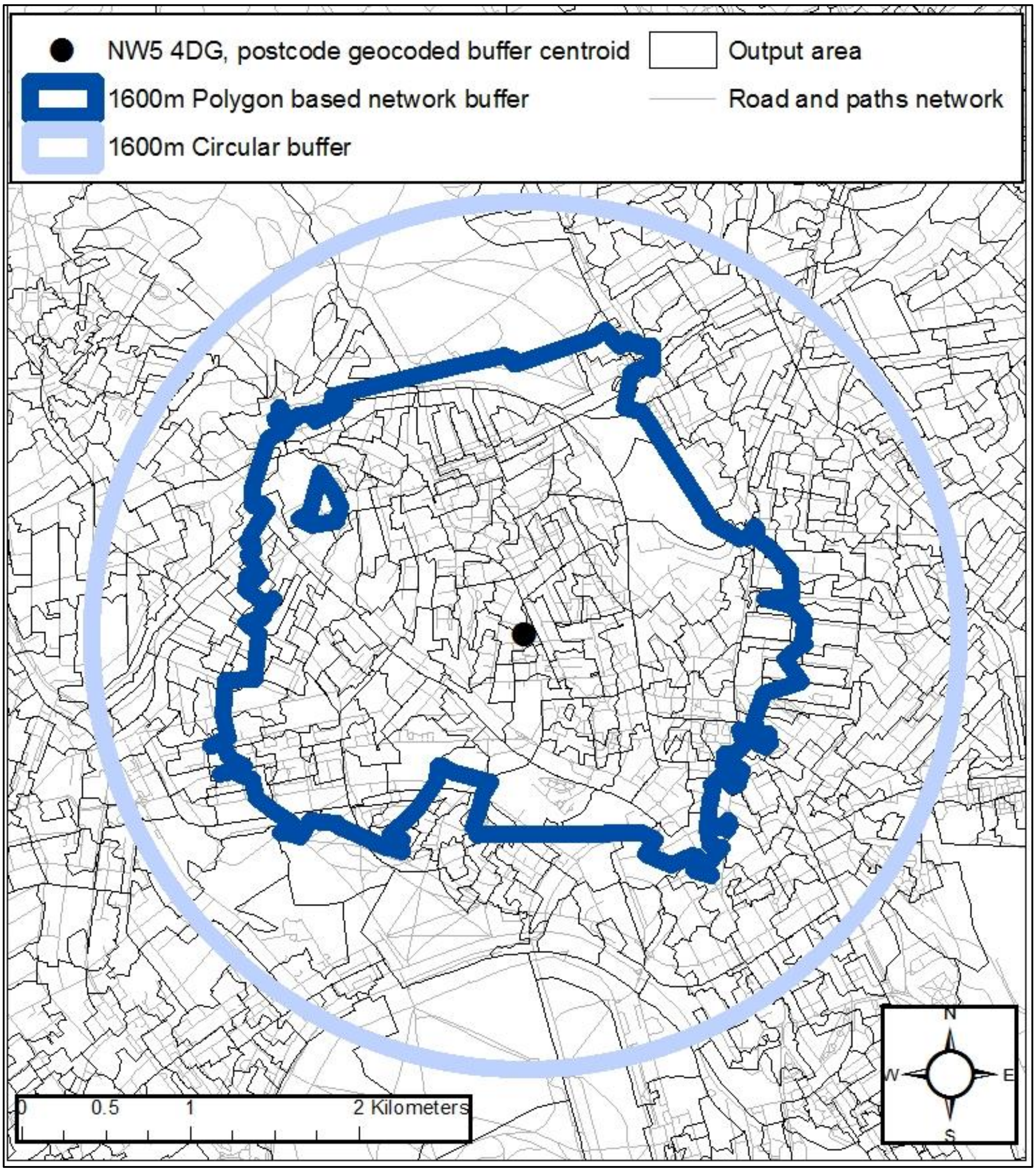
For each OA that was randomly selected using Stata as an example for validating the walkability index, two Google Earth images are provided; an aerial view approximating the extent of the particular OA and a Google Street View. Images are presented in Figure 7.3, Figure 7.4, Figure 7.5 and Figure 7.6 in sets of five with OA level full walkability model 1 (M1Q) scores of 1, 2, 3 and 4 respectively. In each figure the left image is an aerial view whilst the right one is a Google Street View.

Along with images for each of the twenty different OAs, OA walkability scores are given for the three full walkability models – comprising residential dwelling density, junction density and one of the three land use mix models. In addition to those given for OAs, the scores are also provided for the other spatial units in which walkability was measured and in which the OA lay, namely 2000m by 2000m grid cells (GC2000), CAS wards and LAs, not necessarily equalling the OA scores. The numbers overlaying the images in Figure 7.3, Figure 7.4, Figure 7.5 and Figure 7.6 – representing OAs with OA level full walkability model 1 (M1Q) scores of 1, 2, 3 and 4 respectively – constitute these other scores. The walkability model specificity of these scores is found upon alignment with the header for each figure. In Figure 7.3, for example, showing images of the 5 randomly selected OAs with OA M1Q scores of 1, Northumberland Heath in Bexley (OA code 00ADGS0028) also has OA full walkability model M2Q and M3Q scores of 1. In fact, for most neighbourhood spatial units in which this OA falls and for all full walkability models, the walkability score is 1, indicative of the least walkable neighbourhoods. However, Northumberland Heath has a score of 2 indicating slightly greater walkability for the CAS ward in which it falls when the model is specified as either MQ1 or MQ2. The contrast between the walkability scores for different spatial units and walkability models is starker in other cases. For example, in Figure 7.3, for Little Venice in Westminster (OA code 00BKGK019) the OA full walkability model M1Q score of 1 is much lower than the OA scores of 3 for M2Q and M3Q, which differ only in their land use mix dimensions. Also, the walkability of this OA, as indicated by walkability scores of 3 and 4 for the higher spatial units in which it falls – GC2000, CAS wards and LAs – is much greater. Comparison of the Google Street View image for Northumberland Heath with that of Little Venice may provide insight into the more marked disparity in walkability scores for the latter OA. The high rise residences visible in the Little Venice Google Street View indicate higher residential density and possibly greater land use mix, which belie its low OA M1Q walkability score of 1. However, at the fine spatial scale of OAs and with only the basic land use mix included in M1Q, walkability may be assessed as low because the OA area is too small to contain a high mix of land uses. In contrast, the Google Street View image for Northumberland Heath shows a suburban street suggestive of the low residential density, low junction density and low mix of land use which give rise to consistently low walkability scores for OA and higher geographies.

The majority of the aerial images of randomly selected OAs show residential areas. All but one of the Google Street View images taken at ground level are photographs of streets, the vehicle mounted camera access limited in other terrains such as parks.

In the author's opinion, relative walkability is not discernible by comparison of the four sets of five Google Street Views, each set representing a distinct M1Q score at OA level. Most of these views are of residential streets with similarly low levels of pedestrian and vehicular traffic, if any, probably due to the shots being taken during office hours. On first inspection, the similarity of the aerial images regardless of walkability score is also striking, with the vast majority showing neat rows of residential streets, but given the residential density of the London region it is logical that shots of randomly selected OAs will be residential. However, closer examination reveals differences consistent with the walkability scores assigned to them. The aerial photographs give birds' eyes views of OAs which are not those of a person walking on the ground so constitute a less valuable tool for a subjective evaluation than ground-level, street ones. Nevertheless, aerial views are evaluated by the author, with knowledge of objective walkability dimensions, as more indicative of walkability: the OAs with M1Q scores of 1 and 2 appear less diverse in terms of land use than those with scores of 3 and 4. It is interesting that the OAs with the lowest M1Q walkability score of 1 contain less green area than higher scoring OAs despite this model excluding "free recreational land" (which would include green space). This finding aligns with the statistical analysis which showed high correlation between the land use mix models. Differences in residential dwelling density and in junction density - which could be subjective as well as objective indicators of walkability - between OAs differing by M1Q scores are not readily apparent, perhaps due to technical difficulties in matching the scales of aerial shots.

Map 7.6 Example of relative sizes of neighbourhoods around a geocoded postcode point by spatial definition as PBNB, CB and OA. OAs were generally considerably smaller than PBNBs and CBs.



Map 7.7 Example of relative sizes of neighbourhoods around a geocoded postcode point by spatial definition as PBNB, CB and LA.

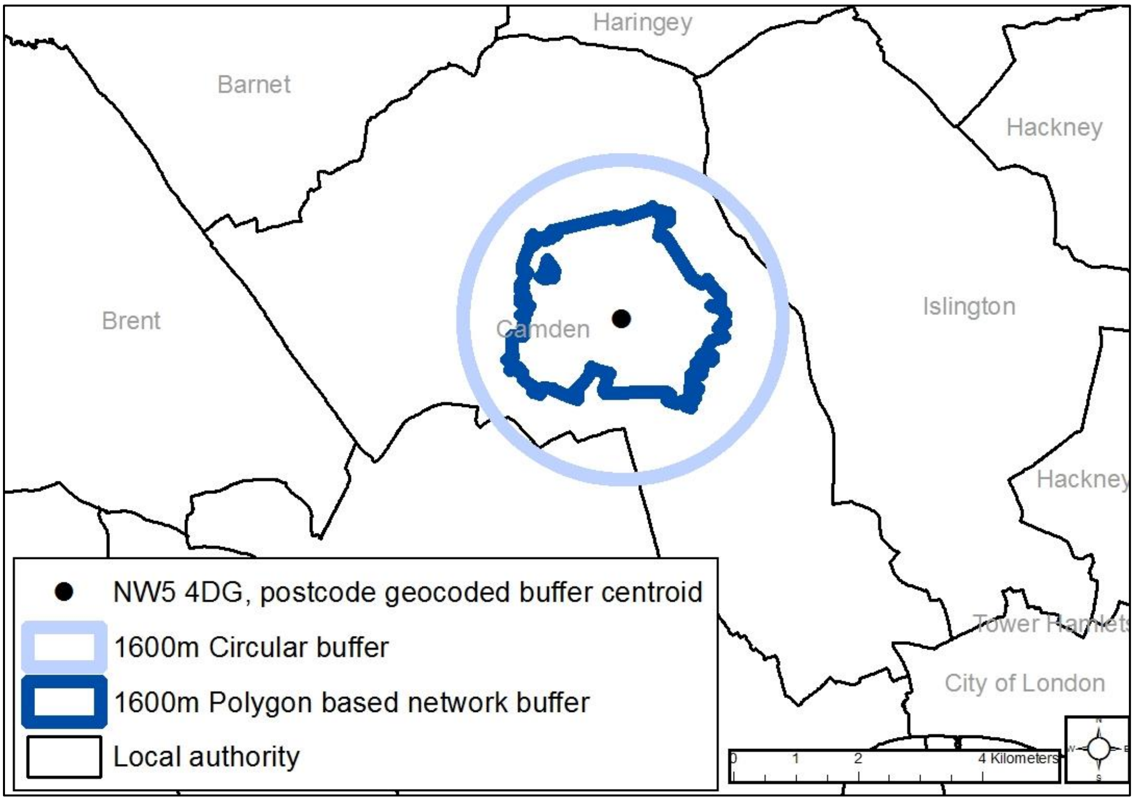


Table 7.9 Frequencies (per cent) of walkability component deciles by spatial unit of enumeration.

Spatial unit	OA					2000m by 2000m grid cell					200m by 200m grid cell					LA					CAS Ward				
Walkability component	L1	L2	L3	RD	JD	L1	L2	L3	RD	JD	L1	L2	L3	RD	JD	L1	L2	L3	RD	JD	L1	L2	L3	RD	JD
Decile 1	32	28	19	10	10	10	10	10	10	10	55	48	34	54	29	12	12	12	12	12	10	10	10	10	10
Decile 2	0	0	1	10	10	10	10	10	10	10	0	0	0	0	0	9	9	9	9	9	10	10	10	10	10
Decile 3	0	2	10	10	10	10	10	10	10	10	0	0	0	0	11	9	9	9	9	9	10	10	10	10	10
Decile 4	8	10	10	10	10	10	10	10	10	10	0	0	6	0	0	12	12	12	12	12	10	10	10	10	10
Decile 5	10	10	10	10	10	10	10	10	10	10	0	2	10	0	17	9	9	9	9	9	10	10	10	10	10
Decile 6	10	10	10	10	10	10	10	10	10	10	5	10	10	6	6	9	9	9	9	9	10	10	10	10	10
Decile 7	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	12	12	12	12	12	10	10	10	10	10
Decile 8	10	10	10	10	10	10	10	10	10	10	10	10	10	10	7	9	9	9	9	9	10	10	10	10	10
Decile 9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	9	9	9	9	9	9	10	10	10	10	10
Decile 10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	9	9	9	9	9	10	10	10	10	10

Table 7.10 Correlations between component and full walkability scores by administrative areas as spatial units of enumeration.

Spatial unit	OA								CAS Ward								LA							
Walkability component	L1D	L2D	L3D	RD	JD	M1Q	M2Q	M3Q	L1D	L2D	L3D	RD	JD	M1Q	M2Q	M3Q	L1D	L2D	L3D	RD	JD	M1Q	M2Q	M3Q
L1D	1								1								1							
L2D	0.91	1							0.76	1							0.84	1						
L3D	0.78	0.87	1						0.64	0.88	1						0.67	0.87	1					
RD	0.21	0.16	0.10	1					0.19	0.14	0.26	1					0.53	0.33	0.22	1				
JD	0.06	0.03	0.06	0.26	1				0.38	0.13	0.04	0.59	1				0.71	0.53	0.28	0.80	1			
M1Q	0.64	0.56	0.48	0.66	0.59	1			0.63	0.28	0.14	0.74	0.84	1			0.82	0.66	0.48	0.85	0.90	1		
M2Q	0.59	0.60	0.52	0.66	0.61	0.96	1		0.60	0.41	0.25	0.70	0.83	0.91	1		0.78	0.68	0.49	0.83	0.91	0.96	1	
M3Q	0.51	0.52	0.57	0.64	0.64	0.90	0.93	1	0.59	0.39	0.33	0.65	0.84	0.89	0.92	1	0.78	0.68	0.56	0.86	0.87	0.97	0.95	1

Table 7.11 Correlations between component and full walkability scores by grid cell (GC) as spatial units of enumeration.

Spatial unit	2000m by 2000m grid cell								200m by 200m grid cell							
Walkability component	L1D	L2D	L3D	RD	JD	M1Q	M2Q	M3Q	L1D	L2D	L3D	RD	JD	M1Q	M2Q	M3Q
L1D	1								1							
L2D	0.79	1							0.88	1						
L3D	0.72	0.89	1						0.72	0.82	1					
RD	0.21	0.03	-0.07	1					0.34	0.30	0.23	1				
JD	0.41	0.27	0.22	0.74	1				0.45	0.43	0.42	0.41	1			
M1Q	0.64	0.42	0.34	0.77	0.87	1			n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M2Q	0.57	0.49	0.38	0.75	0.88	0.92	1		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M3Q	0.56	0.48	0.43	0.73	0.88	0.90	0.95	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

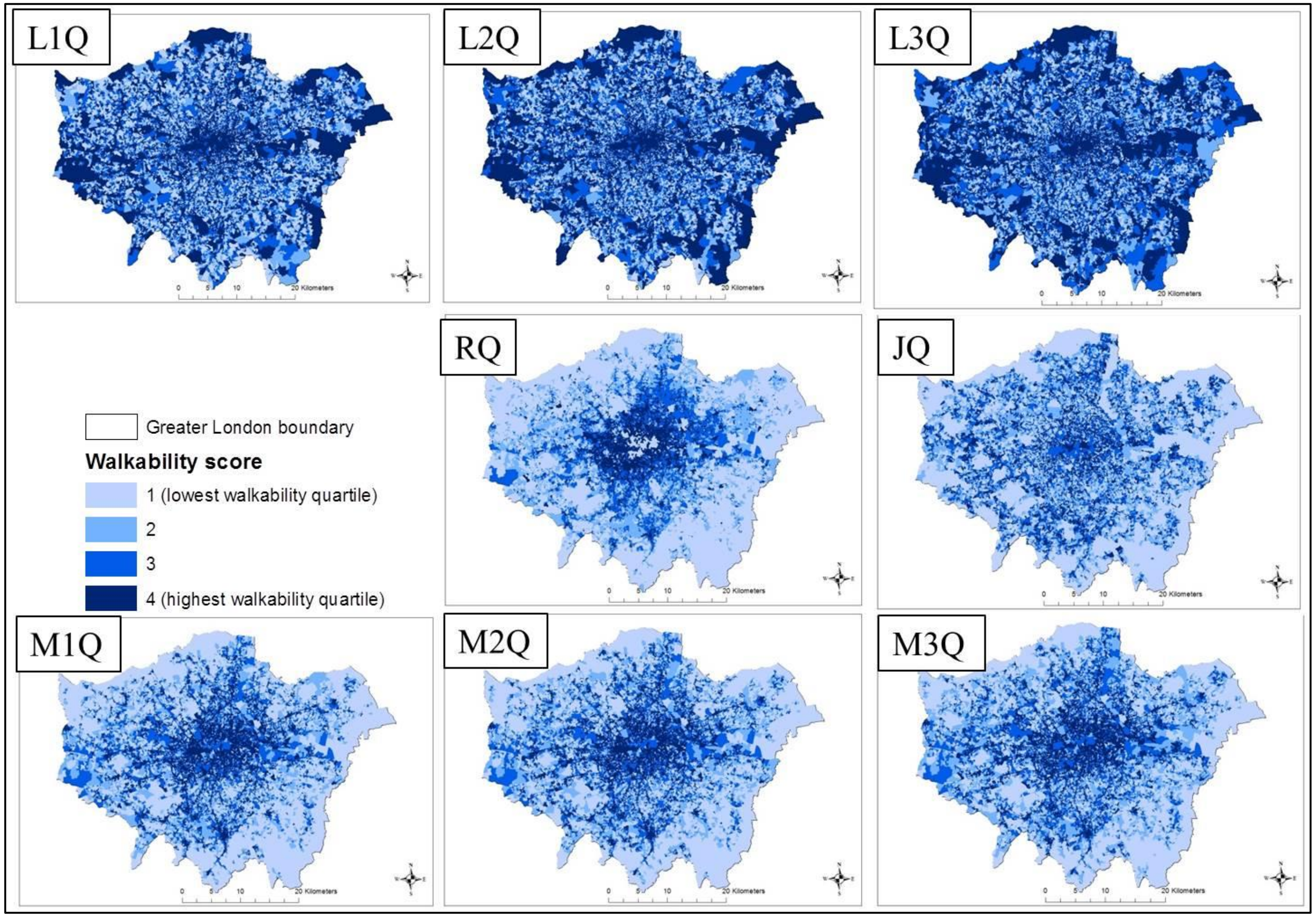
Table 7.12 Correlations between administrative spatial units of enumeration by land use mix decile score for each land use mix model.

Spatial unit	Walkability component	OA	CAS Ward	LA
CAS Ward	L1D	0.24	n/a	n/a
CAS Ward	L2D	0.19	n/a	n/a
CAS Ward	L3D	0.19	n/a	n/a
LA	L1D	0.13	0.45	n/a
LA	L2D	0.11	0.34	n/a
LA	L3D	0.08	0.28	n/a
2000m by 2000m GC	L1D	0.21	0.67	0.53
2000m by 2000m GC	L2D	0.16	0.53	0.46
2000m by 2000m GC	L3D	0.16	0.52	0.38

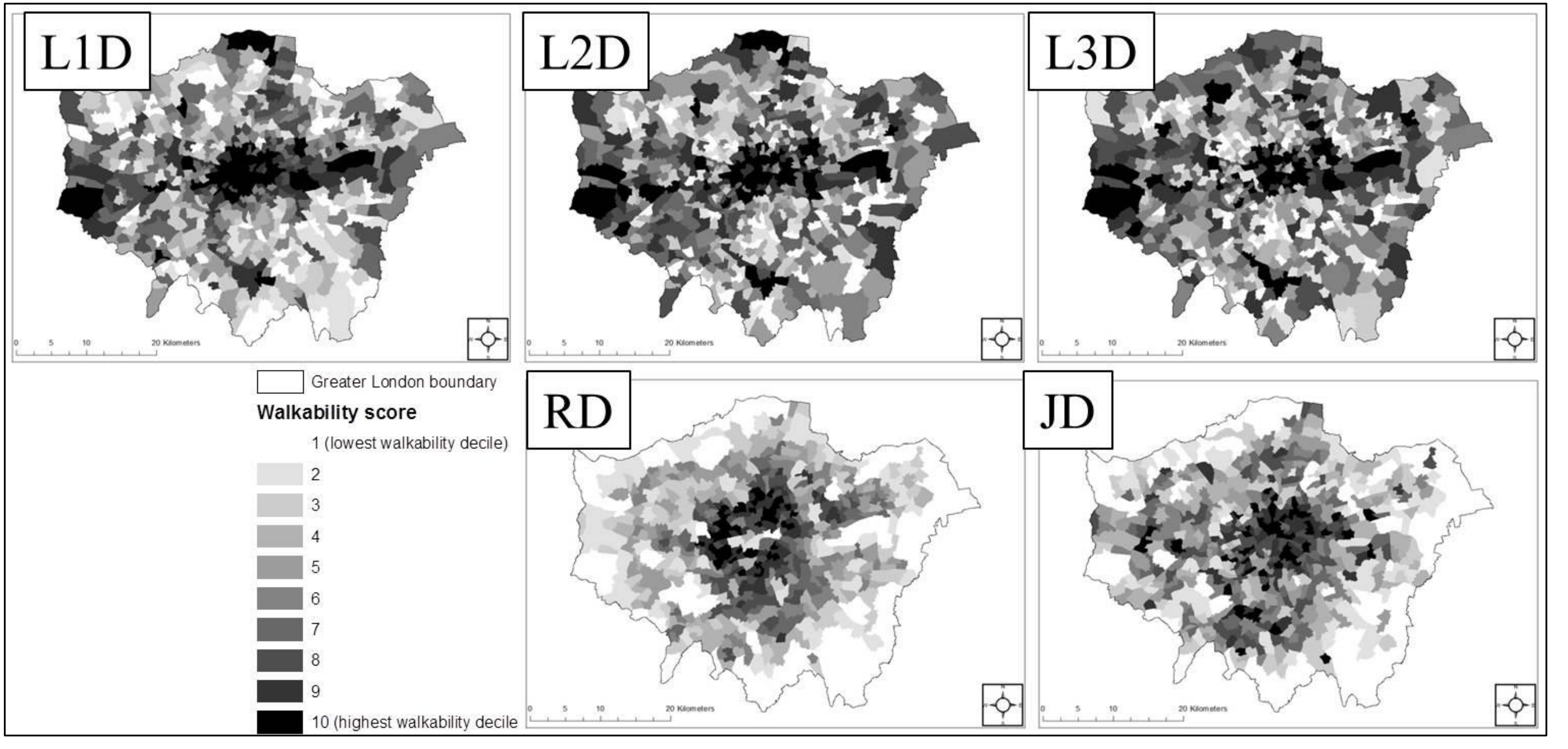
Table 7.13 Correlations between neighbourhood delineation land use mix dimension walkability scores by walkability model.

Neighbourhood delineation	Walkability component	PBNB	CB
CB	L1Q	0.70	n/a
CB	L2Q	0.65	n/a
CB	L3Q	0.63	n/a
OA	L1Q	0.24	0.18
OA	L2Q	0.21	0.16
OA	L3Q	0.20	0.18
CAS Ward	L1Q	0.51	0.52
CAS Ward	L2Q	0.35	0.32
CAS Ward	L3Q	0.37	0.36
LA	L1Q	0.37	0.42
LA	L2Q	0.34	0.40
LA	L3Q	0.24	0.29
2000m by 2000m GC	L1Q	0.51	0.53
2000m by 2000m GC	L2Q	0.40	0.43
2000m by 2000m GC	L3Q	0.36	0.42

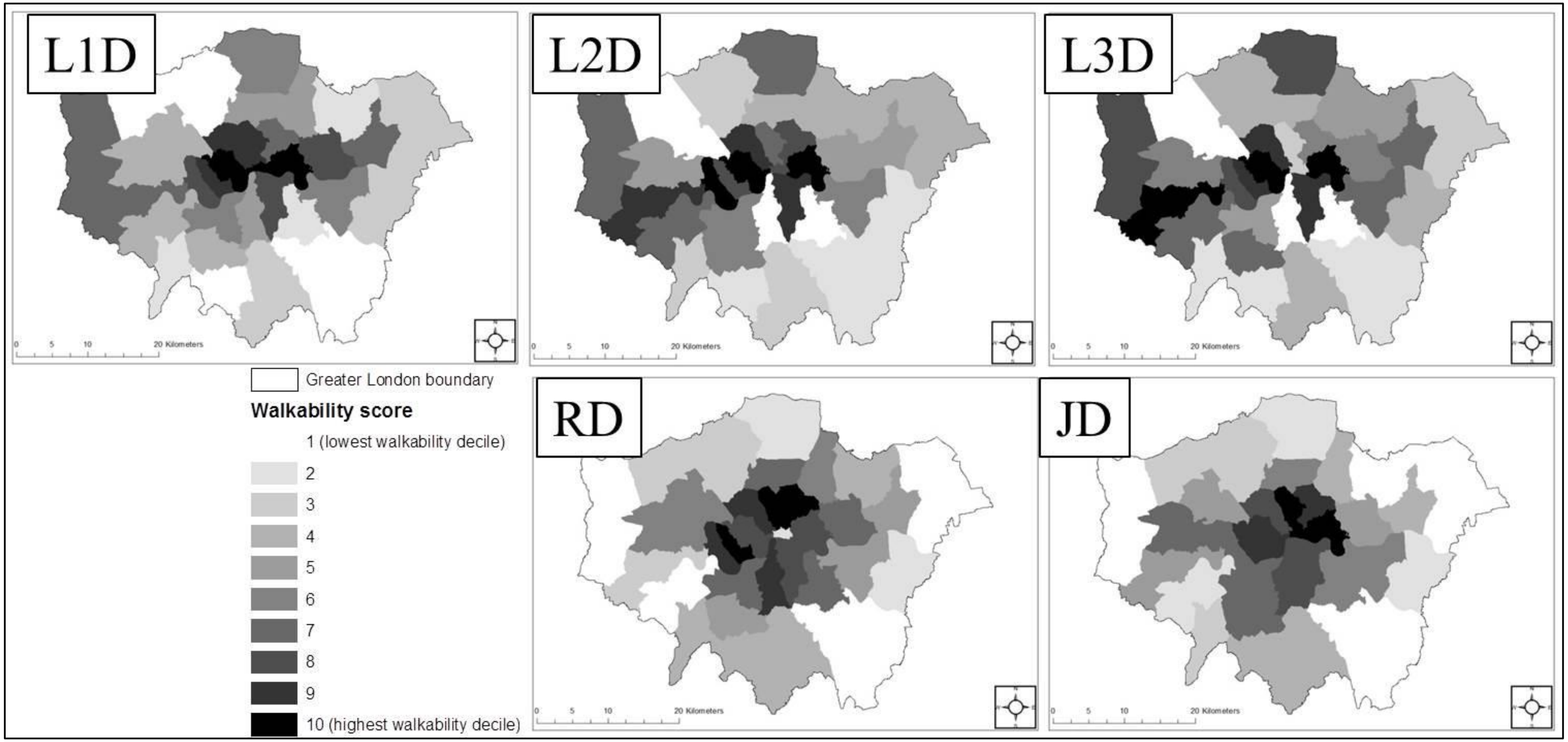
Map 7.8 Walkability of OAs as the land use mix only models (L1Q, L2Q and L3Q) the residential dwelling density model only (RQ), the junction density model only (JQ) and as the full models (M1Q, M2Q and M3Q).



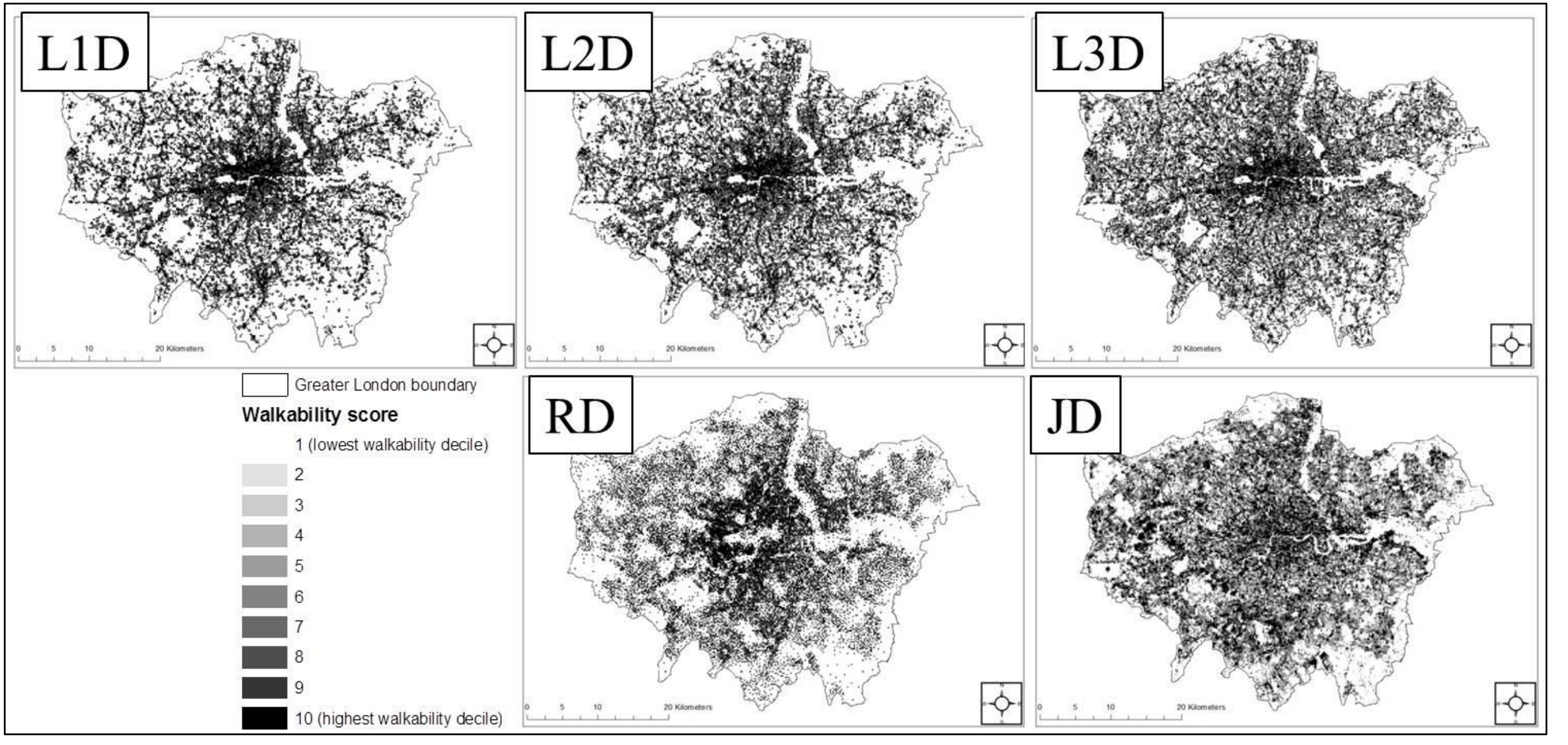
Map 7.9 Walkability of CAS wards as the land use mix only models (L1D, L2D and L3D), the residential dwelling density model only (RD) and the junction density model only (JD).



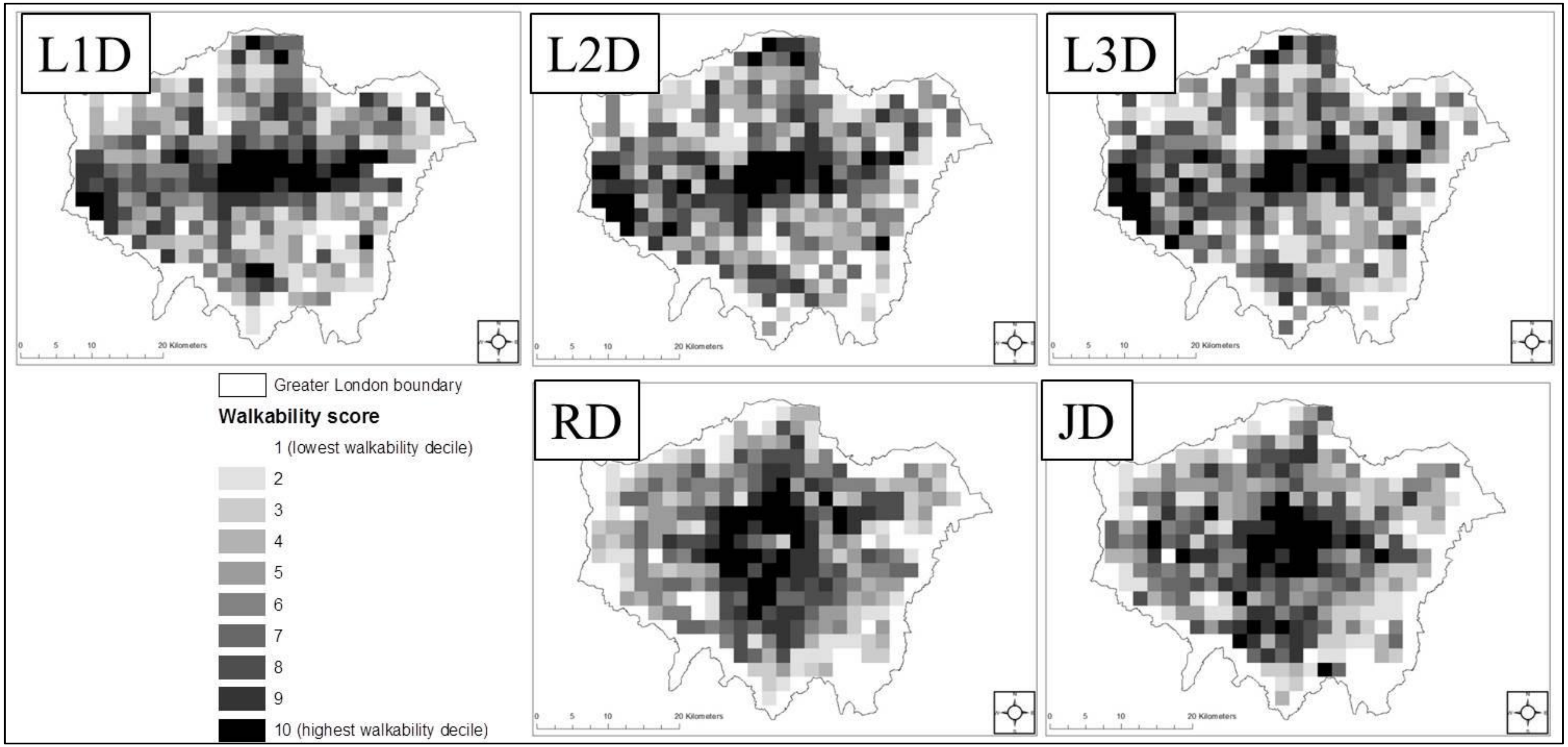
Map 7.10 Walkability of LAs as the land use mix only models (L1D, L2D and L3D), the residential dwelling density model only (RD) and the junction density model only (JD).



Map 7.11 Walkability of 200m by 200m grid cells as the land use mix only models (L1D, L2D and L3D), the residential dwelling density model only (RD) and the junction density model only (JD).



Map 7.12 Walkability of 2000m by 2000m grid cells as the land use mix only models (L1D, L2D and L3D), the residential dwelling density model only (RD) and the junction density model only (JD).



Map 7.13 Difference in CAS ward walkability scores between L1Q and L2Q, between L1Q and L3Q, between M1Q and M2Q, and between M1Q and M3Q.

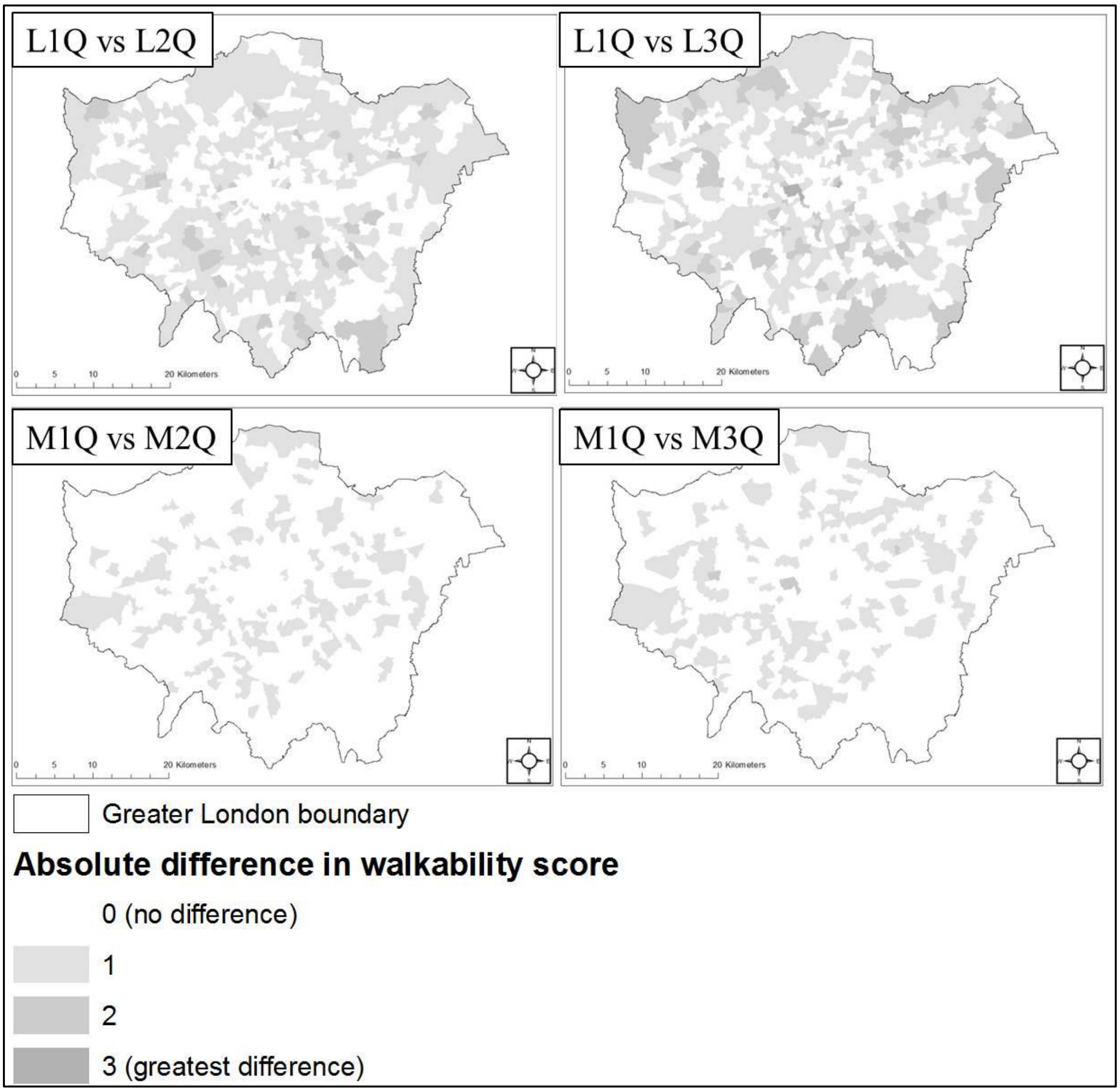


Figure 7.3 Google Earth images of five randomly selected examples of OAs, all scoring 1 (least walkable) for the OA level full walkability model 1 (M1Q), for validating the walkability index.

Neighbourhood spatial unit	OA			GC2000			CAS ward			LA		
Full walkability model:	1	2	3	1	2	3	1	2	3	1	2	3

00ADGS0028
Northumberland Heath, Bexley

00BAGN0022
St Helier, Merton

00AFGZ0043
Plaistow and Sundridge, Bromley

00BKGK0019
Little Venice, Westminster

00BHGF0031
Hatch Lane, Waltham Forest

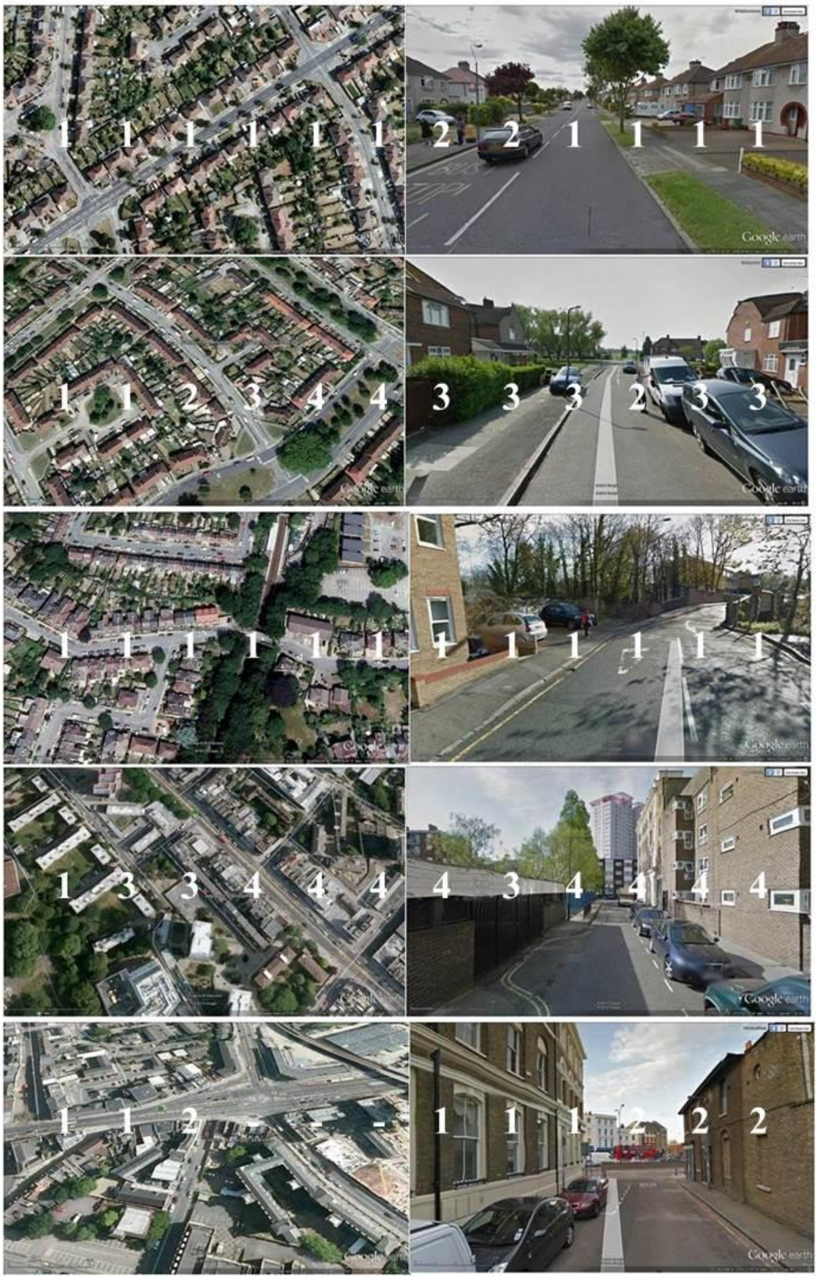


Figure 7.4 Google Earth images of five randomly selected examples of OAs, all scoring 2 for the OA level full walkability model 1 (M1Q), for validating the walkability index.

Neighbourhood spatial unit	OA			GC2000			CAS ward			LA		
Full walkability model:	1	2	3	1	2	3	1	2	3	1	2	3

00AHGZ0055
South Norwood, Croydon

00AUGF0008
Hillrise, Islington

00ALGP0002
Abbey Wood, Greenwich

00BDGB0009
Hampton North, Richmond upon Thames

00ACGP0022
Totteridge, Barnet

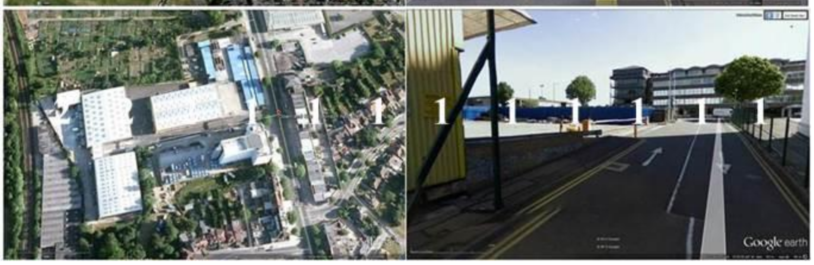


Figure 7.5 Google Earth images of five randomly selected examples of OAs, all scoring 3 for the OA level full walkability model 1 (M1Q), for validating the walkability index.

Neighbourhood spatial unit	OA			GC2000			CAS ward			LA		
Full walkability model:	1	2	3	1	2	3	1	2	3	1	2	3

00ADGC0030
Blackfen and Lamorbey, Bexley

00BCGG0015
Fairlop, Redbridge

00AZGX0035
Whitefoot, Lewisham

00BGGH0011
St Dunstan's and Stepney Green, Tower Hamlets

00ATGD0017
Feltham North, Hounslow

Figure 7.6 Google Earth images of five randomly selected examples of OAs, all scoring 4 (most walkable) for the OA level full walkability model 1 (M1Q), for validating the walkability index.

Neighbourhood spatial unit	OA			GC2000			CAS ward			LA		
Full walkability model:	1	2	3	1	2	3	1	2	3	1	2	3

00BFGE0020

Belmont, Sutton

00AJGY0006

Southall Broadway, Ealing

00BAGJ0024

Merton Park, Merton

00AMGQ0017

Queensbridge, Hackney

00AZGF0031

Brockley, Lewisham

7.4 Discussion

The overall aim of this study was to construct and evaluate a walkability index as a measure of the physical environment pertinent to physical activity and depression. Here the practical challenges of construction of the walkability index are summarised and the effects on walkability scores of inclusion of different in land use mixes in the walkability models are described. This section also reflects on the spatial units in which walkability was enumerated and how neighbourhoods were defined, and the relationship of a spatial unit's size and shape with its walkability score. Also, the extent to which the stated hypotheses held true with regards to the spatial patterning of walkability across London as measured by the walkability index is discussed. Finally, this chapter summarises the ability of the objective walkability index to capture subjective walkability of areas.

7.4.1 Practical challenges of construction of the walkability index

The main practical challenges in constructing a walkability index for London lay in management of geographical data within separate geographical and software packages. It was necessary to conduct geoprocessing, such as the spatial joining of particular land use polygons to OAs polygons within GIS software so that attribute data, including the type and area of the land use polygon, could be assigned to the OA polygon. Because the statistical capabilities of the GIS software were limited, the data had to be exported in spreadsheet format, and then converted to a form compatible with, and imported into separate statistical software. Mathematical operations such as the calculation of land use mix as entropy scores, or the merging of spatial data with geocoded census household count data could then be carried out. Transferring and reformatting data is error prone so ideally both geoprocessing and statistical manipulation would have been performed within one software package. Given that construction of walkability indices is likely a relatively uncommon use of geographical and statistical software, and there was no published protocol for the management of data, it was self-taught through time-consuming trial and error in the present study. Additionally, operations performed within both GIS and statistical software packages were computationally demanding and hence slow due to the large geographical dataset: spatially joining in excess of 11 million land use polygons to London's 240,140 OAs took many hours.

7.4.2 Land use mix and walkability scores

Correlation analysis suggested that whatever the spatial unit of enumeration, its walkability model scores were similar regardless of the land uses included in them. Thus, an area with a high score for the basic model (L1) which included only the basic ones defined as "Residential" "Retail" "Office" "Health, welfare & community", would also have a high score for L2 – extending L1 to include "Entertainment, culture and recreation" – and for L3 – extending L2 to include "Free recreational land". This suggested that wherever basic land uses were well mixed, so too were other land uses hypothesised to facilitate recreational walking and other physical activity. This co–location of land uses fits Wegener and Franz's "Land-use transport feedback cycle" in which land development sparks further development [239]. As

Batty *et al* describe, the number of land uses that a place can contain is dependent on the size of that place [120]. It is also dependent on the size of the land uses so land uses that have relatively small footprints such as shops, offices and houses in L1, and recreational facilities in L2, offer potential for heterogeneity even in the smallest areas. However, land uses that tend to be larger, such as the parks and public open spaces included in L3, may reduce heterogeneity by occupying large proportions of areas and "pushing" other land uses out. Conversely, the inclusion of this "free recreational land" in the land use models may increase the potential for heterogeneity in larger areas, and therefore higher walkability scores. Considerable intra-administrative unit area size variation was found and this may explain the finding that for a given administrative area level walkability score correlations were higher between L1 and L2 than between L1 and L3.

7.4.3 Neighbourhoods delineations and walkability scores

It was hypothesised that the correlation of walkability scores of neighbourhoods that were more similar in scale would be higher and this tendency was observed. Correlations between OAs and all other spatial units of enumeration, and between OAs and buffers, were found to be particularly low, a reflection of their much smaller scale. It may be that GIS-constructed neighbourhoods, such as buffers, provide better representations of the areas in which people move around and they may therefore constitute superior spatial units of enumeration for walkability. However, the benefit of delineating more realistic neighbourhoods could be limited by feasibility, and therefore accuracy, of measuring walkability within them. Unlike administrative areas within which measurement pertinent to walkability such as residential counts are taken, GIS constructed neighbourhoods are not specifically associated with relevant data and depend on aggregation of such administrative unit data, adding methodological complexity and introducing error. In the present study walkability scores of buffers were derived from the means of those of the OAs they contained. It was necessary to derive score from these smallest administrative areas as they were the only ones smaller than the buffers. Yet, as discussed, it was not possible to index OA-level walkability in accordance with the method specified as a result of the particularly small size of OAs relative to areas of land uses included in the land use mix dimension. The use of administrative areas as neighbourhoods for measurement of walkability inevitably presents the modifiable areal unit problem but if this is at least acknowledged, administrative areas such as CAS wards which are the closest administrative units in terms of size to 1600m walkable neighbourhoods – may constitute adequate spatial units of enumeration for walkability. Indeed, in the present study correlation between walkability scores of neighbourhoods defined as buffers and CAS wards was moderately high.

7.4.4 Spatial patterning of walkability across London

As evidenced in Map 7.8, showing the walkability of OAs, the hypothesis that there would be radial decay in walkability of London from the centre to the periphery reflecting reductions in residential density and junction density, two of the core walkability dimensions held true. However, the expectation that sequential addition to the land use mix L1 of recreational facilities in L2 and then greenspace in L3 would reduce the uniformity of the radial decay in walkability was not met. This expectation was based on the observation that the area of these additional land uses does not reduce uniformly from the centre and that spatial variation in greenspace area within Inner London is high. The fact that this trend was not apparent on the maps likely reflects the failure of the smaller OA administrative units to index walkability in terms of land use mix. Also, aggregation of walkability measures into larger administrative areas of CAS wards and LAs which tend to be larger in Outer London may impede the detection of subtle, fine scale changes in walkability.

7.4.5 Objective versus subjective walkability

Validation of the walkability index by subjective evaluation of images of OAs objectively assigned walkability scores was not possible from Google Street View photographs. These photographs did not appear sufficiently different to enable perceptual assessment of the relative walkabilities of OAs. A field "walkabout" may have constituted an effective means of comparing the walkability "feel" of an area to its objectively measured walkability. Aerial views of OAs varied in accordance with their objectively measured walkabilities but subjective assessments of walkability tend not to be made with a bird's eye view.

7.5 Summary

This study highlights a major practical challenge in the construction of a walkability index as the management of geographical data within separate geographical and software packages. Also, using GIS constructed neighbourhoods as enumeration units for walkability is difficult because, unlike administrative units, they are not specifically associated with relevant data.

Contrary to hypotheses, this study reveals that for any given unit of enumeration walkability model scores do not tend to vary dependent on the land uses included in the land use mix dimension. As predicted for the spatial patterning of walkability in London, as indicated by walkability scores, there is a radial decay from the centre to the periphery reflecting reductions in residential density and junction density. However, this pattern is not affected by land use mix specificity of the model: the pattern is the same for models including only the basic land uses as those including, for example, free recreational land.

8 Study 5: Associations between neighbourhood walkability and (A) physical activity, and (B) GHQ-12 depression

8.1 Introduction

8.1.1 Chapter overview

Walkability, the extent to which a place lends itself to walking and cycling as physically active forms of transport and recreation, was the neighbourhood environmental exposure focus of this chapter. This chapter presents results of statistical analyses designed to examine associations between models of neighbourhood walkability, the construction of which was described in Chapter 7, and physical activity and depression outcomes in the Whitehall II Phase 7 London study sub-sample (LWIIP7), and walkability and depression outcomes in the 2008 Health Survey for England London study sub-sample (LHSE2008),.

8.1.2 Research questions

This study addresses the following research questions:

- Is neighbourhood walkability associated with physical activity in the London contingent of the Whitehall II Study Phase 7 (LWIIP7) [182] study sample, and with a score of four or more on the 12-item General Health Questionnaire (GHQ-12 depression) in LWIIP7 and the London contingent of the 2008 Health Survey for England (LHSE2008) [183] study samples?
- How does the specification of the walkability model affect associations?
- How does operationalisation of neighbourhood affect associations?
- Are associations independent of individual-level sociodemographic factors and of area deprivation?

8.1.3 Objectives for Study 5A

The Study 5A objectives are as follows:

- to determine whether odds of physical activity are higher for an individual in the LWIIP7 study population living within a neighbourhood where walkability is higher
- to determine whether associations are affected by neighbourhood operationalisation, as various administrative units at different scales, as GIS software-computed gridded cells and as GIS software-computed walkable 1600m buffers
- to determine whether the association is affected by the model of walkability
- to determine whether the association is affected by individual-level sociodemographic factors and by area deprivation.

8.1.4 Hypotheses for Study 5A

Physical activity is associated with neighbourhood walkability, with individuals living in neighbourhoods where walkability is higher having higher odds of physical activity because walkability is supportive of walking and other physical activity.

The association between physical activity and walkability is stronger for neighbourhoods operationalised as smaller units, representing more walkable areas, than larger units, and stronger for those constructed as GIS software-computed walkable 1600m buffers than those operationalised as administrative units.

The associations between physical activity and walkability where the land use mix (LUM) dimension contains LUM1 will be weaker than those where it contains LUM2 which, in turn will be weaker than those where it contains LUM3. This is because LUM1 contains basic land uses considered supportive of physical activity as personal business destinations potentially reached by foot or bike – "Residential", "Retail", "Office" and "Health, welfare & community" – whereas LUM2 additionally contains "Entertainment, culture and recreation", more destinations to which to walk, and wherein to be physically active, be it in the form of walking or another for recreation as well as personal business. The additional inclusion of "Free recreational land" in LUM3 gives even more destinations and therefore models where LUM dimension contains LUM3 are expected to yield the strongest associations between physical activity and walkability.

The strength of associations is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and for area deprivation.

8.1.5 Objectives for Study 5B

The Study 5B objectives are as follows:

- to determine whether odds of depression are lower for an individual in the LWIIP7 study population living within a neighbourhood where walkability is higher
- to determine whether odds of depression are lower for an individual in the LHSE2008 study population living within a neighbourhood where walkability is higher
- to determine whether associations are affected by neighbourhood operationalisation, as various administrative units at different scales, as GIS software-computed gridded cells and as GIS software-computed walkable 1600m buffers
- to determine whether the association is affected by the model of walkability
- to determine whether the association is affected by individual-level sociodemographic factors and by area deprivation.

8.1.6 Hypotheses for Study 5B

Depression is associated with neighbourhood walkability, with individuals living in neighbourhoods where walkability is higher having lower odds of depression because walkability is supportive of walking and other physical activities which are protective of mental health.

The association between depression and walkability is stronger for neighbourhoods operationalized as smaller units, representing more walkable areas, and those constructed as GIS software-computed walkable 1600m buffers. The associations between depression and walkability where the land use mix (LUM) dimension contains LUM1 will be weaker than those where it contains LUM2 which, in turn will be weaker than those where it contains LUM3 for the reasons outlined for Study 5A.

The strength of associations is reduced but associations remain significant after adjusting for individual-level sociodemographic factors and for area deprivation.

8.2 Methods

8.2.1 Study samples

The study samples for Study 5 were drawn from the seventh wave, conducted in 2004/5, of the Whitehall II study, a longitudinal study of civil servants to examine the social determinants of health [182], and the 2008 Health Survey for England conducted in which drew a nationally representative general population sample of adults living in households in England through multi-stage stratified probability sampling [183] as detailed in Chapter 3. In Study 5A, the investigation of associations between walkability and physical activity, the study sample comprised only Whitehall II phase 7 (WIIP7) participants with a London postcode (LWIIP7) sample (see Section 3.2.1.2). In Study 5B, the investigation of associations between walkability and depression, however, the participants comprised not only LWIIP7 study sample but also a subset of HSE2008 resident in local authorities of the London region (LHSE2008) (see Section 3.2.2.2). There were 6885 participants in the whole WIIP7 study sample but the LWIIP7 study sample comprised 3020 individuals, thirty eight per cent of whom were female. The mean age of the LWIIP7 study sample was 61.0 years ± 6.0. Whilst there were 14,221 participants in the HSE2008 study sample, the LHSE2008 subset comprised 1540 individuals, fifty five per cent of whom were female. The mean age of this sample, at 45.6 years ± 18.4, was notably lower than that of the LWIIP7 sample.

8.2.2 Variables

8.2.2.1 Exposure variables

The exposure variables were various models of walkability measured within neighbourhood defined as particular units, details of which are provided in Chapter 7 (Section 7.2). Briefly, walkability was modelled as single dimensions of walkability, namely each land use mix (LUM 1, LUM2 and LUM3 as models L1, L2 and L3, respectively), residential dwelling density only (model R) and junction density only (model J), and as the full walkability models of M1 (comprising R, J and L1), M2 (comprising R, J and L2) and M3 (comprising R, J and L3). Preliminary statistical analysis was performed to determine the frequencies of walkability scores in the LWIIP7 (Appendix 8.1) and LHSE2008 (Appendix 8.2) samples and thereby whether dichotomisation of them was necessary. Within the LWIIP7 sample it was found that for all models measured within neighbourhoods defined as any of the administrative units – local authorities, CAS wards or Output Areas – the lower walkability quartile scores of 1 and 2 were over-represented as shown in Appendix 8.1. However, representation of scores 3 and 4 was substantive so dichotomisation was not considered necessary. Over representation of the lower walkability quartile scores was likely a consequence of over-representation of LWIIP7 participants in Outer London, as evidenced in Chapter 3 (Map 3.1), where walkability scores tended to be lower, as noted in Chapter 7 (Section 7.3.5).

As explained in Chapter 7 (Section 7.4.3) the GIS software-computed 200m by 200m grid cells (GC200) were too small for quantilation of land use mix and of residential dwelling density, rendering the construction of full walkability models impossible for this geography. Although full walkability models were not constructed for GC200s it was necessary to dichotomise the unevenly distributed walkability scores for GC200s for the LWIIP7 sample (Appendix 8.1), with scores of one and two classified as "low" and those of three and four as "high". All four quartile scores were approximately equally represented in the LWIIP7 study sample for all walkability models when neighbourhoods were defined as GIS software-computed 2000m by 2000m grid cells (GC2000). In contrast to operationalisation of neighbourhoods as administrative units or 2000m by 2000m cells, delineation of neighbourhoods as GIS software-computed 1600m circular and polygon based network buffers gave rise to high degrees of underrepresentation of the lowest quartile score – one, and of the highest –four – for all walkability models. This was a manifestation of their derivation as the means of the scores of the Output Areas they contained. It was therefore deemed necessary to dichotomise walkability quartile scores measured within the 1600m buffer neighbourhoods, as for GC200s. Neighbourhoods for the LHSE2008 sample were only defined as local authorities and it was found that for all models measured within these administrative units the lower walkability quartile scores of 1 and 2 were slightly over-represented as shown in Appendix 8.2 but to a lesser extent than for the LWIIP7 sample, a reflection of the nationally representative design of the LHSE2008. Thus, dichotomisation of walkability model scores was not necessary for the LHSE2008 sample.

8.2.2.2 Outcome variables

For Study 5A the outcome variables specified were as follows: meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity (WHOPA_E); meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast (WHOPA_I); and being in the top tertile of the WIIP7 sample for time spent walking per week (TTW). Chapter 3 (Section 3.3.3.2) provides the rationale for these outcome variable specifications. The WHO recommended physical activity level [47] was defined as follows:

- at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or at least 75 minutes of vigorous intensity aerobic physical activity throughout the week, or an equivalent combination of moderate and vigorous-intensity activity

GHQ-12 depression, defined as a score of four or more on the General Health Questionnaire-12 (GHQ-12), was the outcome specified for Study 5B for reasons given in Chapter 3 (Section 3.3.3.1). The GHQ is a survey tool devised to measure minor psychiatric morbidity (MPM), including anxiety and depression, in research settings [57].

8.2.2.3 Correlates

The individual level correlates entered into the regression models to examine relationships between walkability and physical activity outcomes in Study 5A were sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), car availability (Car available (ref)/No car available), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white). In Study 5B, the investigation of walkability in relation to the GHQ-12 outcome, adjustment was made for the individual-level sociodemographic factors of sex (Male (ref)/ Female), age (50y to <56y (ref)/ >=56y to <60y/ >=60y to <66y/ >=66y to 75y), economic activity (Remaining in Civil Service (ref)/ Working outside Civil Service/ Not working – retired/ Out of work/ Not working as long-term sick), marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) and ethnicity (White (ref)/Non-white) in the WIIP7 study sample, and for sex (Male (ref)/ Female), age (16y-34y (ref)/ 35y-54y/ 55y+), economic activity (Employed (ref)/ Unemployed/ Retired/ Other), income (<=£14918 (ref)/ >£14918-£31871/>£31871), occupational social class (Managerial and professional (ref)/ Intermediate/ Routine and manual/ Other) and marital status (Married (ref)/ Cohabit/ Single/ Divorced/ Widowed) in the HSE2008 study sample. In both studies adjustment was also made for a contextual factor, area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level [197]. Chapter 4 (Section 4.4) provides the justification for adjustment for these individual and area level factors.

8.2.3 Statistical analysis of associations between neighbourhood walkability and physical activity and GHQ-12 depression outcomes

For reasons given in Chapter 3 (Section 3.4) multivariate logistic regression was the method selected to model statistical associations in examination of possible relationships between walkability and the outcomes in the study samples. Bivariate logistic regression models were specified which included only the outcome variable and exposure variable of interest to determine the presence and strength of significant associations. Subsequently, for models in which significant associations were detected, adjustment was made for correlates – firstly individual level sociodemographic factors and then additionally area deprivation. This showed the extent to which the relationships between walkability exposures and outcomes were independent of other factors. Three models were constructed; Model 1, the unadjusted model was univariate analysis between the exposure and the outcome; Model 2 was multivariate analysis adjusting for all individual factors significantly associated with the outcome in the study sample; and Model 3 adjusted for area deprivation, in addition to the individual factors. Results were computed as odds ratios alongside their 95% confidence intervals to indicate with 95% confidence the range within which the true value of the odds ratio lay.

The reference category in each model for neighbourhoods operationalized as LAs, CAS wards and OAs was the lowest walkability quartile score, Quartile score 1, representing the lowest neighbourhood walkability and , for neighbourhoods operationalized as 1600m buffer neighbourhoods – CBs and PBNBs – and for GC200s the reference category was the lowest two walkability quartile scores, Quartile scores 1 and 2. Therefore, an odds ratio indicated the odds of the outcome in those exposed to, or living in a neighbourhood of the lowest walkability relative to those living in a neighbourhood of higher walkability. Tests for trend were performed to evaluate the association between the outcome, be it physical activity or GHQ-12 depression, and walkability variables with respect to the trend for a dose-response effect of the quartile score.

8.3 Results

8.3.1 Results for Study 5A: Associations between walkability and physical activity outcomes in LWIIP7

Statistical tests for trend were performed to evaluate overall patterns in the relationships between the walkability exposures and the physical activity outcomes in the LWIIP7 study sample before adjustment for individual level sociodemographic factors and area level deprivation (Table 8.1). For the majority of walkability models, and for the majority of neighbourhood operationalisations, tests for trend revealed significant dose-response effects of exposure to walkability on all of the physical activity outcomes in the LWIIP7 study sample of meeting the WHO physical activity recommendation excluding brisk walking as a contributory factor (WHOPA_E), meeting the WHO physical activity recommendation including brisk walking as a contributory factor (WHOPA_I), and being in the top tertile for time spent walking per week (TTW). The trend test z statistic for every statistically significant model for the WHOPA_E and WHOPA_I outcomes was negative, indicating that individuals living in neighbourhoods of greater walkability had lower odds of these outcomes than those in neighbourhoods of lower walkability. The reverse was found with respect to the TTW outcome, as indicated by the positive z statistics for all statistically significant models: participants living in the most walkable neighbourhoods had higher odds of being in the top tertile of the LWIIP7 study sample for time spent walking per week.

The full walkability models, M1, M2 and M3, comprising land use mix (LUM) 1, LUM2 and LUM 3, respectively, together with residential dwelling density (RDD) and junction density (JD) components, were consistently significantly associated with the WHOPA_E and WHOPA_I outcomes at all neighbourhood operationalisations to which they were applied, namely 1600m circular buffers (CBs), 1600m polygon based network buffers (PBNBs), local authorities (LAs), CAS wards, Output Areas

(OAs) and 2000m by 2000m grid cells (GC2000s). Of the full models, only M2 with neighbourhoods operationalised as CBs was not significantly associated with WHOPA_E. Similarly, the residential dwelling density only model (R) and the junction density only model (J) were consistently significantly associated with WHOPA_E and WHOPA_I as indicated by tests for trend. Only for the model J with neighbourhoods operationalised either as PBNBs and OAs were no significant associations with WHOPA_E found, and only for the model R with neighbourhoods operationalised as 200m by 200m grid cells (GC200s) was no significant association with WHOPA_I detected. However, the models comprising a land use mix only, namely L1, L2 and L3, were less consistently associated with the WHOPA_E and WHOPA _I outcomes. Of the three land use mix only models, the greatest number of statistically significant associations with WHOPA_E and WHOPA_I was found for L1, and only with neighbourhoods operationalised as PBNBs was a significant association found for neither outcome. With respect to walkability associations with the TTW outcome as indicated by tests for trend, fewer models were statistically significantly associated at any given neighbourhood operationalisation but as for the other physical activity outcomes, where found they tended to be for the full walkability models of M1, M2 and M3. It was notable that the neighbourhood operationalisations for which the most walkability models were significantly associated with all three outcomes, before adjustment for individual level sociodemographic factors and area deprivation, were CAS wards. When walkability was measured in neighbourhoods defined as CAS wards, the residential dwelling density only model (R), the junction density only model (J), and the full walkability models M1, M2 and M3 models were significantly associated with all three physical activity outcomes.

Whilst the results of trend analysis in Table 8.1 reveal significant associations between walkability and the WHOPA_E or WHOPA_I outcomes for most models, and for any neighbourhood delineation, adjustment for correlates reduced the number of, or entirely eliminated, significant associations. Appendix 8.3 and Appendix 8.4, respectively, provide summaries of the walkability models for which significant associations with WHOPA_E and WHOPA_I were found, as indicated by tests for trend, with each neighbourhood operationalisation before and after adjustment for individual level sociodemographic factors and area level deprivation. After adjustment for individual level factors, no models of walkability were significantly associated with WHOPA_E or WHOPA_I with neighbourhood operationalised as PBNBs. Also, after this adjustment, no walkability models were significantly associated with WHOPA_E with neighbourhoods defined as CBs and, with this neighbourhood operationalisation, only the junction density only model (J) was significantly associated with WHOPA_I. This significant association, however, was rendered non-significant after additional adjustment for area level deprivation. Some walkability models remained significantly associated with both the WHOPA_E and WHOPA_I outcomes for all three administrative unit neighbourhood operationalisations of LAs, CAS wards and OAs after adjustment for individual level factors. More models remained significantly associated for the smaller administrative unit neighbourhood operationalisations. For example, for the WHOPA_E outcome, only the residential dwelling density only model (R) and the junction density only model (J) were significantly associated with it, after adjustment for individual level factors, with neighbourhood defined as LAs. However, all three of the land use mix only models (L1, L2 and L3), the residential dwelling density only model (R), and all three of the full walkability models (M1, M2 and M3), were significantly associated with this outcome with neighbourhoods defined as OAs. After additional adjustment for area level deprivation, OAs proved the most robust neighbourhood operationalised with regards to significant associations remaining: the land use mix 2 and land use mix 3 only models (L2 and L3), the residential dwelling density only model (R), and all full walkability models (M1, M2 and M3) remained significantly associated with WHOPA_E whilst models R, M1, M2, M3 remained significantly associated with WHOPA_I.

Appendix 8.5 provides a summary of the walkability models for which significant associations with TTW were found, as indicated by tests for trend, with each neighbourhood operationalisation before and after adjustment for individual level sociodemographic factors and area level deprivation. It is evident that significant associations remained for walkability models, after adjustment, for more neighbourhood operationalisations for the TTW outcome than for the WHOPA_E and WHOPA_I outcomes. Significant associations remained with the TTW outcome after adjustment for individual level factors and area level deprivation for walkability models with neighbourhood operationalised as CBs, PBNBs, CAS wards, GC2000s and GC200s. The greatest number of walkability models remained significantly associated with TTW after this adjustment with neighbourhoods defined as CAS wards, namely the land use mix 1 only model (L1), the junction density only model (J) and all of the full walkability models (M1, M2 and M3).

As discussed in Chapter 7 (Section 7.4.3), modelling walkability within particularly small neighbourhoods, and in those that were not administratively defined, was flawed. In the former case, land use mix entropy scores could not be quantiled due to overrepresentation of neighbourhoods with entropies of zero: there was insufficient variation to make quantiling possible. In the latter case of non-administratively defined units, attribution of Output Area household counts from those whose centroids fell within the defined neighbourhood – for calculation of the residential dwelling density dimension of walkability – was not necessarily accurate: the centroid of the overlapping Output Area may not have "contained" the concentration of the Output Area's dwellings. Therefore, odds ratios for associations between walkability and physical activity outcomes before and after adjustment for correlates, together with tests for trend, are only presented for CAS wards, a neighbourhood operationalisation in which walkability was deemed adequately measurable, for the WHOPA_E, WHOPA_I and TTW outcomes in Table 8.2, Table 8.3 and Table 8.4, respectively. As shown in Table 8.2, before adjustment for individual level factors, the lowest odds of the WHOPA_E outcome was for the walkability model M2. Those living in the most walkable CAS ward-defined neighbourhoods as indicated by a model M2 walkability quartile score of 4 had lower odds than those in the least walkable CAS ward-defined neighbourhoods as indicated by a score of 1 (OR= 0.57, $p<0.001$). This significant association remained, albeit weaker, after adjustment for individual level sociodemographic factors (OR=0.66, $p=0.01$). However, after additional adjustment for area level deprivation this association was no longer statistically significant at the 95% confidence level (OR=0.73, $p=0.077$). Also, no other models of walkability remained significantly associated with WHOPA_E with the CAS ward neighbourhood operationalisation after additional adjustment for area level deprivation. As for the WHOPA_E outcome, significant associations between walkability and WHOPA_I, with neighbourhoods operationalised as CAS wards were found with many walkability models before adjustment for correlates (Table 8.3). The lowest odds of the WHOPA_I outcome before adjustment was for the junction density only model (J): those living in the most walkable neighbourhoods as indicated by a model J walkability quartile score of 4 had lower odds than those in the least walkable neighbourhoods (OR=0.61, p<0.001). This significant association remained, albeit weaker, after adjustment for individual level sociodemographic factors (OR=0.72, p<0.01) and additional adjustment for area level deprivation (OR=0.76, p<0.05). However, the test for trend statistic was not significant after this additional adjustment for area level deprivation. Aside from that for the junction density only model (J), and in line with the results for the WHOPA_E outcome, there was no evidence of significant associations of any other models of walkability with WHOPA_I after additional adjustment for area level deprivation.

In contrast to the findings for the World Health Organisation physical activity recommendation outcomes of WHOPA_E and WHOPA_I, several walkability models remained significantly associated with being in the top tertile of the LWIIP7 study sample for time spent walking per week (TTW) after adjustment for individual level factors and area deprivation (Table 8.4). Before adjustment for individual level factors, the highest odds of the TTW outcome was for the full walkability model 1 (M1). Those living in the most walkable CAS ward-defined neighbourhoods as indicated by a model M1 walkability quartile score of 4 had higher odds of more time spent walking than those in the least walkable CAS ward-defined neighbourhoods as indicated by a score of 1 (OR= 1.51, p<0.001). After adjustment for individual level sociodemographic factors the odds ratio was 1.39 (p<0.05) and, after additional adjustment for area level deprivation, it was 1.42 (p<0.05).

Table 8.1 Significance of trends in associations between walkability models and the physical activity outcomes of WHOPA_E, WHOPA_I and TTW in the LWIIP7 study sample.

Outcome		WHOPA_E[1]		WHOPA_I[2]		TTW[3]	
Neighbourhood delineation	Model	z	p	z	p	z	p
	L1	-1.96	0.050	-2.91	0.004	1.78	0.074
	L2	-1.26	0.208	-1.72	0.086	1.74	0.082
	L3	-1.55	0.122	-2.30	0.022	-0.07	0.947
	R	-2.40	0.016	-2.18	0.029	2.53	0.012
	J	-2.56	0.010	-3.47	0.001	2.36	0.018
	M1	-1.99	0.047	-2.42	0.016	3.01	0.003
	M2	-2.06	0.039	-2.30	0.022	3.10	0.002
1600m circular buffer	M3	-2.33	0.020	-2.59	0.010	2.66	0.008
	L1	-0.62	0.538	-0.59	0.555	0.99	0.324
	L2	-0.38	0.703	0.66	0.509	1.42	0.156
	L3	-0.37	0.709	-0.79	0.430	-0.05	0.961
	R	-2.56	0.010	-2.27	0.023	2.94	0.003
1600m polygon based network buffer	J	-1.43	0.152	-2.69	0.007	2.69	0.007

Outcome		WHOPA_E[1]		WHOPA_I[2]		TTW[3]	
Neighbourhood delineation	Model	z	p	z	p	z	p
	M1	-1.88	0.061	-2.15	0.031	2.88	0.004
	M2	-2.09	0.037	-2.26	0.024	3.37	0.001
	M3	-2.81	0.005	-2.71	0.007	3.70	<0.001
Local authority	L1	-1.30	0.195	-2.74	0.006	1.40	0.163
	L2	-0.54	0.589	-1.95	0.051	1.23	0.217
	L3	-0.10	0.918	-2.41	0.016	1.10	0.272
	R	-4.17	<0.001	-3.58	<0.001	0.74	0.461
	J	-3.19	0.001	-3.55	<0.001	1.65	0.099
	M1	-2.74	0.006	-3.63	<0.001	1.71	0.088
	M2	-2.66	0.008	-3.48	0.001	2.02	0.043
	M3	-2.35	0.019	-3.63	<0.001	1.84	0.065
CAS ward	L1	-2.41	0.016	-2.96	0.003	3.04	0.002
	L2	-1.95	0.051	-2.25	0.024	0.20	0.838
	L3	-1.36	0.175	-2.43	0.015	1.03	0.302
	R	-4.03	<0.001	-2.95	0.003	2.14	0.033
	J	-3.41	0.001	-3.98	<0.001	2.89	0.004
	M1	-3.57	<0.001	-3.62	<0.001	3.74	<0.001
	M2	-3.66	<0.001	-3.87	<0.001	3.17	0.002
	M3	-4.12	<0.001	-4.15	<0.001	3.18	0.001
Output Area	L1	-2.59	0.010	-2.97	0.003	-0.77	0.440
	L2	-2.66	0.008	-2.51	0.012	-0.43	0.664
	L3	-2.68	0.007	-2.37	0.018	-1.02	0.306
	R	-5.27	<0.001	-5.35	<0.001	3.22	0.001
	J	-1.30	0.192	-3.05	0.002	2.11	0.034
	M1	-4.52	<0.001	-5.14	<0.001	2.08	0.038
	M2	-4.38	<0.001	-4.92	<0.001	1.72	0.085
	M3	-4.82	<0.001	-4.97	<0.001	2.22	0.026
2000m by 2000m grid cell (GC2000)	L1	-2.30	0.022	-2.81	0.005	2.66	0.008
	L2	-2.07	0.039	-1.97	0.049	0.80	0.426
	L3	-0.38	0.702	-1.58	0.114	0.01	0.996
	R	-4.30	<0.001	-3.59	<0.001	2.23	0.026
	J	-3.28	0.001	-3.27	0.001	2.14	0.032
	M1	-3.56	<0.001	-3.99	<0.001	2.76	0.006
	M2	-3.67	<0.001	-3.81	<0.001	2.26	0.024
	M3	-3.26	0.001	-3.48	0.001	2.12	0.034
200m by 200m grid cell (GC200)	L1	-2.27	0.023	-2.26	0.024	0.97	0.330
	L2	-2.18	0.029	-2.01	0.044	-0.05	0.957
	L3	-3.14	0.002	-2.22	0.027	0.79	0.427
	R	-2.69	0.007	-1.91	0.056	2.84	0.005
	J	-4.20	<0.001	-3.87	<0.001	1.27	0.204

[1]Meeting the World Health Organisation (WHO) recommended physical activity level excluding any walking as a contributory physical activity ;[2]Meeting the WHO recommended physical activity level including walking for those who reported their normal pace was brisk or fast; [3]Being in the top tertile for

Outcome		WHOPA_E[1]		WHOPA_I[2]		TTW[3]	
Neighbourhood delineation	Model	z	p	z	p	z	p
time spent walking per week							

Table 8.2 Association between various models of walkability at CAS ward level and meeting the WHOPA_E in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
			2918			2889			2889	
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
L1	1	1	REF		1	REF		1	REF	
L1	2	0.95	0.77-1.18	0.663	0.98	0.79-1.23	0.908	1	0.80-1.25	0.986
L1	3	0.90	0.71-1.15	0.408	1	0.78-1.29	0.98	1.05	0.81-1.37	0.692
L1	4	0.71	0.55-0.93	0.011	0.79	0.60-1.04	0.089	0.85	0.64-1.14	0.275
		Test for trend p < 0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L2	1	1	REF		1	REF		1	REF	
L2	2	0.90	0.72-1.12	0.339	0.88	0.70-1.11	0.296	0.88	0.70-1.10	0.255
L2	3	0.86	0.68-1.09	0.227	0.87	0.68-1.10	0.25	0.87	0.69-1.11	0.276
L2	4	0.78	0.60-1.01	0.055	0.8	0.62-1.05	0.105	0.83	0.64-1.09	0.181
		Test for trend p = 0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L3	1	1	REF		1	REF		1	REF	
L3	2	0.91	0.72-1.14	0.403	0.88	0.70-1.12	0.31	0.88	0.70-1.12	0.314
L3	3	0.95	0.75-1.19	0.647	0.91	0.72-1.16	0.452	0.91	0.72-1.15	0.438
L3	4	0.81	0.63-1.05	0.11	0.81	0.63-1.06	0.126	0.85	0.65-1.11	0.226
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
R	1	1	REF		1	REF		1	REF	
R	2	0.83	0.67-1.04	0.099	0.91	0.72-1.14	0.393	0.91	0.72-1.14	0.408
R	3	0.73	0.58-0.92	0.008	0.85	0.66-1.08	0.187	0.88	0.68-1.14	0.319
R	4	0.59	0.45-0.78	<0.001	0.72	0.54-0.95	0.022	0.79	0.57-1.09	0.148
		Test for trend p < 0.001			Test for trend p < 0.05			Tr tst non-sig (p>0.05)		
J	1	1	REF		1	REF		1	REF	
J	2	0.73	0.59-0.92	0.006	0.8	0.64-1.00	0.055	0.81	0.64-1.02	0.076
J	3	0.74	0.59-0.93	0.011	0.88	0.69-1.12	0.3	0.91	0.71-1.18	0.486
J	4	0.65	0.49-0.84	0.001	0.77	0.58-1.02	0.068	0.83	0.61-1.12	0.218
		Test for trend p = 0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M1	1	1	REF		1	REF		1	REF	

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
		2918			2889			2889		
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
M1	2	0.86	0.69-1.08	0.192	0.95	0.76-1.20	0.692	0.99	0.78-1.25	0.915
M1	3	0.81	0.64-1.02	0.07	0.96	0.75-1.22	0.731	1.02	0.79-1.32	0.892
M1	4	0.60	0.46-0.80	<0.001	0.74	0.55-0.99	0.041	0.82	0.59-1.14	0.233
		Test for trend p < 0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M2	1	1	REF		1	REF		1	REF	
M2	2	0.86	0.69-1.07	0.179	0.91	0.72-1.14	0.407	0.93	0.73-1.17	0.528
M2	3	0.81	0.65-1.01	0.065	0.99	0.78-1.26	0.95	1.05	0.81-1.35	0.717
M2	4	0.57	0.42-0.77	<0.001	0.66	0.48-0.91	0.01	0.73	0.52-1.03	0.077
		Test for trend p < 0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M3	1	1	REF		1	REF		1	REF	
M3	2	0.84	0.68-1.04	0.113	0.92	0.74-1.14	0.447	0.95	0.75-1.19	0.638
M3	3	0.68	0.53-0.88	0.003	0.82	0.63-1.07	0.143	0.87	0.66-1.15	0.319
M3	4	0.61	0.47-0.80	<0.001	0.73	0.55-0.97	0.028	0.79	0.58-1.09	0.155
		Test for trend p < 0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
		[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

Table 8.3 Association between various models of walkability at CAS ward level and WHOPA_I in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation.

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
		2918			2889			2889		
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
L1	1	1	REF		1	REF		1	REF	
L1	2	0.84	0.69-1.01	0.061	0.86	0.71-1.04	0.121	0.86	0.71-1.05	0.143
L1	3	0.82	0.66-1.01	0.060	0.89	0.71-1.11	0.302	0.93	0.74-1.17	0.560
L1	4	0.73	0.58-0.91	0.005	0.84	0.67-1.06	0.147	0.91	0.71-1.16	0.441
		Test for trend p <0.01			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L2	1	1	REF		1	REF		1	REF	
L2	2	0.97	0.80-1.17	0.737	0.97	0.79-1.19	0.78	0.96	0.79-1.18	0.711

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
			2918			2889			2889	
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
L2	3	0.85	0.69-1.04	0.122	0.84	0.68-1.04	0.119	0.85	0.69-1.05	0.128
L2	4	0.80	0.64-1.00	0.049	0.87	0.69-1.09	0.216	0.90	0.71-1.13	0.371
		Test for trend p <0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L3	1	1	REF		1	REF		1	REF	
L3	2	0.89	0.73-1.09	0.265	0.90	0.73-1.10	0.302	0.90	0.73-1.10	0.302
L3	3	0.85	0.69-1.03	0.103	0.81	0.66-1.00	0.045	0.81	0.66-0.99	0.041
L3	4	0.77	0.62-0.96	0.019	0.83	0.66-1.04	0.105	0.86	0.68-1.08	0.203
		Test for trend p <0.05			Test for trend p <0.05			Test for trend p <0.05		
R	1	1	REF		1	REF		1	REF	
R	2	0.89	0.73-1.08	0.231	0.97	0.79-1.18	0.738	0.98	0.80-1.20	0.810
R	3	0.72	0.59-0.88	0.002	0.82	0.66-1.02	0.074	0.88	0.70-1.10	0.259
R	4	0.78	0.62-0.98	0.032	0.95	0.74-1.21	0.655	1.10	0.84-1.45	0.486
		Test for trend p <0.01			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
J	1	1	REF		1	REF		1	REF	
J	2	0.78	0.65-0.95	0.013	0.84	0.69-1.03	0.097	0.86	0.70-1.05	0.136
J	3	0.81	0.67-0.99	0.044	0.93	0.75-1.15	0.500	0.96	0.77-1.20	0.745
J	4	0.61	0.49-0.77	<0.001	0.72	0.56-0.92	0.008	0.76	0.59-0.99	0.043
		Test for trend p <0.001			Test for trend p <0.05			Tr tst non-sig (p>0.05)		
M1	1	1	REF		1	REF		1	REF	
M1	2	0.79	0.66-0.96	0.019	0.89	0.73-1.09	0.244	0.92	0.74-1.13	0.431
M1	3	0.82	0.67-0.99	0.044	0.95	0.77-1.17	0.617	1.01	0.81-1.27	0.898
M1	4	0.65	0.52-0.82	<0.001	0.79	0.62-1.02	0.068	0.89	0.68-1.18	0.428
		Test for trend p <0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M2	1	1	REF		1	REF		1	REF	
M2	2	0.92	0.76-1.11	0.379	0.96	0.79-1.17	0.685	0.98	0.80-1.21	0.872
M2	3	0.76	0.62-0.92	0.005	0.91	0.74-1.13	0.392	0.97	0.78-1.21	0.794
M2	4	0.66	0.52-0.84	0.001	0.77	0.60-1.00	0.053	0.87	0.65-1.16	0.337
		Test for trend p <0.001			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M3	1	1	REF		1	REF		1	REF	
M3	2	0.88	0.73-1.06	0.179	0.92	0.76-1.12	0.405	0.95	0.78-1.17	0.640
M3	3	0.69	0.56-0.85	0.001	0.83	0.66-1.04	0.097	0.88	0.69-1.12	0.304
M3	4	0.68	0.54-0.85	0.001	0.80	0.63-1.01	0.064	0.88	0.67-1.15	0.358
		Test for trend p <0.001			Test for trend p <0.05			Tr tst non-sig (p>0.05)		
		[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

Table 8.4 Association between various models of walkability at CAS ward level and TTW in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation.

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n		2756			2736			2736		
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
L1	1	1	REF		1	REF		1	REF	
	2	1.16	0.95-1.42	0.155	1.13	0.92-1.39	0.255	1.14	0.92-1.41	0.235
	3	1.23	0.98-1.54	0.074	1.17	0.92-1.47	0.197	1.16	0.91-1.48	0.223
	4	1.42	1.12-1.78	0.003	1.34	1.05-1.71	0.017	1.34	1.04-1.72	0.025
		Test for trend p <0.01			Test for trend p <0.05			Test for trend p <0.05		
L2	1	1	REF		1	REF		1	REF	
	2	0.87	0.70-1.08	0.198	0.86	0.69-1.07	0.176	0.86	0.69-1.06	0.155
	3	0.82	0.66-1.02	0.076	0.79	0.63-1.00	0.045	0.79	0.63-0.99	0.042
	4	1.09	0.87-1.38	0.450	1.03	0.82-1.31	0.778	1.01	0.80-1.29	0.916
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L3	1	1	REF		1	REF		1	REF	
	2	1.02	0.82-1.26	0.889	1.02	0.81-1.27	0.885	1.01	0.81-1.26	0.927
	3	0.96	0.78-1.20	0.736	0.93	0.75-1.17	0.548	0.93	0.75-1.16	0.543
	4	1.18	0.94-1.49	0.158	1.15	0.90-1.46	0.257	1.14	0.89-1.45	0.299
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
R	1	1	REF		1	REF		1	REF	
	2	0.96	0.78-1.19	0.727	0.97	0.78-1.20	0.773	0.96	0.77-1.20	0.738
	3	1.08	0.87-1.34	0.501	1.04	0.83-1.30	0.758	1.03	0.81-1.31	0.835
	4	1.29	1.02-1.64	0.034	1.16	0.90-1.49	0.262	1.13	0.85-1.50	0.411
		Test for trend p <0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
J	1	1	REF		1	REF		1	REF	
	2	1.13	0.91-1.39	0.264	1.11	0.89-1.37	0.348	1.11	0.89-1.39	0.350
	3	1.27	1.02-1.57	0.033	1.27	1.01-1.60	0.037	1.28	1.01-1.62	0.040
	4	1.37	1.08-1.74	0.009	1.26	0.98-1.63	0.071	1.24	0.94-1.63	0.124
		Test for trend p <0.01			Test for trend p <0.05			Test for trend p <0.05		
M1	1	1	REF		1	REF		1	REF	
	2	1.04	0.84-1.28	0.729	1.02	0.82-1.27	0.882	1.03	0.82-1.30	0.782
	3	1.29	1.04-1.60	0.018	1.24	0.99-1.56	0.057	1.27	1.00-1.61	0.055
	4	1.51	1.19-1.91	0.001	1.39	1.08-1.79	0.011	1.42	1.07-1.89	0.015
		Test for trend p <0.001			Test for trend p <0.01			Test for trend p <0.01		

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n			2756			2736			2736	
Model	Quart score	OR	CI	p	OR	CI	p	OR	CI	p
M2	1	1	REF		1	REF		1	REF	
	2	1.18	0.96-1.45	0.119	1.15	0.93-1.43	0.188	1.17	0.93-1.45	0.175
	3	1.25	1.01-1.54	0.039	1.19	0.95-1.49	0.123	1.21	0.95-1.53	0.121
	4	1.47	1.14-1.88	0.003	1.33	1.02-1.73	0.035	1.33	0.99-1.80	0.058
		Test for trend p <0.01			Test for trend p <0.05			Test for trend p <0.05		
M3	1	1	REF		1	REF		1	REF	
	2	1.18	0.96-1.44	0.115	1.15	0.94-1.42	0.176	1.18	0.95-1.47	0.140
	3	1.18	0.94-1.48	0.146	1.11	0.88-1.42	0.379	1.13	0.88-1.46	0.339
	4	1.49	1.18-1.88	0.001	1.40	1.09-1.79	0.009	1.42	1.07-1.87	0.014
		Test for trend p <0.01			Test for trend p <0.05			Test for trend p <0.05		
		[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level								

8.3.2 Results for Study 5B: Associations between walkability and GHQ-12 depression in LWIIP7 and in LHSE2008

For the LWIIP7 study sample, statistical tests for trend were performed to evaluate overall patterns in the relationships between the walkability exposures and the GHQ-12 depression outcome before adjustment for individual level sociodemographic factors and area level deprivation, as shown in Table 8.5. Additionally, Appendix 8.6 provides a summary of the walkability models for which significant associations with GHQ-12 depression were found in the LWIIP7 study sample, as indicated by tests for trend, with each neighbourhood operationalisation before and after adjustment for individual level sociodemographic factors and area level deprivation.

Table 8.5 shows that, before adjustment for individual level sociodemographic factors, at least one walkability model for each neighbourhood operationalisation had a significant dose-response association with the GHQ-12 depression outcome in the LWIIP7 study sample, as indicated by tests for trend. Also, as evidenced by the positive trend test z statistics, all significant associations of walkability with GHQ-12 depression were positive: individuals living in more walkable neighbourhoods had higher odds of this outcome. For the LWIIP7 study sample, neighbourhood operationalisations for which the greatest number of walkability models was significantly associated with the GHQ-12 depression outcome before adjustment were LAs, with only one model, the land use mix 3 only model (L3), not significantly associated with the outcome. Table 8.5 also shows that across the neighbourhood operationalisations, the models most consistently associated with GHQ-12 depression in the LWIIP7 study sample before adjustment for individual level factors were residential dwelling density only model (R) and the full walkability model 3 (M3). However, Appendix 8.6 shows that after adjustment for these factors, all significant associations as indicated by tests for trend, for all neighbourhood operationalisations except LAs, were rendered not significant. For neighbourhoods defined as LAs only the land use mix 2 only model (L2) remained statistically significantly positively associated with GHQ-12 depression, and after adjustment for area deprivation the significant association was eliminated.

Odds ratios for associations between walkability and the GHQ-12 depression outcome in the LWIIP7 study sample before and after adjustment for correlates, together with tests for trend, are presented only for CAS wards (Table 8.6). Reasons for the restriction of presentation of results to this neighbourhood operationalisation are the same as those for the presentation of results for associations of walkability with physical activity outcomes, and discussed in Chapter 7 in Section 7.4.3. For the LHSE2008 study sample, odds ratios for associations between walkability and the GHQ-12 depression outcome before and after adjustment for correlates, together with tests for trend, are presented for LAs, the only neighbourhood operationalisation for these individuals (Table 8.7). Table 8.6 shows that the before adjustment for individual level factors, the highest odds of the GHQ-12 depression outcome in the LWIIP7 study sample was for the junction density only model (J). Among these individuals, those living in the most walkable CAS ward-defined neighbourhoods as indicated by a model J walkability quartile score of 4 had higher odds than those in the least walkable CAS ward-defined neighbourhoods as indicated by a score of 1 for this model (OR= 1.55, $p<0.01$). Odds were also significantly higher for those living in the most walkable CAS ward-defined neighbourhoods as indicated by a residential dwelling density only model (R) walkability quartile score of 4 relative to those in the least walkable CAS ward-defined neighbourhoods as indicated by a model R walkability score of 1 (OR= 1.52, $p<0.01$). As for the LWIIP7 study sample with neighbourhoods operationalised as CAS wards, after adjustment for individual level factors no significant associations were found for the LHSE2008 study sample with neighbourhoods defined as LAs. In fact, Table 8.7 shows that even before this adjustment, only the R model of walkability was significantly associated with GHQ-12 depression in this study sample. As for the LWIIP7 study sample, walkability was positively associated with GHQ-12 depression in the LHSE2008 study sample. Those living in the most walkable LA-defined neighbourhoods as indicated by a model R walkability quartile score of 4 had higher odds than those in the least walkable neighbourhoods as indicated by a score of 1 for this model (OR= 1.50, $p<0.05$). However, the non-significant trend test statistic for this model indicated there was no significant dose effect.

Table 8.5 Significance of trends in associations between walkability models and the GHQ-12 depression outcome in the LWIIP7 study sample.

Outcome		GHQ-12 depression	
Neighbourhood delineation	Model	z	p
	L1	0.83	0.409
	L2	0.00	0.999
	L3	0.23	0.820
1600m circular buffer	R	2.52	0.012

Outcome		GHQ-12 depression	
Neighbourhood delineation	Model	z	p
	J	1.61	0.108
	M1	1.40	0.162
	M2	1.44	0.150
	M3	2.18	0.030
	L1	0.61	0.544
	L2	0.12	0.901
	L3	-0.80	0.425
	R	2.40	0.016
	J	0.19	0.847
	M1	0.96	0.336
	M2	1.83	0.068
1600m polygon based network buffer	M3	2.56	0.010
	L1	2.81	0.005
	L2	2.76	0.006
	L3	1.84	0.065
	R	2.56	0.011
	J	1.96	0.050
	M1	2.59	0.010
	M2	2.87	0.004
Local authority	M3	2.40	0.016
	L1	1.26	0.208
	L2	1.17	0.242
	L3	0.55	0.582
	R	2.58	0.010
	J	2.58	0.010
	M1	2.25	0.025
	M2	2.54	0.011
CAS ward	M3	2.56	0.010
	L1	0.91	0.363
	L2	0.89	0.376
	L3	0.31	0.754
	R	3.33	0.001
	J	0.70	0.483
	M1	3.16	0.002
	M2	3.32	0.001
Output Area	M3	2.40	0.016
	L1	2.45	0.014
	L2	1.44	0.150
	L3	0.79	0.429
	R	2.46	0.014
	J	2.03	0.042
2000m by 2000m grid cell (GC2000)	M1	2.82	0.005

Outcome		GHQ-12 depression	
Neighbourhood delineation	Model	z	p
	M2	3.14	0.002
	M3	2.93	0.003
200m by 200m grid cell (GC200)	L1	1.21	0.225
	L2	2.07	0.039
	L3	1.59	0.111
	R	1.65	0.099
	J	0.57	0.570

Table 8.6 Association between various models of walkability at CAS ward level and GHQ-12 depression in the LWIIP7 study sample before and after adjustment for individual factors and area level deprivation.

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n		2888			2878			2878		
Model	Quartile score	OR	CI	p	OR	CI	p	OR	CI	p
L1	1	1	REF		1	REF		1	REF	
	2	0.94	0.72-1.22	0.645	0.90	0.69-1.18	0.459	0.90	0.69-1.19	0.473
	3	1.12	0.85-1.50	0.419	1.01	0.75-1.35	0.971	0.99	0.73-1.34	0.965
	4	1.17	0.87-1.56	0.291	1.00	0.74-1.35	0.995	0.96	0.70-1.32	0.813
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L2	1	1	REF		1	REF		1	REF	
	2	0.95	0.72-1.25	0.705	0.92	0.70-1.22	0.569	0.94	0.71-1.24	0.658
	3	1.34	1.02-1.76	0.035	1.31	1.00-1.73	0.053	1.31	0.99-1.73	0.055
	4	1.04	0.77-1.41	0.805	0.96	0.70-1.31	0.782	0.94	0.69-1.29	0.724
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L3	1	1	REF		1	REF		1	REF	
	2	1.20	0.91-1.59	0.196	1.20	0.90-1.59	0.213	1.20	0.91-1.60	0.198
	3	1.18	0.89-1.55	0.250	1.17	0.88-1.54	0.273	1.17	0.88-1.54	0.279
	4	1.07	0.79-1.45	0.665	1.02	0.75-1.39	0.904	1.00	0.73-1.37	0.999
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
R	1	1	REF		1	REF		1	REF	
	2	1.10	0.83-1.45	0.506	0.98	0.74-1.30	0.905	1.00	0.76-1.34	0.975
	3	1.12	0.84-1.49	0.428	0.96	0.71-1.28	0.760	0.98	0.71-1.34	0.883
	4	1.52	1.13-2.05	0.005	1.28	0.94-1.74	0.121	1.28	0.90-1.81	0.173
		Test for trend p =0.01			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
J	1	1	REF		1	REF		1	REF	
	2	1.21	0.92-1.58	0.167	1.10	0.84-1.45	0.485	1.13	0.85-1.50	0.416
	3	1.15	0.86-1.53	0.341	0.97	0.72-1.30	0.844	0.97	0.72-1.33	0.868
	4	1.55	1.15-2.09	0.004	1.30	0.96-1.77	0.094	1.32	0.94-1.84	0.111

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n			2888			2878			2878	
Model	Quartile score	OR	CI	p	OR	CI	p	OR	CI	p
		Test for trend p =0.01			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M1	1	1	REF		1	REF		1	REF	
	2	0.93	0.70-1.22	0.581	0.79	0.60-1.05	0.110	0.79	0.58-1.06	0.116
	3	1.15	0.88-1.51	0.300	0.94	0.71-1.25	0.666	0.92	0.68-1.25	0.591
	4	1.38	1.03-1.85	0.030	1.13	0.83-1.54	0.427	1.08	0.76-1.53	0.676
		Test for trend p <0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M2	1	1	REF		1	REF		1	REF	
	2	0.84	0.64-1.11	0.229	0.75	0.57-1.00	0.051	0.75	0.56-1.01	0.061
	3	1.17	0.90-1.53	0.232	0.93	0.70-1.22	0.586	0.91	0.68-1.23	0.547
	4	1.45	1.07-1.96	0.016	1.23	0.90-1.69	0.193	1.20	0.83-1.72	0.329
		Test for trend p <0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M3	1	1	REF		1	REF		1	REF	
	2	1.10	0.84-1.43	0.479	0.97	0.74-1.27	0.823	0.97	0.73-1.28	0.808
	3	1.34	1.01-1.78	0.041	1.08	0.80-1.45	0.608	1.07	0.78-1.48	0.662
	4	1.39	1.03-1.87	0.030	1.16	0.85-1.57	0.360	1.12	0.78-1.59	0.539
		Test for trend p <0.05			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		

[a]Adjustment for sex, age, economic activity, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

Table 8.7 Association between various models of walkability at local authority level and having depression in the LHSE2008 study sample as indicated by a GHQ-12 score of four or more, before and after adjustment for individual factors and area level deprivation.

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n			1540			1150			1150	
Model	Quartile score	OR	CI	p	OR	CI	p	OR	CI	p
L1	1	1	REF		1	REF		1	REF	
L1	2	1.03	0.72-1.48	0.851	0.94	0.61-1.44	0.761	0.89	0.57-1.39	0.610
L1	3	1.11	0.77-1.61	0.567	0.87	0.55-1.38	0.555	0.79	0.49-1.29	0.353
L1	4	1.41	0.93-2.14	0.105	1.21	0.72-2.04	0.480	1.08	0.62-1.88	0.798
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L2	1	1	REF		1	REF		1	REF	
L2	2	1.33	0.94-1.89	0.106	1.14	0.74-1.74	0.554	1.16	0.75-1.79	0.493
L2	3	1.13	0.77-1.64	0.541	0.97	0.61-1.54	0.881	0.95	0.60-1.52	0.841
L2	4	1.41	0.92-2.16	0.118	1.35	0.80-2.28	0.268	1.31	0.76-2.25	0.327

		No adjustment			Adjustment for individual-level factors[a]			Adjustment for individual-level factors & area deprivation[b]		
n		1540			1150			1150		
Model	Quartile score	OR	CI	p	OR	CI	p	OR	CI	p
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
L3	1	1	REF		1	REF		1	REF	
L3	2	1.15	0.79-1.67	0.474	0.99	0.63-1.57	0.972	0.99	0.63-1.58	0.982
L3	3	1.33	0.90-1.96	0.149	1.08	0.68-1.72	0.746	1.02	0.64-1.64	0.926
L3	4	1.20	0.78-1.84	0.400	1.13	0.67-1.90	0.652	1.06	0.62-1.81	0.832
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
R	1	1	REF		1	REF		1	REF	
R	2	1.32	0.92-1.89	0.138	1.28	0.82-2.00	0.273	1.26	0.80-1.97	0.313
R	3	0.95	0.65-1.41	0.813	0.96	0.60-1.56	0.879	0.87	0.52-1.44	0.582
R	4	1.50	1.00-2.24	0.047	1.20	0.72-2.00	0.476	1.04	0.60-1.82	0.885
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
J	1	1	REF		1	REF		1	REF	
J	2	1.21	0.85-1.71	0.286	1.17	0.76-1.81	0.467	1.09	0.69-1.71	0.707
J	3	1.00	0.68-1.48	0.988	1.08	0.67-1.74	0.752	1.00	0.61-1.64	0.995
J	4	1.26	0.84-1.89	0.265	1.13	0.68-1.87	0.634	0.99	0.57-1.73	0.973
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M1	1	1	REF		1	REF		1	REF	
M1	2	1.18	0.84-1.65	0.350	1.09	0.71-1.68	0.678	0.95	0.60-1.50	0.817
M1	3	0.77	0.50-1.17	0.219	0.79	0.48-1.28	0.336	0.66	0.38-1.14	0.135
M1	4	1.40	0.95-2.07	0.088	1.17	0.71-1.92	0.536	0.95	0.54-1.67	0.858
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M2	1	1	REF		1	REF		1	REF	
M2	2	1.08	0.77-1.50	0.661	0.99	0.66-1.50	0.972	0.85	0.53-1.36	0.493
M2	3	0.83	0.52-1.33	0.446	0.87	0.50-1.51	0.631	0.79	0.45-1.40	0.426
M2	4	1.40	0.95-2.07	0.088	1.18	0.72-1.93	0.522	0.99	0.56-1.73	0.964
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		
M3	1	1	REF		1	REF		1	REF	
M3	2	1.11	0.78-1.60	0.558	0.95	0.60-1.51	0.841	0.82	0.49-1.35	0.434
M3	3	0.90	0.61-1.31	0.576	0.96	0.61-1.49	0.845	0.84	0.52-1.37	0.481
M3	4	1.40	0.95-2.07	0.088	1.17	0.72-1.92	0.527	0.98	0.56-1.72	0.949
		Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)			Tr tst non-sig (p>0.05)		

[a]Adjustment for sex, age, economic activity, income, occupational social class and marital status; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level

8.4 Discussion

The aim of this chapter was to examine associations between neighbourhood walkability and physical activity outcomes in LWIIP7, and between neighbourhood walkability and depression in LWIIP7 and LHSE2008, and to investigate how operationalisation of neighbourhood and modelling of walkability affected associations. This study also sought to determine whether associations were independent of individual-level sociodemographic factors and of area deprivation.

In Study 5A physical activity outcomes were specified as follows: meeting the WHO physical activity recommendation excluding brisk walking as a contributory factor (WHOPA_E); meeting the WHO recommendation including brisk walking (WHOPA_I); and being in the top tertile in LWIIP7 for time spent walking per week (TTW). It was hypothesised, in Study 5A, that all physical activity outcomes would be associated with neighbourhood walkability, with individuals living in neighbourhoods where walkability was higher having higher odds of the specified physical activity outcomes because walkability is supportive of walking and other physical activities. The association between the physical activity outcomes and walkability was expected to be significantly stronger for neighbourhoods operationalized as smaller units, representing walkable areas, and those constructed as GIS software-computed walkable 1600m buffers. The associations between physical activity and walkability where the land use mix (LUM) dimension contained LUM1were hypothesised to be weaker than those where it contains LUM2 which, in turn were predicted to be weaker than those where it contained LUM3. This is because LUM1 contained basic land uses considered supportive of physical activity as personal business destinations potentially reached primarily by foot whereas LUM2 additionally contained land uses expected to be supportive of physically activity, in the form of walking or another, for recreation as well as personal business, and LUM3 included even more such destinations providing support for such behaviours. A reduction in the strength of these significant associations was hypothesised after adjustment for individual-level sociodemographic factors and for area deprivation.

Only some of the hypotheses tested held true. Many significant associations did remain, albeit as hypothesised with reduced strength, after adjustment for correlates, indicating independent associations between walkability and physical activity outcomes. However, whilst it was found that physical activity was significantly associated with several models of neighbourhood walkability, individuals living in more walkable neighbourhoods had higher odds of physical activity only when this outcome was specified as walking. In fact, when specified as meeting the WHO recommended physical activity level, which included a range of activities of at least moderate intensity – either excluding or including brisk walking as a contributory one – individuals living in more walkable neighbourhoods as defined by several of the models actually had lower odds, where significant, of the outcome. The finding of only significant positive associations between models of walkability and walking, and of not only a lack of such associations between models of walkability and the other physical activity outcomes, but of significant negative associations between them instead can perhaps be explained in light of Study 2A. Study 2A, an investigation into associations between neighbourhood greenspace and physical activity, found positive associations with greenspace within administratively defined neighbourhoods and meeting the WHO recommended physical activity level as outcomes but negative associations with TTW as the outcome. Associations were stronger when domestic garden was included in the greenspace exposure which fitted with the study's hypothesis that administrative areas with a higher proportion of domestic garden would be more supportive of gardening, contributing to meeting WHO recommended physical activity levels, but that such areas might be less supportive of walking, perhaps due to having inherently lower residential dwelling densities. The walkability results are consistent with the greenspace study's hypothesis: higher walkability areas are likely to contain lower proportions of land designated as domestic garden because the greater the residential dwelling density – residences per unit area of residential-specific land – a dimension of walkability, the lower the area of domestic garden.

Contrary to the hypothesis that significant associations between the physical activity outcomes and walkability would be stronger for neighbourhoods operationalized as smaller units, representing walkable areas, and those constructed as GIS software-computed walkable 1600m buffers, this pattern did not emerge for any given model of walkability. Although odds ratios are only presented within this chapter for CAS wards, the neighbourhood delineation in which walkability was deemed adequately measurable, analysis revealed the strength of significant associations was not dependent on the size of the neighbourhood delineated, and significant associations of similar magnitudes were found for both administratively and non-administratively defined neighbourhoods. It was notable that a greater number of walkability models remained significantly associated with the two World Health Organisation physical activity level recommendations outcomes (WHOPA_E and WHOPA_I) as indicated by tests for trend, after adjustment for correlates, for the smaller administrative unit neighbourhood operationalisations of OAs and CAS wards than of LAs. However, fewer, if any, significant associations remained with these outcomes with neighbourhoods operationalised as buffers –generally smaller than CAS wards but bigger than OAs – and GC200s – generally smaller than OAs, suggesting that smaller areas do not necessarily better represent physical activity-relevant neighbourhoods.

Another of the hypotheses tested, that the strength of the associations between walkability and physical activity outcomes would vary dependent on the model specified, was also largely rejected. There was no discernable pattern with respect to the relative strengths of associations of walkability models that differed in the land use mix component with the physical activity outcomes. Whilst data is only shown for CAS wards, this was the case for all operationalisations of neighbourhoods. To an extent the findings of the present study align with those of Christian *et al*; they found a statistically significant trend for higher odds of walking in more walkable neighbourhoods, irrespective of the model specified [129]. However, Christian *et al*'s physical activity outcomes, which were recreation and transport-related walking also showed distinct relationships with particular walkability models. Transport-related walking was, overall, more strongly associated than recreational walking, with walkability. Transport-related walking was also more strongly related to the walkability model that included "Entertainment, culture and recreation" than it was with the core land uses of "Residential", "Retail", "Office" and "Health, welfare and community" alone. Walking for recreation, by contrast, was most strongly associated with the walkability model which additionally contained the recreational walking-supportive "Public open space, sporting infrastructure, and primary and rural" land uses. Thus, in their study, inter-neighbourhood variation in land use mix was sufficient to produce differential associations with walking. It was not possible in the present study to differentiate walking outcomes but the lack of an additive effect on the strength of associations by the addition of the hypothetically recreation walking-supportive land uses to models suggests that walking was mostly non-recreational in the LWIIP7 sample. Alternatively, another explanation for the lack of a stronger association between models of walkability including recreational walking supportive land uses and the walking outcome, as per Christian *et al*'s study, may lie in geographic differences between the study areas. The population density of London is approximately thirteen fold that of Perth Metropolitan Area, the geographical area from which Christian *et al*'s study population was drawn: of the 875 urban areas of the world, defined as places with populations of over 500,000, London is ranked 44th in terms of population density whilst Perth is placed 823rd [240]. The implication of the substantially higher population density of London is that it may limit land use mix variation between areas because there is less "room" for variation in land use mix within areas of London; deviation from the high mean land use mix in different areas of London is likely to be far more limited than from the relatively much lower mean land use mix found in different areas of Perth. It is possible that between area variation in land use mix – the inter-neighbourhood variation – in any given walkability model was of insufficient magnitude to give differential associations in physical activity outcomes. The high density of London also explains why correlations between walkability index models 1, 2 and 3 scores within a given neighbourhood, however delineated, were found to be high in Chapter 7. It is likely that the majority of neighbourhoods would contain all the land uses specified in model 3, limiting potential for variation in the model scores within a given neighbourhood.

Whilst differences in associations with regard to land use mix variation in walkability models were not marked, it is notable that the full models, M1, M2 and M3, which each contained all three core dimensions of walkability, namely residential dwelling density, junction density and a land use mix, were the most robust models in terms of retaining significant associations with the outcomes after adjustment for correlates as evidenced in Appendix 8.3, Appendix 8.4 and Appendix 8.5. This finding is consistent with Frank *et al*'s study which found that a combined walkability index explained more of the variance in physical activity, even after adjustment for sociodemographic variables than any single one of its constituent physical environment factors [113].

In Study 5B GHQ-12 depression was defined as a score of four or more on the GHQ-12 and it was hypothesised that this outcome would be associated with neighbourhood walkability, with individuals living in neighbourhoods where walkability was higher having lower odds of GHQ-12 depression because walkability is supportive of walking and other physical activities which are protective of mental health. In line with the relationships hypothesised in Study 5A, the association between walkability and GHQ-12 depression was expected to be significantly stronger for neighbourhoods operationalized as smaller units, representing walkable areas, and those constructed as GIS software-computed walkable 1600m buffers, with a reduction in the strength of these significant associations hypothesised after adjustment for individual-level sociodemographic factors and for area deprivation. In Study 5B the finding that walkability was positively associated with GHQ-12 depression among the LWIIP7 sample, countering the stated hypothesis, suggests that the physical activity theorised to protect against this outcome must reach a threshold of intensity given that walkability was found to be negatively associated with meeting the WHO recommended physical activity. However, as the significance of the association was rendered non-significant in this sample upon adjustment for individual level factors, it is likely that the characteristics of individuals living in more walkable areas explain differences in depression outcomes. In light of the multiple significant associations found in the LWIIP7 sample, the lack of significant associations between GHQ-12 depression in the LHSE2008 sample in neighbourhoods delineated as LAs with all but one of the models of walkability, namely residential dwelling density, even before adjustment for individual level factors, may be attributable to the smaller sample size of LSHE2008, approximately half that of LWIIP7. Nevertheless, the failure to detect more than one significant association between walkability and GHQ-12 depression in the LHSE2008 sample, reinforces the apparent negligible role of walkability in regard to this outcome.

In conclusion, the main finding of this part of the present study – that walkability is positively associated with walking – broadly aligns with that of the Australian study on which it is based [129]. Due to differences in the specification of the outcomes between the present study and that one, it is impossible to make direct comparisons. The present study drew a sample from a cohort of older adults of limited socioeconomic diversity, resulting in skewed distributions of physical activity outcomes, and restricting outcome specification to dichotomised variables for statistical analysis, and the Whitehall II study was not designed with walking as an outcome of particular interest, in contrast to the study from which Christian *et al*'s sample was drawn. However, it is improbable that identical specification of outcomes and exposures in a study conducted in a city such as Perth, Australia, would yield the same relationships between walking and walkability as one conducted in London. London has evolved over a longer period and its greater residential dwelling density makes it likely that most neighbourhoods, however delineated, have reached a threshold of walkability above which there is little effect on walking. Therefore, retrofitting London for greater walkability may be futile. Land use has been defined as "the arrangements, activities and inputs by people to produce, change or maintain a certain land cover type" but in London there may be no room, or indeed need, for rearrangement of land and the human activities it underpins. Harnessing the power to exploit the support for physical activity offered by the relatively high walkability of London may be possible and necessary. Land, whatever its use, is immobile but people are not and more walkable environments could be insufficiently "un-driveable" to deter motorised, non-physically active modes of transport that may discourage active travel. Thus, it may be people's interaction with land, driven in part by social norms, that shapes others' use of land. Walkability is a measure of the potential of land to make people walk but land can only offer support: without a will there is no way.

8.5 Summary

In summary this study shows that the nature of neighbourhood walkability and physical activity associations are not entirely as predicted. Many models of neighbourhood walkability are positively associated with physical activity, independent of individual level factors, but only when this outcome is specified as being in the top tertile of the sample for minutes spent walking per week (TTW). However, contrary to the hypothesis, some models of neighbourhood walkability were negatively associated with physical activity as indicated by meeting the WHO recommended physical activity level. Also in contrast to the hypotheses, there was neither tendency for associations between the physical activity outcomes and walkability to be stronger for neighbourhoods operationalized as smaller units, nor for non-administratively defined neighbourhoods. Furthermore, model specification in terms of land uses included has no apparent differential effect on associations, regardless of neighbourhood operationalisations. Lastly, this study shows that walkability was not significantly associated with depression after taking individual level factors into account, again countering expectations.

8.6 Appendices

Appendix 8.1 Frequencies of walkability score quartiles for the land use mix (L1, L2 and L3), the residential dwelling density (R) and the junction density (J) only models, and for the full models (M1, M2 and M3) in the LWIIP7 sample for various neighbourhoods delineations.

Neighbourhood delineation	N	Model	L1	L2	L3	R	J	M1	M2	M3
		Quartile score	%	%	%	%	%	%	%	%
LA	3020	1	46	46	32	31	34	39	39	39
		2	22	16	32	33	27	28	31	22
		3	17	23	19	20	24	18	15	24
		4	14	15	17	16	15	15	15	15
CAS ward	3020	1	37	30	29	30	33	38	38	37
		2	27	28	25	28	27	24	25	27
		3	19	24	27	24	23	22	24	19
		4	17	19	19	18	17	16	14	16
OA	3020	1	36	30	28	43	32	46	44	43
		2	22	27	29	26	28	22	24	26
		3	25	23	23	18	23	17	17	18
		4	17	19	20	12	17	15	14	14
GC2000	2786	1	28	28	29	12	16	22	22	22
		2	30	30	30	22	25	18	19	24
		3	21	23	24	29	27	30	29	27
		4	21	19	16	37	32	30	29	27
GC200	3009	1	42	37	27	26	9	n/a	n/a	n/a
		2	0	4	25	0	34	n/a	n/a	n/a
		3	33	36	27	29	27	n/a	n/a	n/a
		4	25	23	21	46	30	n/a	n/a	n/a
1600m CB	2750	1	0	0	0	16	0	7	6	4
		2	70	64	64	44	56	57	57	60
		3	30	36	36	30	44	36	37	37
		4	0	0	0	11	0	0	0	0
1600m PBNB	2859	1	0	0	0	18	1	9	8	6
		2	66	60	60	43	53	54	55	56
		3	33	39	40	29	46	36	36	37
		4	0	0	0	10	0	1	1	1

Appendix 8.2 Frequencies of walkability score quartiles for the land use mix (L1, L2 and L3), the residential dwelling density (R) and the junction density (J) only models, and for the full models (M1, M2 and M3) in the LHSE2008 neighbourhoods delineated as local authorities (N=1540).

Model	L1	L2	L3	R	J	M1	M2	M3
Quartile score	%	%	%	%	%	%	%	%
1	35	36	26	31	36	38	38	38
2	27	28	31	27	28	28	33	23
3	24	23	25	25	21	19	13	23
4	14	14	19	17	16	16	16	16

Appendix 8.3 Summary of significant negative associations between walkability models in neighbourhoods defined as each spatial unit and WHOPA_E in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation.

Neighbourhood delineation	No adjustment	Adjustment for individual-level factors[a]	Adjustment for individual-level factors & area deprivation[b]
1600m circular buffer	L1, R, J, M1, M2, M3		
1600m polygon based network buffer	R, M2, M3		
Local authority	R, J, M1, M2, M3	R, J	
CAS ward	L1, R, J, M1, M2, M3	R, M3	
Output Area	L1, L2, L3, R, M1, M2, M3	L1, L2, L3, R, M1, M2, M3	L2, L3, R, M1, M2, M3
2000m by 2000m grid cell (GC2000)	L1, L2, R, J, M1, M2, M3	R, M2	R
200m by 200m grid cell (GC200)	L1, L2, L3, R, J	L3, R, J	L3, J
	[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level		

Appendix 8.4 Summary of significant negative associations between all walkability models in neighbourhoods defined as each spatial unit and WHOPA_I in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation.

Neighbourhood delineation	No adjustment	Adjustment for individual-level factors[a]	Adjustment for individual-level factors & area deprivation[b]
1600m circular buffer	L1, L3, R,J, M1, M2, M3	J	
1600m polygon based network buffer	R, J, M1, M2, M3		

Neighbourhood delineation	No adjustment	Adjustment for individual-level factors[a]	Adjustment for individual-level factors & area deprivation[b]
Local authority	L1, L3, R, J, M1, M2, M3	R, M1	
CAS ward	L1, L2, L3, R, J, M1, M2, M3	L3, J, M3	
Output Area	L1, L2, L3, R, J, M1, M2, M3	L1, L2, R, M1, M2, M3	R, M1, M2, M3
2000m by 2000m grid cell (GC2000)	L1, L2, R, J, M1, M2, M3	M2	
200m by 200m grid cell (GC200)	L1, L2, L3, J	J	J
	[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level		

Appendix 8.5 Summary of significant positive associations between all walkability models in neighbourhoods defined as each spatial unit and TTW in the LWIIP7 study sample, before and after adjustment for individual factors and area level deprivation.

Neighbourhood delineation	No adjustment	Adjustment for individual-level factors[a]	Adjustment for individual-level factors & area deprivation[b]
1600m circular buffer	R, J, M1, M2, M3	J, M1, M2	J, M1, M2
1600m polygon based network buffer	R, J, M1, M2, M3	R, J, M1, M2, M3	R, J, M2, M3
Local authority	M2		
CAS ward	L1, R, J, M1, M2, M3	L1, J, M1, M2, M3	L1, J, M1, M2, M3
Output Area	R, J, M1, M3	R	
2000m by 2000m grid cell (GC2000)	L1, R, J, M1, M2, M3	L1, M1	L1, M1
200m by 200m grid cell (GC200)	R	R	R
	[a]Adjustment for sex, age, economic activity, car availability, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level		

Appendix 8.6 Summary of significant positive associations between all walkability models in neighbourhoods defined as each spatial unit and GHQ-12 depression in the LWIIP7 study sample before and after adjustment for individual factors and area level deprivation.

Neighbourhood delineation	No adjustment	Adjustment for individual-level factors[a]	Adjustment for individual-level factors & area deprivation[b]
1600m circular buffer	R, M3		
1600m polygon based network buffer	R, M3		
Local authority	L1, L2, R, J, M1, M2, M3	L2	
CAS ward	R,J, M1, M2, M3		
Output Area	R, M1, M2, M3		
2000m by 2000m grid cell (GC2000)	L1, R, J, M1, M2, M3		
200m by 200m grid cell (GC200)	L2		
	[a]Adjustment for sex, age, economic activity, marital status and ethnicity; [b]Additional adjustment for area deprivation as England-based quintiles of the 2004 Index of Multiple Deprivation (IMD2004) at Lower Super Output Area (LSOA) level		

9 Discussion

9.1 Introduction

9.1.1 Chapter overview

This thesis was formulated in response to the practical need to understand which aspects of neighbourhood physical environments are important for both physical activity and mental health and what constitutes a health-relevant neighbourhood. Its conceptualisation was also based on gaps identified in research literature. An examination of the relationship between physical activity and depression in the study population was presented in the first results chapter, establishing whether it was plausible that independent associations between depression and particular attributes of the physical environment were mediated to some extent by physical activity, where physical activity was also independently associated with those attributes. The following chapters explored separate hypotheses regarding distinct aspects of neighbourhood physical environments – greenspace, environmental quality and walkability – but all shared an underlying theme; to investigate the effect of neighbourhood operationalisation. This chapter aims to provide an interpretation of the results, explain potential implications for public health and suggest directions for future research.

9.1.2 Research questions

At the beginning of this research project the following research questions were posed:

1. Which non-physical environment factors identified in the literature review as putative confounders of relationships between the neighbourhood physical environment exposure variables and depression are associated with depression? (Chapter 4)
2. Which non-physical environment factors identified in the literature review as putative confounders of relationships between the neighbourhood physical environment exposure variables and physical activity are associated with physical activity? (Chapter 4)
3. Is physical activity independently associated with depression? (Chapter 4)
4. Is proportion of neighbourhood greenspace independently associated with physical activity, and with depression, exploring the role of neighbourhood operationalisation (Chapter 5)?
5. Are objective and subjective measures of neighbourhood environmental quality independently associated with physical activity, and with depression, exploring the role of neighbourhood operationalisation? (Chapter 6)
6. How does a walkability index capture the walkability of London, exploring the roles of model component specification and spatial units of enumeration? (Chapter 7)
7. Is neighbourhood walkability independently associated with physical activity, and with depression, and how does the specification of the walkability model affect the associations, exploring the role of neighbourhood operationalisation? (Chapter 8)

9.1.3 Summary of findings

This thesis showed that several factors identified in the literature review as putative confounders of relationships between the neighbourhood physical environment exposure variables and depression, and between the neighbourhood physical environment exposure variables and physical activity, were significantly associated with the outcomes under investigation. Also, this thesis demonstrated that physical activity was related to depression, and that various physical environmental attributes were related to both physical activity and depression, generally in neighbourhoods operationalized as smaller units, but not consistently as hypothesised. Chapter 4 showed that depression was negatively associated with physical activity, independently of individual-level sociodemographic factors and of area-level deprivation. The role of neighbourhood greenspace, and greenspace plus domestic garden, in physical activity and depression was explored in Chapter 5, with both these exposures found to be associated with physical activity independently of individual-level sociodemographic factors and of area-level deprivation, and with modelling of the exposures within smaller units tending to strengthen associations. However, whilst positively associated with meeting WHO recommended physical activity levels, greenspace, and greenspace plus domestic garden, were negatively associated with being in the top tertile of the sample for minutes spent walking per week. Although greenspace, and greenspace plus domestic garden, were negatively associated with depression, these relationships were not independent of individual and area-level deprivation. Chapter 6, an investigation of environmental quality in relation to physical activity, showed that poor (objectively measured) environmental quality was associated with lower odds of meeting WHO recommended physical activity levels, independently of individual-level sociodemographic factors and of area-level deprivation but was not associated with walking as the physical activity outcome. Modelling objective environmental quality exposure within smaller units, as seen for the greenspace exposure, tended to strengthen associations. Multilevel modelling showed the low level of variation in meeting the WHO recommended physical activity level that was found between areas, relative to the variation between individuals, was partly attributable to objective environmental quality. In contrast to the objective measure, perceived environmental quality was not associated with any of the physical activity outcomes. Both objectively measured and perceived environmental quality within neighbourhoods (operationalized as local authorities) were associated with depression, with lower quality giving higher odds of depression. These associations were independent of individual-level sociodemographic factors but only the objective measure was associated with depression independently of area-level deprivation. Chapter 7 and 8 focused specifically on the London region in an exploration of walkability, its measurement and its associations with physical activity and depression. Chapter 7 showed a pattern of radial decay of walkability in London, as indicated by walkability index scores, with Inner London tending to be more walkable than Outer London. This tendency was substantively affected by specificity of neither the land use mix dimension of the walkability index model nor the spatial units of enumeration. Walkability modelled in neighbourhoods operationalized as Census Area Statistics (CAS) wards, administrative units of which there are 633 in London, was shown to be associated with physical activity, independently of individual-level factors and area-level deprivation. However, whilst walkability was positively associated with being in the top tertile of the sample for minutes spent walking per week, it was negatively associated with meeting WHO recommended physical activity levels. Walkability was positively associated with depression, with individuals living in more walkable areas more likely to be depressed, but not after statistical adjustment for individual-level factors and area-level deprivation.

9.2 Physical activity and the physical environment of the neighbourhood

A number of interesting, unexpected relationships between attributes of neighbourhood physical environments and physical activity outcomes are evidenced by this thesis. The two main physical activity outcomes specified, meeting the WHO recommended physical activity level and being in the top tertile for time spent walking per week were qualitatively quite different and this was reflected in their differential associations with neighbourhood physical environment attributes. The WHO recommended physical activity level outcome was derived from measures of participants' self-reported activities classified as of moderate or vigorous intensity. Although two forms of this variable were constructed – one excluding walking and one including walking if there was indication that it was classifiable as moderate for a given individual – associations with any given physical environment attribute exposure were generally of similar magnitude, suggesting among those for whom it was classifiable as moderate, it was a minor component. Thus, the WHO recommended physical activity level variables captured physical activity quite distinct from the walking variable, with those spending the most time walking per week not necessarily meeting the physical activity recommendations. The qualitative distinction between these physical activity outcomes is highlighted in the opposite polarities of their associations, independently of individual level factors and area-deprivation, with neighbourhood greenspace; the association was negative for the walking outcome but positive for the WHO recommended physical activity level one. Causation cannot be inferred but these differential associations raise the question of whether greenspace, as an environmental attribute is facilitating non-walking physical activity whilst also impeding walking. As discussed previously, countering the simple hypothesis that more greenspace promotes all physical activity by providing better access to recreational space to be physically active, the relationships may be more subtle. It seems likely that the greater likelihood of having more garden space among those who live in greener neighbourhoods may account for a greater propensity for doing non-walking physical activity, such as gardening, and thereby meeting the WHO recommended physical activity level. However, the garden theory does not offer a plausible explanation for the putative negative effect of greenspace on walking, commonly regarded as a park-based activity. This relationship may be explained in light of the positive association found between walkability and walking, and the negative one between walkability and meeting the WHO recommended physical activity level. More walkable areas, as operationalized in this thesis, are those with greater residential dwelling density, higher street connectivity and greater land use mix. The positive association found between walkability and walking aligns with the hypothesis that more walkable neighbourhoods do encourage more walking. However, the more walkable neighbourhoods that theoretically promote walking are inherently less green, and those with a higher proportion of greenspace are inherently less walkable. Therefore, it is not greater proportion of greenspace per se that accounts for lower likelihood of being a "heavy" walker; rather it is the associated lower walkability of the neighbourhood. Likewise, it is not greater neighbourhood walkability that explains lower likelihood of meeting the WHO recommended physical activity level but the lower proportion of greenspace associated with such neighbourhoods.

There is evidence of varying strengths for the importance of a multitude of physical environment attributes in relation to physical environments and health outcomes, and the present study has to some extent added to this with respect to greenspace in Chapter 4. People, however, are not exposed to and subject to the influences of a particular aspect of an environmental feature in isolation. For example, the experience of an individual cycling across a park may be simultaneously shaped by not only the size of the park but its cleanliness, the weather, and the air quality. Chapters 5, 6 and 8 addressed this issue with an exploration of more complex, comprehensive, composite indicators of environmental quality. It is interesting that objective but not subjective environmental quality was found to be associated with physical activity in Chapter 5. Lower objectively measured environmental quality within a neighbourhood, with simultaneous account of air quality, greenspace, temperature, sunlight and proximity to industry, was associated with a lower likelihood of meeting the WHO recommended physical activity level, whereas lower perceived environmental quality as local authority-level satisfaction with parks and open spaces, was not. The physical environment attributes evaluated in the perceived measure used in this thesis are not as comprehensive as those of the objective measure and it may be that, regardless of the perspective of the evaluation, a more holistic environmental measure, encompassing more potential influences, is more salient. Also given that it may be the higher level of garden space associated with neighbourhoods with a greater proportion of greenspace, that provides opportunity for physical activity, rather than the greenspace per se, satisfaction with parks and open spaces may be irrelevant.

This thesis developed models of walkability that captured physical environment attributes hypothesised to provide pedestrian support. However, the model that included greenspace, through incorporation of this land use within the land use mix dimension, differed little from the most basic model that did not include it as indicated by high correlation of walkability score between these models for any given unit of enumeration. This suggest that where there is a high degree of land use mix of core land uses such as shops, schools and residences, there will be a co-existence of non-core land uses such as greenspace, and the high degree of land use mix will extend to include these as well. The exploration of modelling walkability in this thesis also revealed that, given the high correlation of walkability scores, operationalisation of neighbourhoods as the administrative units of CAS wards constituted an adequate and much simpler alternative to operationalisation of neighbourhoods as, and deriving walkability scores for 1600m buffers. Therefore, whilst the exploitation of advances in GIS software to model neighbourhoods is attractive and may be useful, it may be unnecessary and its dependence on precise geo-locational information of study participants renders it impossible in many studies.

Results of this thesis show that living in a more walkable neighbourhood is associated with spending more time per week walking, independently of individual-level sociodemographic factors and of area-level deprivation. Whilst longitudinal studies are necessary to determine causal pathways, this suggests that the walkability index is fit for purpose. However, results here also suggest that living in a more walkable neighbourhood, as indicated by walkability index score, may actually discourage other types of physical activity: walkability was negatively associated, independently of individual-level sociodemographic factors and of area-level deprivation, with meeting the WHO recommended physical activity level. It had been hypothesised that all physical activity outcomes would be positively associated with walkability because, as well as providing pedestrian support, more walkable places would give better access to destinations where physical activity could take place. More walkable neighbourhoods, as modelled in the present study, however, have higher residential dwelling densities and therefore less garden space and greenspace more generally, environmental attributes which the results suggest may be supportive of physical activity other than walking.

The opposite polarities of the association of the WHO recommended physical activity level outcome with the greenspace exposure variable (investigated in Chapter 4) and with the walkability variable (investigated in Chapter 8) also highlight the critical importance of consideration about exposure variable measurement in explaining associations. The greenspace exposure variable was simply the proportion of greenspace within a neighbourhood. However, greenspace was also a component of walkability models and – although as discussed its inclusion was perhaps superfluous – its measurement was different, as an indicator of land use mix. Thus, a highly walkable neighbourhood, with a "good" mix of land uses including greenspace, could also have a very low total proportion of greenspace, accounting for distinct associations of a particular physical activity outcome with two variables sharing a common theme of greenspace.

In the sample in which physical activity outcomes were investigated here, LWII7, walking was not a major contributory component for meeting the WHO physical activity level recommendation. Whilst meeting physical activity targets as indicated by objective measurement among older adults has been shown to be associated with self-reported walking for transport [241], there is scant evidence on the contribution of walking to meeting physical activity guidelines in samples more representative than the occupation-specific LWIIP7. However, one recent study found that it made only a small contribution [242]. Therefore, whilst extrapolation of the findings of this thesis to the general population cannot be made, the results could imply that only neighbourhood physical environment attributes positively associated with meeting the WHO physical activity recommendation are conducive to better health. However, physical activity of any intensity promotes better mental health and, regardless of intensity, time spent doing physical activity reduces sedentary behaviour, an outcome independently associated with diseases such as diabetes and cardiovascular disease [243][244]. Thus, attributes of neighbourhood physical environments that are positively associated with walking have potential health benefits as well, particularly among older adults in whom participation in sports is less likely and sedentary time is higher [6].

9.3 Depression and the physical environment of the neighbourhood

This thesis examined relationships between attributes of the neighbourhood physical environment and depression based on the hypothesis that the physical environment would influence mental health directly by, for example provision of opportunity for contact with nature, and indirectly by provision of opportunity for physical activity. Where independent associations were significant they tended to be for the larger, nationally representative study sample of HSE2008, for whom geographical information was restricted to local authorities, precluding inferences regarding the effect of operationalisation on associations. Far fewer independent associations of physical environment attributes were found with depression than with physical activity, perhaps testament to the lesser role of attributes of the physical environment relative to other factors in influencing mental health.

Depression is a complex, non-physical disease and, whilst multiple factors, such as self-efficacy, also determine physical activity, the impact of physical exposure on a physical outcome is easier to fathom: physical attributes of the environment as barriers or otherwise can literally affect an individual's ability to be physically active, regardless of propensity. Nevertheless, this thesis does not reject the hypothesis that attributes of neighbourhood physical environments influence depression outcomes as some significant associations were found, independently of individual-level sociodemographic factors and area-level deprivation. Further, associations found that were only independent of individual-level factors but not of area-level deprivation cannot be used as evidence to reject the theory since area-level deprivation itself includes physical environment dimensions.

People living in greener neighbourhoods, as indicated by the greenspace variable investigated in Chapter 5, were less likely to be depressed, even after accounting for individual-level sociodemographic factors, but statistical analysis suggested that it was area-level deprivation that accounted for this geographical variation in outcome. Inconsistent with the findings here, however, another study found higher levels of neighbourhood greenspace were associated lower odds of depression, independent of individual level and neighbourhood level socioeconomic factors [168]. That study, in which neighbourhoods were defined as census blocks, was conducted in Wisconsin, United States, an area with different greenspace characteristics from those of England, the present study area. It may be that greenspace variation between neighbourhoods in Wisconsin is greater both in terms of quality and quantity, accounting for the apparent differential effect on depression outcomes, whereas between English neighbourhoods there is insufficient variation for differential effects. However, a recent longitudinal study which analysed variation in depression among participants of the British Household Panel Survey (n=65,407) in which neighbourhoods were operationalized as CAS wards found greenspace was associated with better mental health in men and in women, but with distinct interactions with age [245]. It may be that the older age range of the WIIP7 study sample for which neighbourhoods were operationalized as CAS wards precluded detection of significant associations. Whilst the present study operationalized neighbourhoods for greenspace measurement at multiple administrative geographies for the WIIP7 study sample, only operationalisation as local authorities was possible for the nationally representative HSE2008 study sample due to data limitations. Had it been possible to operationalise neighbourhoods as the smaller administrative units of CAS wards for this representative sample, significant independent associations between depression and greenspace may have been found.

The greenspace exposure variable specified was simply a crude measure of the proportion of greenspace within an administrative area so the putative relationship of depression with greenspace, indirectly through physical activity or directly by contact with nature, was perhaps too simplistically modelled. A more sophisticated measure of greenspace pertinent to depression outcomes might include aspects of accessibility, such as its spatial relationship with busy roads that might influence an individual's ability to reach a greenspace for physical activity, or to experience mental health enhancing tranquillity within it. Indeed, results of Chapter 6 suggest this may be the case, with objectively measured lower environmental quality – as indicated by a variable comprising greenspace area alongside factors such as monthly cloud cover and proximity to industry – giving higher odds of depression, independently both of individual-level sociodemographic factors and of area-level deprivation. As for greenspace, walkability as investigated in Chapter 8 does not seem to determine depression outcome. The walkability variable was designed to model all physical activity, but as discussed previously successfully captured only walking and was actually negatively associated with meeting the WHO recommended physical activity level. Thus, the apparent absence of a role for walkability in depression suggests that living in a more walkable neighbourhood that promotes physical activity in the form of walking is not necessarily supportive of mental health. However as evidenced in Chapter 4, physical activity including walking specifically, is independently negatively associated with depression. Thus, it may be that by facilitating walking whilst deterring other physical activity, more walkable neighbourhoods negate the mental health benefits of physical activity. Another explanatory theory is that the walkability measure does not capture aesthetic dimensions of neighbourhood physical environments which would likely produce a differential effect of neighbourhood walking on mental wellbeing and health. For example, walking in a more walkable neighbourhood in terms of navigability and attractiveness might be more protective against depression than one that was less so.

9.4 Physical activity: a mediator for associations between neighbourhood physical environments and depression?

As expected, Chapter 4 showed depression to be negatively associated with physical activity, independent of individual-level sociodemographic factors. Due to its cross-sectional design the investigation of this association could not clarify the exact nature of the relationship and infer causation but, based on the large body of research literature, it is likely to be a bi-directional one: people are more likely to be depressed as a result of lower physical activity levels and also more likely to do less physical activity as a consequence of being depressed. Focussing on the former relational direction, the finding potentiates a mediatory role for physical activity in relationships between physical environments and depression, with certain aspects of physical environments either encouraging or deterring physical activity, which in turn impacts on mental health.

In light of the positive association greenspace showed with physical activity, as indicated by meeting the WHO recommended level, an attractive explanation for the finding that physical activity variables were negatively associated with depression in the WIIP7 study sample, is that greenspace facilitates physical activity and thereby protects against depression. Similarly it is plausible that, by facilitating walking, more walkable areas promote better mental health. However, as discussed the relationships of each of these environmental attributes, as measured in this thesis, with depression did not show independence from individual-level socioeconomic factors and area level deprivation, suggesting that any mental health-enhancing physical activity is not attributable to them. It is nevertheless important not to dismiss potential roles for greenspace and walkability in determining depression outcomes, with physical activity as a mediator, because as discussed, the failure to detect significant independent associations in this thesis may have been due to inadequate modelling of expected effects. This thesis in fact only found one physical environmental attribute exposure variable, the objective measure of environmental quality investigated in Chapter 5, to be significantly associated both with depression, independently of individual-level sociodemographic factors and of area-level deprivation, and with a physical activity outcome. Those living in local authorities of worse objectively measured environmental quality were more likely to be depressed and less likely to meet the WHO physical activity level recommendation, although there was no association between this neighbourhood physical environment attribute and the walking outcome. Thus, it may be that depression is at least partly attributable to low environmental support, as indicated by objective environmental quality, for physical activity.

The differential association a particular neighbourhood physical environment attribute, such as objective environmental quality, with different domains of physical activity may have important implications for any physical activity-mediated effects on mental health. Hamer *et al* found types of physical activity to be differentially associated with a lower risk of psychological distress, as indicated by a score of 4 or more on the GHQ-12 scale, after adjustment for individual level factors [70]. Whilst all activities they studied were independently associated with lower odds of psychological distress, stronger effects were observed for sports than for walking. In Chapter 6, the physical activities contributing to meeting the WHO recommended physical activity level were all moderate to vigorous physical activities which would encompass all types including sports. In view of Hamer *et al*'s research and the findings of Chapter 6 it may be that environmental support for sport is particularly salient in explaining physical-activity mediated relationships between objective environmental quality and depression. However, Hamer *et al*'s research did not take account of the environmental context of the activity and, given evidence for synergistic effects of natural environments and physical activity [106], it is likely that neighbourhood physical environment attributes supportive of mild intensity walking in parks – "green exercise" – for example, are important too.

9.5 Neighbourhood operationalisation

Neighbourhood operationalisation is a critical issue but one that is often overlooked in studies examining the relationships of physical environment attributes with physical activity, and with health [246]. This thesis addressed it through an exploration of the influence of definitions of neighbourhood on these associations. Neighbourhoods were operationalized at three levels of administrative geography; local authorities (LAs) (as the largest), CAS wards and Output Areas (OAs) (as the smallest). Neighbourhood operationalisation was restricted to LAs for the HSE2008 sample as participants' residential locations could not be identified to a lower geography. However, the predominantly London-based WIIP7 study sample was identifiable to postcode-level, enabling operationalisation of neighbourhoods and comparisons of associations at all three administrative geographies. In addition, the ability to geographically "pinpoint" the WIIP7 participants allowed construction of bespoke, individualised neighbourhoods centred on each participant's residence, as simple circular buffers (CBs) or as polygon based network based buffers (PBNBs). These were designed to delineate and represent the area that was "walkable" from an individual's home, which was defined as 1600m in all directions "as the crow flies" for CBs or "as a person walks" for PBNBs.

It was hypothesised that neighbourhood physical environment attributes measured closer to home would have a greater influence on outcomes, and that the bespoke buffers – and more so the one that modelled human rather than avian movement – would represent the extent of the health-relevant neighbourhood most accurately, manifesting in the strongest associations. To an extent this thesis supports this theory because it found associations tended to be stronger for all neighbourhoods smaller than the largest operationalisation, LAs. Across the United Kingdom, and outside London in particular, LAs areas may exceed the areal extent of the neighbourhood to which an individual has daily routine exposure that is likely to impact on their physical activity and health. Nevertheless, it may be valuable to aggregate environmental data as exposure variables in large administrative units, particularly when there is limited geographical data available for the study sample, because individuals will be exposed to and potentially affected by a substantial proportion. Indeed the relevance and potential importance of the LA neighbourhood is evidenced in this thesis by detection of significant independent associations of neighbourhood physical environment exposures at this level, albeit weaker than at lower geographies, with the outcomes.

Although this thesis finds evidence suggesting smaller units may be better at representing health-relevant neighbourhoods, at least with respect to the outcomes investigated – it also concludes that some operationalisations may be too small. This was particularly apparent in construction of the walkability index for London with the spatial units of enumeration as Output Areas (OAs), and the assignment of walkability scores to individuals as the score of the OA in which their residential postcode was located. As discussed in Chapter 7, there were technical difficulties associated with measuring walkability within such small units, leading to possible inaccuracies in the construction of the index. Neighbourhoods operationalized as OAs, especially in London's residentially denser parts can constitute very small areas which, for example, could have the width of just a couple of streets and be walked in only a few paces, therein underrepresenting the likely extent of the area of environmental exposure. The construction of residential postcode centred-buffers as neighbourhoods, particularly that of polygon based network buffers, and measurement of the walkability of each one was complex and computationally demanding. The strength of significant associations between walkability and the physical activity outcomes was not dependent on the size of the neighbourhood delineated, and magnitudes of significant associations for administratively-defined neighbourhoods were similar to those for buffers, making the effort of neighbourhood operationalisation as buffers redundant. Therefore, the simplicity of operationalisation of neighbourhoods as administrative units should not necessarily be construed as inferior to operationalisation as GIS-computed buffers and may be more efficient.

9.6 Evaluation

9.6.1 Limitations

A number of weaknesses have been identified which may have impacted on the findings of this thesis. Whilst significant relationships between attributes of the physical environment and physical activity, and between attributes of the physical environment and depression, were evidenced by this thesis, the criteria for evidence of causality [247] are not met, in part due to the cross-sectional study design. For a relationship to be causal, the effect must happen before the outcome, but this temporality cannot be demonstrated in a cross-sectional study as the “effect” is measured at the same time as the “outcome”. A cross-sectional study only permits one snapshot view in time of a characteristic of an individual, such as their health status, that is hypothesised to be related to an exposure variable. Without knowledge of this characteristic prior to the exposure, it is not possible to determine whether the outcome has changed and, therefore, whether the exposure has affected the outcome.

Another limitation of this thesis concerned exposure to neighbourhood physical environmental attributes. It was not possible to discern the length for which participants in either WIIP7 or HSE2008 resided at the same address. Thus, whilst the study enabled examination of association between outcomes and a particular environmental attribute, the exposure could be controlled only with respect to its spatial dimension – as the specification of neighbourhood operationalisation – but not to its temporal dimension. Thus, two individuals exposed to a particular environmental attribute within a specified neighbourhood could receive a different “dose” in terms of years of exposure. Whilst the cross-sectional design of this study does not permit causal inferences anyway, the lack of duration of residency data prevents standardisation of exposures and thereby compromises the validity of statistical associations.

It is possible that findings and conclusions of this study regarding physical activity outcomes may be inaccurate as evidence suggests self-reports of physical activity overestimate physical activity levels when compared to levels measured objectively by accelerometry [248]. Indeed, for example, in this thesis a high level of accuracy of the self-reported time spent walking per week within the WIIP7 study sample is doubtful: among the 891 individuals in the top tertile for time spent walking, the median time was eight hours per week and ranged from six to sixty three. Even in a younger population the median time would seem an inordinately long time per week to spend walking, and the maximum, averaging nine hours per day, almost impracticable. It is likely that individuals’ recall of their walking behaviours was wrong, or their interpretation of the relevant questionnaire item deviated from the intended one. Also, physical activity data was available for HSE2008 but time constrained the investigation of physical activity outcomes in relation to neighbourhood physical environment attributes to the WIIP7 study sample. The WIIP7 sample comprised only older adults presently or formerly employed in the Civil Service who were geographically over-represented in London and the South East of England, limiting the generalizability of findings. Another limitation of this thesis is that, a part from walking, there was no investigation of specific physical activity domains in relation to the exposure variables which may have given different associations, again a constraint of time in this thesis. Statistical adjustment was not made for physical functioning for which there is evidence for association with physical activity and depression. However, adjustment was made for other related factors such as whether an individual was unemployed due to long-term sickness, reducing the impact of this omission on findings.

A limitation in the measurement of the depression outcome is also identified in this thesis. To establish a diagnosis of depression in accordance with diagnostic criteria [56], a structured diagnostic interview must be carried out. Given that assessment of depression in the present study was made using a self-administered survey instrument, the 12-item General Health Questionnaire (GHQ), it is therefore not certain that those classified in the present study as having the depressed outcome were definitively clinically depressed. However, a recent study found the GHQ to have good criterion validity as a measure of depression [249], suggesting that the use of the survey instrument rather than a clinical interview in the present study would not have had a substantive effect on results.

In modelling relationships of each outcome with each exposure, this study employed only single and multi-level multiple logistic regression, with one dichotomous dependent variable and multiple independent variables which were either dichotomous or continuous ordinal. The continuous variable of time spent walking per week was not normally distributed and was split into tertiles with variables recoded as dummies and then dichotomised around the top tertile into high and low groups, producing a dichotomous outcome variable; individuals falling in the high group were in the top tertile whereas those in the low group were in the bottom or middle This binning approach may have impacted on findings since there was no clear cut off between the top and middle tertiles, with some individuals classified in the "low" walking group spending only a few minutes less walking per week than others in the "high" walking group, leading to over or under estimations of the strength of statistical associations with the exposure variable. However, the approach taken was considered preferable to classifying individuals as walkers and non-walkers because, among walkers, there was large variation in minutes spent walking per week; the dichotomisation performed allowed the possible impact of the neighbourhood environmental attributes on the level of walking to be evaluated.

Finally, a limitation of this thesis relates to measurement of physical environment exposures. It was not always possible to use, or – in the case of the walkability variable – to develop, environmental exposure variables from geographical datasets whose construction matched the date of outcome measurement in the study samples. For, example, whilst the perceived environmental quality variable was derived from a dataset constructed in 2008, and the walkability index contained information pertaining to land use areas measured in 2010, the WIIP7 participant data was extracted in 2004/5. Thus, it is possible that the environmental attributes to which WIIP7 participants were exposed in 2005 were quantitatively and qualitatively differed from those for which statistical associations with their outcomes were tested.

However, it is unlikely that changes over only a few years would have been substantial, rendering the impact on findings minimal.

9.6.2 Strengths

This thesis has several strengths which have enabled significant contributions to the research field. One of the most important strengths was its use of two datasets, which were large both in terms of sample size and participant data. This limited the influence of outliers as extreme observations and allowed detection of statistically significant associations that may not have been detectable with smaller samples. The rich datasets also allowed adjustment for a multitude of sociodemographic factors for which there is evidence for association with physical activity and depression. The HSE2008 dataset, from which depression outcomes were extracted, was nationally representative with regard to participant characteristics and their geographical distribution: the sampling strategy gave an even geographical spread of participants across England and also across London – with approximately the same population within each administrative unit – an asset for the present study. The national representativeness of HSE2008 rendered extrapolation of findings to the English population possible. Whilst participants in HSE2008 were identifiable only to the geographical level of local authorities, the participants of the other study sample employed, WIIP7, and from whom both depression and physical activity outcomes were extracted, could be identified to their residential postcodes. This enabled examination of the effects on associations of neighbourhood operationalisation at a wide range of scales, a privilege enjoyed by few researchers using large study samples in this field of study.

A preliminary review of literature established that it was plausible that the particular neighbourhood environmental attribute variables examined in this thesis caused the specified outcomes of physical activity and of depression. Although all Bradford Hill guidelines [247] should be met to infer causation in relationships between “exposures” and “outcomes”, fulfilment of plausibility lends weight to causal inferences made from statistically significant associations identified in this thesis. Another strength of this study lies in its use of multilevel modelling. This statistical technique allowed the quantification of variation in outcomes between individuals relative to that between the areas, be it local authorities or other administrative units, in which they were “nested”, or resided. Also, the extent to which the measured physical environmental attributes of these areas relative to unmeasured environmental attributes accounted for variation in outcomes between areas could be evaluated. Thereby, it was possible to discern the relative importance of the specified exposure variables. Studies which do not employ multilevel modelling risk over-estimating the putative effects of exposures.

The strengths of this study also relate to the neighbourhood physical environment exposure variables used. In many studies of the effects of perceived attributes of neighbourhood on physical activity and health, perception as an exposure variable is measured in the same individual as the outcome. Given that perceptions are likely to be influenced by individual level sociodemographic factors and health status such as are self-reported by the same individuals confounding bias is highly probably, whereby it will be wrongly assumed that the self-reported perception is associated with the outcome. Instead, the reality may be that both the perception and the outcome are related to the individual sociodemographic factors but not to one another. This study, however, employed an area-aggregated assessment of perceived environmental quality as a measure of local authority-level resident satisfaction derived from the Place Survey that was entirely independent of the study sample, an approach which has been used elsewhere to eliminate single-source bias [250]. With regard to modelling walkability and investigating its relationship with physical activity and depression outcomes, the strength of large size of the WIIP7 study sample lay not only in the provision of better population representativeness, enabling detection of significant differences and reducing quantitative estimation error: it also reduced the chances of discovery failure, an important goal in this exploratory investigation. For example, if the sample comprised only one hundred WIIP7 participants, variation in the walkability of their neighbourhoods would be limited and might not have been sufficient to have differential associations with, and putative effects, on outcomes.

9.7 Future research

This thesis provides evidence that supports the idea that neighbourhood physical environments can influence physical activity behaviours and, to some extent, mental health but the relationships revealed are not straightforward. The challenge then for urban planners and public health professionals is to characterise environments that simultaneously support walking, other physical activities, and mental health, rather than one of these outcomes at the expense of the others. Additionally, effort should be made to identify or create environments that support brisk walking and thereby meeting physical activity level recommendations. Such environments might have lower levels of motorised traffic and more space for pedestrians: less congestion of both people and vehicles would enable a faster pace to be maintained.

Most notably, a potential role of domestic gardens in helping people meet recommended physical activity levels is evidenced in this thesis. This finding warrants further research in the form of longitudinal studies to clarify the nature of the relationship between exposure to domestic gardens and physical activity. Should domestic gardens be found to encourage physical activity such that people meet recommended levels, policy makers would be advised to ensure gardens are integral parts of new housing developments.

Understanding what constitutes a neighbourhood is necessary before examining neighbourhood physical environment effects but it is impossible to provide a definition and operationalisation of neighbourhood that will “fit” any two individuals with regard to physical environmental influences, no matter how similar their characteristics. Nevertheless, within a particular sociodemographic group, such as older adults, relevant neighbourhoods are likely to be more similar in terms of scale than without. Advances in geographical information system (GIS) technologies, which now allow tracking of individuals’ physical activity in space and time, should be exploited to study small samples to gauge neighbourhood scales appropriate for particular sociodemographic groups. Armed with this requisite information, researchers can then conduct longitudinal studies, for which it is feasible to use larger study populations, to examine the effects of physical environments within neighbourhoods defined as administrative units approximating the appropriate scale.

A walkability index constructed in this research for the London region was found to be fit for purpose, with individuals living in more walkable neighbourhoods being significantly more likely to spend more time walking per week, independently of individual-level sociodemographic factors and area-level deprivation. Therefore, further research could replicate its construction and test it for another city or town in the United Kingdom.

9.8 Final conclusions

This thesis found evidence that suggests neighbourhood physical environments play a much lesser role than individual level factors in determining physical activity and mental health. Therefore, evidence-based modification of neighbourhood physical environments to promote healthy behaviour is likely to have only a small effect at a population level. Nevertheless, it is imperative that research efforts continue to identify salient aspects because, as this thesis also demonstrates, apparently obvious relationships are not necessarily so – interventions to promote healthy behaviour could be counter-productive if misguided – and neighbourhoods that are more conducive to physical activity have higher footfall, a step in the right direction on the path to critical mass for behaviour change.

10 References

1 Annual report of the Chief Medical Officer 2009. London: : Department of Health 2009. http://www.empho.org.uk/viewResource.aspx?id=11815 (accessed 20 May2014).

2 WHO The world health report 2002 - Reducing Risks, Promoting Healthy Life. WHO. http://www.who.int/whr/2002/en/ (accessed 4 Dec2013).

3 World Health Organization. WHO Global health risks. Geneva, Switzerland: : World Health Organization 2009. http://www.who.int/healthinfo/global_burden_disease/global_health_risks/en/ (accessed 7 Jan2014).

4 Chaudhury M, Shelton N. Physical activity among 60-69-year-olds in England: knowledge, perception, behaviour and risk factors. *Ageing Soc* 2010;**30**:1343–55.

5 The Chief Medical Officer. At least five a week: evidence on the impact of physical activity and its relationship to health. London: : Department of Health 2004.

6 Health Survey for England 2008 - Volume 1: Physical activity and fitness. London: : National Centre for Social Research 2010. www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/health-survey-for-england/health-survey-for-england--2008-physical-activity-and-fitness

7 Bauman AE, Reis RS, Sallis JF, *et al.* Correlates of physical activity: why are some people physically active and others not? *Lancet* 2012;**380**:258–71. doi:10.1016/S0140-6736(12)60735-1

8 Maxwell JA. *Qualitative research design an interactive approach*. Thousand Oaks (Calif.): : SAGE Publications 2013.

9 Solar O, Irwin A. Towards a Conceptual Framework for Analysis and Action on the Social Determinants of Health: Discussion paper for the Commission on Social Determinants of Health. Geneva: Commission on Social Determinants of Health 2005.

10 Krieger N. A glossary for social epidemiology. *J Epidemiol Community Health* 2001;**55**:693–700. doi:10.1136/jech.55.10.693

11 Sallis JF, Owen N. Ecological models. In: *Health Behavior and Health Education: Theory, Research, and Practice*. San Francisco: Jossey-Bass 1996. 403–24.http://www.med.upenn.edu/hbhe4/part5-ch20.shtml

12 Bronfenbrenner U. *The Ecology of Human Development: Experiments by Nature and Design*. New Ed edition. Cambridge, Mass.: : Harvard University Press 1979.

13 Dahlgren G, Whitehead M. European strategies for tackling social inequities in health: levelling up, Part 2. Copenhagen: : WHO Regional Office for Europe 2007. http://www.euro.who.int/__data/assets/pdf_file/0018/103824/E89384.pdf

14 Evans T, Whitehead M, Diderichsen F, *et al. Challenging Inequities in Health*. Oxford University Press 2001. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195137408.001.0001/acprof-9780195137408 (accessed 1 May2014).

15 Mackenbach JP, van de Mheen H, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in The Netherlands. *Soc Sci Med 1982* 1994;**38**:299–308.

16 Acheson SD. *Independent Inquiry Into Inequalities in Health: Report*. Stationery Office Books 1998.

17 Wendel-Vos W, Droomers M, Kremers S, *et al.* Potential environmental determinants of physical activity in adults: a systematic review. *Obes Rev Off J Int Assoc Study Obes* 2007;**8**:425–40. doi:10.1111/j.1467-789X.2007.00370.x

18 Health Canada. Natural and Built Environments. Ottawa: : Division of Childhood and Adolescence, Health Canada 2002.

19 Schaefer-McDaniel N, Caughy MO, O'Campo P, *et al.* Examining methodological details of neighbourhood observations and the relationship to health: a literature review. *Soc Sci Med 1982* 2010;**70**:277–92. doi:10.1016/j.socscimed.2009.10.018

20 Sugiyama T, Cerin E, Owen N, *et al.* Perceived neighbourhood environmental attributes associated with adults′ recreational walking: IPEN Adult study in 12 countries. *Health Place* 2014;**28C**:22–30. doi:10.1016/j.healthplace.2014.03.003

21 Gebel K, Bauman A, Owen N. Correlates of non-concordance between perceived and objective measures of walkability. *Ann Behav Med Publ Soc Behav Med* 2009;**37**:228–38. doi:10.1007/s12160-009-9098-3

22 Hoehner CM, Brennan Ramirez LK, Elliott MB, *et al.* Perceived and objective environmental measures and physical activity among urban adults. *Am J Prev Med* 2005;**28**:105–16. doi:10.1016/j.amepre.2004.10.023

23 Arvidsson D, Kawakami N, Ohlsson H, *et al.* Physical activity and concordance between objective and perceived walkability. *Med Sci Sports Exerc* 2012;**44**:280–7. doi:10.1249/MSS.0b013e31822a9289

24 MacKerron G, Mourato S. Life satisfaction and air quality in London. *Ecol Econ* 2009;**68**:1441–53.

25 McGinn AP, Evenson KR, Herring AH, *et al.* Exploring associations between physical activity and perceived and objective measures of the built environment. *J Urban Health Bull N Y Acad Med* 2007;**84**:162–84. doi:10.1007/s11524-006-9136-4

26 Adams MA, Ryan S, Kerr J, *et al.* Validation of the Neighborhood Environment Walkability Scale (NEWS) items using geographic information systems. *J Phys Act Health* 2009;**6 Suppl 1**:S113–123.

27 Gebel K, Bauman AE, Sugiyama T, *et al.* Mismatch between perceived and objectively assessed neighborhood walkability attributes: prospective relationships with walking and weight gain. *Health Place* 2011;**17**:519–24. doi:10.1016/j.healthplace.2010.12.008

28 McCormack GR, Cerin E, Leslie E, *et al.* Objective Versus Perceived Walking Distances to Destinations Correspondence and Predictive Validity. *Environ Behav* 2008;**40**:401–25. doi:10.1177/0013916507300560

29 Boehmer TK, Hoehner CM, Wyrwich KW, *et al.* Correspondence Between Perceived and Observed Measures of Neighborhood Environmental Supports for Physical Activity. *J Phys Act Health* 2006;**3**:22–36.

30 Perry C. The Neighborhood Unit. In: *Neighborhood and Community Planning*. New York: Arno Press 1929. 21–140.

31 Boruff BJ, Nathan A, Nijënstein S. Using GPS technology to (re)-examine operational definitions of 'neighbourhood' in place-based health research. *Int J Health Geogr* 2012;**11**:22. doi:10.1186/1476-072X-11-22

32 Thornton LE, Pearce JR, Kavanagh AM. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. *Int J Behav Nutr Phys Act* 2011;**8**:71. doi:10.1186/1479-5868-8-71

33 Kwan M. GIS methods in time-geographic research: geocomputation and geovisualization of human activity patterns. *Geogr Ann* 2005;**86**:267–80.

34 Openshaw S. *The Modifiable Areal Unit Problem*. Norwich: Geo Books 1984.

35 Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? *Soc Sci Med 1982* 2008;**66**:1241–55. doi:10.1016/j.socscimed.2007.11.042

36 Cummins S, Macintyre S, Davidson S, *et al.* Measuring neighbourhood social and material context: generation and interpretation of ecological data from routine and non-routine sources. *Health Place* 2005;**11**:249–60. doi:10.1016/j.healthplace.2004.05.003

37 Longley PA, Batty M. *Spatial Analysis: Modelling in a GIS Environment*. Cambridge: John Wiley & Sons 1996.

38 Moudon A, Lee C, Cheadle A, *et al.* Operational definitions of walkable neighborhood: theoretical and empirical insights. *J Phys Act Health* 2006;**3**:S99–S117.

39 Stafford M, De Silva M, Stansfeld S, *et al.* Neighbourhood social capital and common mental disorder: testing the link in a general population sample. *Health Place* 2008;**14**:394–405. doi:10.1016/j.healthplace.2007.08.006

40 Coulton CJ, Korbin J, Chan T, *et al.* Mapping residents' perceptions of neighborhood boundaries: a methodological note. *Am J Community Psychol* 2001;**29**:371–83.

41 Lee C, Moudon AV. Correlates of walking for transportation or recreation purposes. *J Phys Act Health* 2006;**3**:77.

42 Lee C. *Activity-friendly Communities: Correlates of Transportation Or Recreation Walking, and Correlates of Physical Activity for Different Sub-populations*. University of Washington 2004.

43 Oliver LN, Schuurman N, Hall AW. Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. *Int J Health Geogr* 2007;**6**:41. doi:10.1186/1476-072X-6-41

44 Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health* 2001;**91**:1783–9.

45 Krizek KJ. Operationalizing Neighborhood Accessibility for Land Use-Travel Behavior Research and Regional Modeling. *J Plan Educ Res* 2003;**22**:270–87. doi:10.1177/0739456X02250315

46 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep Wash DC 1974* 1985;**100**:126–31.

47 WHO Global Recommendations on Physical activity for Health. WHO. http://www.who.int/dietphysicalactivity/factsheet_recommendations/en/index.html (accessed 10 Oct2013).

48 Chief Medical Office. UK physical activity guidelines - Publications - GOV.UK. https://www.gov.uk/government/publications/uk-physical-activity-guidelines (accessed 16 Apr2014).

49 Ainsworth BE, Haskell WL, Whitt MC, *et al.* Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000;**32**:S498–504.

50 Adams J. Prevalence and socio-demographic correlates of ‘active transport’ in the UK: analysis of the UK time use survey 2005. *Prev Med* 2010;**50**:199–203. doi:10.1016/j.ypmed.2010.01.006

51 Health and Social Care Information Centre. Statistics on Obesity, Physical Activity and Diet - England, 2014. 2010.http://www.hscic.gov.uk/catalogue/PUB13648 (accessed 13 Apr2014).

52 Coyle N FJ. Health survey for England 2006 London boost: health and lifestyle in London main report. 2009.http://www.thehealthwell.info/node/26547 (accessed 2 Jan2014).

53 World Health Organisation. Mental health: a state of well-being. WHO. http://www.who.int/features/factfiles/mental_health/en/ (accessed 28 Apr2014).

54 World Health Organisation. Mental health: strengthening our response. WHO. http://www.who.int/mediacentre/factsheets/fs220/en/ (accessed 24 Jul2014).

55 World Health Organisation. The global burden of disease: 2004 update. WHO. http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/ (accessed 24 Jul2014).

56 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Dsm-5*. Arlington, VA. Amer Psychiatric Pub Incorporated 2013.

57 Goldberg DP, Gater R, Sartorius N, *et al.* The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychol Med* 1997;**27**:191–7.

58 Kendler KS, Thornton LM, Gardner CO. Genetic Risk, Number of Previous Depressive Episodes, and Stressful Life Events in Predicting Onset of Major Depression. *Am J Psychiatry* 2001;**158**:582–6. doi:10.1176/appi.ajp.158.4.582

59 Cassano P, Fava M. Depression and public health: an overview. *J Psychosom Res* 2002;**53**:849–57.

60 Psychiatric Morbidity Among Adults Living in Private Households. 2001.http://www.ons.gov.uk/ons/rel/psychiatric-morbidity/psychiatric-morbidity-among-adults-living-in-private-households/2000/index.html (accessed 26 Oct2013).

61 Lorant V, Deliège D, Eaton W, *et al.* Socioeconomic Inequalities in Depression: A Meta-Analysis. *Am J Epidemiol* 2003;**157**:98–112. doi:10.1093/aje/kwf182

62 Weich S, Nazroo J, Sproston K, *et al.* Common mental disorders and ethnicity in England: the EMPIRIC study. *Psychol Med* 2004;**34**:1543–51.

63 Hauck K, Rice N. A longitudinal analysis of mental health mobility in Britain. *Health Econ* 2004;**13**:981–1001. doi:10.1002/hec.939

64 Weich S, Twigg L, Lewis G, *et al.* Geographical variation in rates of common mental disorders in Britain: prospective cohort study. *Br J Psychiatry J Ment Sci* 2005;**187**:29–34. doi:10.1192/bjp.187.1.29

65 Halpern D. *Mental health and the built environment: More than bricks and mortar?* Philadelphia, PA, US: : Taylor & Francis 1995.

66 Weich S, Blanchard M, Prince M, *et al.* Mental health and the built environment: cross-sectional survey of individual and contextual risk factors for depression. *Br J Psychiatry* 2002;**180**:428–33. doi:10.1192/bjp.180.5.428

67 Boyle PJ, Gatrell AC, Duke-Williams O. The effect on morbidity of variability in deprivation and population stability in England and Wales: an investigation at small-area level. *Soc Sci Med 1982* 1999;**49**:791–9.

68 Kuper H, Marmot M, Hemingway H. Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease. *Semin Vasc Med* 2002;**2**:267–314. doi:10.1055/s-2002-35401

69 Carson AJ, MacHale S, Allen K, *et al.* Depression after stroke and lesion location: a systematic review. *The Lancet* 2000;**356**:122–6. doi:10.1016/S0140-6736(00)02448-X

70 Hamer M, Stamatakis E, Steptoe A. Dose-response relationship between physical activity and mental health: the Scottish Health Survey. *Br J Sports Med* 2009;**43**:1111–4. doi:10.1136/bjsm.2008.046243

71 Krishnan KRR, Delong M, Kraemer H, *et al.* Comorbidity of depression with other medical diseases in the elderly. *Biol Psychiatry* 2002;**52**:559–88.

72 Joynt KE, Whellan DJ, O'Connor CM. Depression and cardiovascular disease: mechanisms of interaction. *Biol Psychiatry* 2003;**54**:248–61.

73 Rosendorff C, Beeri MS, Silverman JM. Cardiovascular risk factors for Alzheimer's disease. *Am J Geriatr Cardiol* 2007;**16**:143–9.

74 Hamer M, Steptoe A. Association between physical fitness, parasympathetic control, and proinflammatory responses to mental stress. *Psychosom Med* 2007;**69**:660–6. doi:10.1097/PSY.0b013e318148c4c0

75 Kendler KS, Karkowski LM, Prescott CA. Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* 1999;**156**:837–41.

76 Hamer M, Stamatakis E. Objectively assessed physical activity, fitness and subjective wellbeing. *Ment Health Phys Act* 2010;**3**:67–71. doi:10.1016/j.mhpa.2010.09.001

77 Brosse AL, Sheets ES, Lett HS, *et al.* Exercise and the treatment of clinical depression in adults: recent findings and future directions. *Sports Med Auckl NZ* 2002;**32**:741–60.

78 De Moor MHM, Boomsma DI, Stubbe JH, *et al.* Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Arch Gen Psychiatry* 2008;**65**:897–905. doi:10.1001/archpsyc.65.8.897

79 O'Donovan G, Blazevich AJ, Boreham C, *et al.* The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. *J Sports Sci* 2010;**28**:573–91. doi:10.1080/02640411003671212

80 Foster C, Hillsdon M, Thorogood M. Environmental perceptions and walking in English adults. *J Epidemiol Community Health* 2004;**58**:924–8. doi:10.1136/jech.2003.014068

81 Sugiyama T, Giles-Corti B, Summers J, *et al.* Initiating and maintaining recreational walking: a longitudinal study on the influence of neighborhood green space. *Prev Med* 2013;**57**:178–82. doi:10.1016/j.ypmed.2013.05.015

82 Humpel N, Owen N, Leslie E. Environmental factors associated with adults’ participation in physical activity: a review. *Am J Prev Med* 2002;**22**:188–99.

83 Stronegger WJ, Titze S, Oja P. Perceived characteristics of the neighborhood and its association with physical activity behavior and self-rated health. *Health Place* 2010;**16**:736–43. doi:10.1016/j.healthplace.2010.03.005

84 Humpel N, Owen N, Leslie E, *et al.* Associations of location and perceived environmental attributes with walking in neighborhoods. *Am J Health Promot AJHP* 2004;**18**:239–42.

85 Saelens BE, Sallis JF, Black JB, *et al.* Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health* 2003;**93**:1552–8.

86 Craig CL, Marshall AL, Sjöström M, *et al.* International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;**35**:1381–95. doi:10.1249/01.MSS.0000078924.61453.FB

87 Shigematsu R, Sallis JF, Conway TL, *et al.* Age differences in the relation of perceived neighborhood environment to walking. *Med Sci Sports Exerc* 2009;**41**:314–21. doi:10.1249/MSS.0b013e318185496c

88 De Melo LL, Menec V, Porter MM, *et al.* Personal factors, perceived environment, and objectively measured walking in old age. *J Aging Phys Act* 2010;**18**:280–92.

89 Bracy NL, Millstein RA, Carlson JA, *et al.* Is the relationship between the built environment and physical activity moderated by perceptions of crime and safety? *Int J Behav Nutr Phys Act* 2014;**11**:24. doi:10.1186/1479-5868-11-24

90 Cervero R. Accessible Cities and Regions: A Framework for Sustainable Transport and Urbanism in the 21st Century. Institute of Transportation Studies, UC Berkeley 2005. http://econpapers.repec.org/paper/cdlitsrrp/qt27g2q0cx.htm (accessed 16 Apr2014).

91 Levine J, Grengs J, Shen Q, *et al.* Does Accessibility Require Density or Speed? *J Am Plann Assoc* 2012;**78**:157–72. doi:10.1080/01944363.2012.677119

92 Brownson RC, Hoehner CM, Day K, *et al.* Measuring the Built Environment for Physical Activity: State of the Science. *Am J Prev Med* 2009;**36**:S99–S123.e12. doi:10.1016/j.amepre.2009.01.005

93 Frank LD, Pivo G. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, walking. *Transp Res Rec* Published Online First: 1994.http://pubsindex.trb.org/view.aspx?id=425321 (accessed 5 Jul2013).

94 Department for Communities and Local Government. (2006b). Planning policy Statement 3 (PPS3): Housing. London, UK: : Department for Communities and Local Government 2006.

95 English Nature. Accessible natural greenspace in towns and cities: A review of appropriate size and distance criteria. English Nature Research Report 153. Peterborough, UK: English Nature 1995.

96 Maas J, Verheij RA, de Vries S, *et al.* Morbidity is related to a green living environment. *J Epidemiol Community Health* 2009;**63**:967–73. doi:10.1136/jech.2008.079038

97 Mitchell R, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study. *The Lancet* 2008;**372**:1655–60. doi:10.1016/S0140-6736(08)61689-X

98 Lee ACK, Maheswaran R. The health benefits of urban green spaces: a review of the evidence. *J Public Health* 2011;**33**:212–22. doi:10.1093/pubmed/fdq068

99 Coombes E, Jones AP, Hillsdon M. The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Soc Sci Med* 2010;**70**:816–22. doi:10.1016/j.socscimed.2009.11.020

100 Wilson L-AM, Giles-Corti B, Burton NW, *et al.* The association between objectively measured neighborhood features and walking in middle-aged adults. *Am J Health Promot* 2011;**25**:e12–21.

101 Kaczynski AT, Potwarka LR, Saelens BE. Association of park size, distance, and features with physical activity in neighborhood parks. *Am J Public Health* 2008;**98**:1451–6. doi:10.2105/AJPH.2007.129064

102 Giles-Corti B, Broomhall MH, Knuiman M, *et al.* Increasing walking: how important is distance to, attractiveness, and size of public open space? *Am J Prev Med* 2005;**28**:169–76. doi:10.1016/j.amepre.2004.10.018

103 Hansen WG. How Accessibility Shapes Land Use. *J Am Inst Plann* 1959;**25**:73–6. doi:10.1080/01944365908978307

104 Hanibuchi T, Kawachi I, Nakaya T, *et al.* Neighborhood built environment and physical activity of Japanese older adults: results from the Aichi Gerontological Evaluation Study (AGES). *BMC Public Health* 2011;**11**:657. doi:10.1186/1471-2458-11-657

105 Leslie E, Sugiyama T, Ierodiaconou D, *et al.* Perceived and objectively measured greenness of neighbourhoods: Are they measuring the same thing? *Landsc Urban Plan* 2010;**95**:28–33. doi:10.1016/j.landurbplan.2009.11.002

106 Barton J, Pretty J. What is the best dose of nature and green exercise for improving mental health? A multi-study analysis. *Environ Sci Technol* 2010;**44**:3947–55. doi:10.1021/es903183r

107 Forsyth A, Oakes JM, Schmitz KH, *et al.* Does Residential Density Increase Walking and Other Physical Activity? *Urban Stud* 2007;**44**:679–97. doi:10.1080/00420980601184729

108 Steiner RL. Residential density and travel patterns: Review of the literature. *Transp Res Rec* Published Online First: 1994.http://trid.trb.org/view.aspx?id=425320 (accessed 22 Apr2014).

109 Alexander ER. Density measures: A review and analysis. *J Archit Plan Res* 1993;**10**:181–202.

110 Glazier RH, Creatore MI, Weyman JT, *et al.* Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada. *PLoS ONE* 2014;**9**:e85295. doi:10.1371/journal.pone.0085295

111 Twin Cities Walking Study. Design For Health. http://designforhealth.net/resources/other/twin-cities-walking-study/ (accessed 2 May2014).

112 Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med Publ Soc Behav Med* 2003;**25**:80–91.

113 Frank LD, Schmid TL, Sallis JF, *et al.* Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ. *Am J Prev Med* 2005;**28**:117–25. doi:10.1016/j.amepre.2004.11.001

114 Molotch H. The City as a Growth Machine: Toward a Political Economy of Place. In: Kleniewski N, ed. *Cities and Society*. Malden, MA: Blackwell Publishing Ltd 2005. 15–27. http://onlinelibrary.wiley.com/doi/10.1002/9780470752814.ch2/summary (accessed 24 Jul2014).

115 Oakes JM, Forsyth A, Schmitz KH. The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities walking study. *Epidemiol Perspect Innov EPI* 2007;**4**:16. doi:10.1186/1742-5573-4-16

116 Giles-Corti B, Robertson-Wilson J, Wood L, *et al.* The role of the changing built environment in shaping our shape. In Geographies of Obesity: environmental understandings of the obesity epidemic. In: *Geographies of Obesity: environmental understandings of the obesity epidemic*. Surrey: Ashgate Publishing, Ltd.

117 Cervero R, Kockelman K. Travel demand and the 3Ds: Density, diversity, and design. *Transp Res Part Transp Environ* 1997;**2**:199–219. doi:10.1016/S1361-9209(97)00009-6

118 Duncan MJ, Winkler E, Sugiyama T, *et al.* Relationships of land use mix with walking for transport: do land uses and geographical scale matter? *J Urban Health Bull N Y Acad Med* 2010;**87**:782–95. doi:10.1007/s11524-010-9488-7

119 Song Y, Merlin L, Rodriguez D. Comparing measures of urban land use mix. *Comput Environ Urban Syst* 2013;**42**:1–13. doi:10.1016/j.compenvurbsys.2013.08.001

120 Batty M, Besussi E, Maat K, *et al.* Representing Multifunctional Cities: Density and Diversity in Space and Time. *Built Environ* 2004;**30**:324–37. doi:10.2148/benv.30.4.324.57156

121 Ewing R, Cervero R. Travel and the Built Environment. *J Am Plann Assoc* 2010;**76**:265–94. doi:10.1080/01944361003766766

122 Gibbs JW. *Scientific Papers: Thermodynamics*. New York: Dover Publications 1906.

123 Cerin E, Leslie E, du Toit L, *et al.* Destinations that matter: associations with walking for transport. *Health Place* 2007;**13**:713–24. doi:10.1016/j.healthplace.2006.11.002

124 Forsyth A, Hearst M, Oakes JM, *et al.* Design and Destinations: Factors Influencing Walking and Total Physical Activity. *Urban Stud* 2008;**45**:1973–96. doi:10.1177/0042098008093386

125 Frank LD, Sallis JF, Saelens BE, *et al.* The development of a walkability index: application to the Neighborhood Quality of Life Study. *Br J Sports Med* Published Online First: 29 April 2009. doi:10.1136/bjsm.2009.058701

126 Frank LD, Engelke PO. The Built Environment and Human Activity Patterns: Exploring the Impacts of Urban Form on Public Health. *J Plan Lit* 2001;**16**:202–18. doi:10.1177/08854120122093339

127 Maghelal PK, Capp CJ. Walkability: A Review of Existing Pedestrian Indices. *J Urban Reg Inf Syst Assoc* 2011;**23**:5.

128 Frank LD. Land Use and Transportation Interaction Implications on Public Health and Quality of Life. *J Plan Educ Res* 2000;**20**:6–22. doi:10.1177/073945600128992564

129 Christian HE, Bull FC, Middleton NJ, *et al.* How important is the land use mix measure in understanding walking behaviour? Results from the RESIDE study. *Int J Behav Nutr Phys Act* 2011;**8**:55. doi:10.1186/1479-5868-8-55

130 Hajna S, Dasgupta K, Halparin M, *et al.* Neighborhood walkability: field validation of geographic information system measures. *Am J Prev Med* 2013;**44**:e51–55. doi:10.1016/j.amepre.2013.01.033

131 Sallis JF, Frank LD, Saelens BE, *et al.* Active transportation and physical activity: opportunities for collaboration on transportation and public health research. *Transp Res-Part Policy Pract* 2004;**38**:249–68.

132 Badland H, Schofield G. Transport, urban design, and physical activity: an evidence-based update. *Transp Res Part Transp Environ* 2005;**10**:177–96. doi:10.1016/j.trd.2004.12.001

133 Owen N, Humpel N, Leslie E, *et al.* Understanding environmental influences on walking; Review and research agenda. *Am J Prev Med* 2004;**27**:67–76. doi:10.1016/j.amepre.2004.03.006

134 Wendel-Vos GCW, Schuit AJ, de Niet R, *et al.* Factors of the physical environment associated with walking and bicycling. *Med Sci Sports Exerc* 2004;**36**:725–30.

135 Department for Environment, Food and Rural Affairs. Causes of air pollution. 2014. http://uk-air.defra.gov.uk/air-pollution/causes (accessed 25 Apr2014).

136 Cogliano VJ, Baan R, Straif K, *et al.* Preventable exposures associated with human cancers. *J Natl Cancer Inst* 2011;**103**:1827–39. doi:10.1093/jnci/djr483

137 Stieb DM, Judek S, Burnett RT. Meta-analysis of time-series studies of air pollution and mortality: effects of gases and particles and the influence of cause of death, age, and season. *J Air Waste Manag Assoc 1995* 2002;**52**:470–84.

138 Bell ML, Dominici F, Samet JM. A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. *Epidemiol* 2005;**16**:436–45.

139 Ito K, De Leon SF, Lippmann M. Associations between ozone and daily mortality: analysis and meta-analysis. *Epidemiol* 2005;**16**:446–57.

140 Levy JI, Hammitt JK, Spengler JD. Estimating the mortality impacts of particulate matter: what can be learned from between-study variability? *Environ Health Perspect* 2000;**108**:109–17.

141 Schwartz J. Air pollution and daily mortality: a review and meta analysis. *Environ Res* 1994;**64**:36–52. doi:10.1006/enrs.1994.1005

142 COMEAP. Cardiovascular disease and air pollution. Department for Health. 2014 https://www.gov.uk/government/publications/comeap-cardiovascular-disease-and-air-pollution (accessed 18 Apr2014).

143 Day R. Place and the experience of air quality. *Health Place* 2007;**13**:249–60. doi:10.1016/j.healthplace.2006.01.002

144 Marshall JD, Brauer M, Frank LD. Healthy neighborhoods: walkability and air pollution. *Environ Health Perspect* 2009;**117**:1752–9. doi:10.1289/ehp.0900595

145 Brunekreef B, Holgate ST. Air pollution and health. *Lancet* 2002;**360**:1233–42. doi:10.1016/S0140-6736(02)11274-8

146 Basu R, Samet JM. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev* 2002;**24**:190–202.

147 Martens WJM. Climate change, thermal stress and mortality changes. *Soc Sci Med* 1998;**46**:331–44. doi:10.1016/S0277-9536(97)00162-7

148 Merrill RM, Shields EC, White GL Jr, *et al.* Climate conditions and physical activity in the United States. *Am J Health Behav* 2005;**29**:371–81.

149 Tucker P, Gilliland J. The effect of season and weather on physical activity: A systematic review. *Public Health* 2007;**121**:909–22. doi:10.1016/j.puhe.2007.04.009

150 Stamatakis E, Nnoaham K, Foster C, *et al.* The influence of global heating on discretionary physical activity: an important and overlooked consequence of climate change. *J Phys Act Health* 2013;**10**:765–8.

151 Rodrigue J-P, Comtois C, Slack B. *The Geography of Transport Systems*. Routledge 2013.

152 Forsyth A, Krizek KJ. Promoting Walking and Bicycling: Assessing the Evidence to Assist Planners. *Built Environ* 2010;**36**:429–46. doi:10.2148/benv.36.4.429

153 Van Holle V, Deforche B, Van Cauwenberg J, *et al.* Relationship between the physical environment and different domains of physical activity in European adults: a systematic review. *BMC Public Health* 2012;**12**:807. doi:10.1186/1471-2458-12-807

154 Panter JR, Jones A. Attitudes and the environment as determinants of active travel in adults: what do and don't we know? *J Phys Act Health* 2010;**7**:551–61.

155 Saelens BE, Handy SL. Built Environment Correlates of Walking: A Review. *Med Sci Sports Exerc* 2008;**40**:S550–S566. doi:10.1249/MSS.0b013e31817c67a4

156 Duncan MJ, Spence JC, Mummery WK. Perceived environment and physical activity: a meta-analysis of selected environmental characteristics. *Int J Behav Nutr Phys Act* 2005;**2**:11. doi:10.1186/1479-5868-2-11

157 Ried LD, Tueth MJ, Handberg E, *et al.* Validating a self-report measure of global subjective well-being to predict adverse clinical outcomes. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2006;**15**:675–86. doi:10.1007/s11136-005-3515-2

158 Schmidt L, Garratt A, Fitzpatrick R. Instruments for Mental Health: a Review Report from the Patient-reported Health Instruments Group (formerly the Patient-assessed Health Outcomes Programme). Department of Health 2000.

159 Diez-Roux AV, Mair C. Neighborhoods and health. *Ann N Y Acad Sci* 2010;**1186**:125–45. doi:10.1111/j.1749-6632.2009.05333.x

160 Berke EM, Gottlieb LM, Moudon AV, *et al.* Protective association between neighborhood walkability and depression in older men. *J Am Geriatr Soc* 2007;**55**:526–33. doi:10.1111/j.1532-5415.2007.01108.x

161 Wood L, Frank LD, Giles-Corti B. Sense of community and its relationship with walking and neighborhood design. *Soc Sci Med 1982* 2010;**70**:1381–90. doi:10.1016/j.socscimed.2010.01.021

162 Cooper R, Boyko C, Codinhoto MR. *State-of-Science Review: SR-DR2 The Effect of the Physical Environment on Mental Wellbeing*. The Government Office for Science, London. www. foresight. gov. uk/Mental% 20Capital/SR-DR2_MCW. pdf 2008. http://www.dti.gov.uk/assets/foresight/docs/mental-capital/sr-dr2_mcw.pdf (accessed 10 Jan2013).

163 Howard SE. *Garden cities of to-morrow*. Cambridge, MA: MIT Press 1965.

164 Jacobs J. *The death and life of great American cities*. London: Vintage Books 1961.

165 Wilson EO. *Biophilia*. Cambridge: Harvard University Press 1984.

166 Kaplan R, Kaplan S. *The experience of nature: a psychological perspective*. CUP Archive 1989.

167 Aspinall P, Mavros P, Coyne R, *et al.* The urban brain: analysing outdoor physical activity with mobile EEG. *Br J Sports Med* Published Online First: 6 March 2013. doi:10.1136/bjsports-2012-091877

168 Beyer KMM, Kaltenbach A, Szabo A, *et al.* Exposure to neighborhood green space and mental health: evidence from the survey of the health of Wisconsin. *Int J Environ Res Public Health* 2014;**11**:3453–72. doi:10.3390/ijerph110303453

169 Kaplan R. The Nature of the View from Home Psychological Benefits. *Environ Behav* 2001;**33**:507–42. doi:10.1177/00139160121973115

170 Vries S de, Verheij RA, Groenewegen PP, *et al.* Natural environments -- healthy environments? An exploratory analysis of the relationship between greenspace and health. *Environ Plan A* 2003;**35**:1717–31. doi:10.1068/a35111

171 Groenewegen PP, van den Berg AE, de Vries S, *et al.* Vitamin G: effects of green space on health, well-being, and social safety. *BMC Public Health* 2006;**6**:149. doi:10.1186/1471-2458-6-149

172 Grahn P, Stigsdotter UK. The relation between perceived sensory dimensions of urban green space and stress restoration. *Landsc Urban Plan* 2010;**94**:264–75. doi:10.1016/j.landurbplan.2009.10.012

173 Thompson Coon J, Boddy K, Stein K, *et al.* Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *Environ Sci Technol* 2011;**45**:1761–72. doi:10.1021/es102947t

174 Trees and Air Quality. Calif. Air Resour. Board. 2010.http://www.arb.ca.gov/research/ecosys/tree-aq/tree-aq.htm

175 Powe NA, Willis KG. Mortality and morbidity benefits of air pollution (SO2 and PM10) absorption attributable to woodland in Britain. *J Environ Manage* 2004;**70**:119–28.

176 Darçın M. Association between air quality and quality of life. *Environ Sci Pollut Res Int* 2014;**21**:1954–9. doi:10.1007/s11356-013-2101-3

177 Welsch H. Environment and Happiness: Valuation of Air Pollution in Ten European Countries. DIW Berlin, German Institute for Economic Research 2003. http://ideas.repec.org/p/diw/diwwpp/dp356.html (accessed 28 Feb2012).

178 Ferreira S (Susana), Moro M, Clinch JP. Valuing the environment using the life-satisfaction approach. University College Dublin. School of Geography, Planning and Environmental Policy 2006. http://researchrepository.ucd.ie/handle/10197/887 (accessed 7 May2014).

179 Rehdanz K, Maddison D. Local environmental quality and life-satisfaction in Germany. *Ecol Econ* 2008;**64**:787–97. doi:10.1016/j.ecolecon.2007.04.016

180 StataCorp. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp 2009. http://www.stata.com/

181 Esri Products. A Complete GIS Mapping Software System. http://www.esri.com/products (accessed 29 Nov2012).

182 Marmot M, Brunner E. Cohort Profile: the Whitehall II study. *Int J Epidemiol* 2005;**34**:251–6.

183 Craig R, Mindell J, Hirani V. Health Survey for England 2008: physical activity and fitness. Volume 2 - methods and documentation. Leeds: 2009. http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/health-survey-for-england/health-survey-for-england--2008-physical-activity-and-fitness

184 Mindell J, Biddulph JP, Hirani V, *et al.* Cohort Profile: The Health Survey for England. *Int J Epidemiol* Published Online First: 9 January 2012. doi:10.1093/ije/dyr199

185 Van Rossum CT, Shipley MJ, van de Mheen H, *et al.* Employment grade differences in cause specific mortality. A 25 year follow up of civil servants from the first Whitehall study. *J Epidemiol Community Health* 2000;**54**:178–84.

186 Office of the Deputy Prime Minister. National Land Use Database: Land Use and Land Cover Classification (Version 4.4). 2006. http://www.communities.gov.uk/documents/planningandbuilding/pdf/144275.pdf

187 Jansen LJM, Di Gregorio A. Land-use and land cover characterisation and classification: the need for baseline data sets to assess future planning. In: *Proceedings of the IGBP-DIS and IGBP/IHDP-LUCC Data Gathering and Compilation Workshop*. Barcelona: Institute for Geography of Catalonia 1998.

188 Wightman A. Generalised Land Use Database. Update 2006. London: Department for Communities and Local Government 2009. http://www.andywightman.com/docs/GLUD_2006.pdf

189 Richardson EA, Mitchell RJ, Shortt NK, *et al.* Evidence-based selection of environmental factors and datasets for measuring multiple environmental deprivation in epidemiological research. *Environ Health* 2009;**8**:S18. doi:10.1186/1476-069X-8-S1-S18

190 Richardson, EA. Multiple Environmental Deprivation Index metadata. 2011, *pers comm*.

191 Turner EL, Dobson JE, Pocock SJ. Categorisation of continuous risk factors in epidemiological publications: a survey of current practice. *Epidemiol Perspect Innov* 2010;**7**:9. doi:10.1186/1742-5573-7-9

192 Jackson C. The General Health Questionnaire. *Occup Med* 2007;**57**:79–79. doi:10.1093/occmed/kql169

193 Goldberg D, Williams P. *A User's Guide to the General Health Questionnaire*. NFER-Nelson 1991.

194 Sabia S, Dugravot A, Kivimaki M, *et al.* Effect of intensity and type of physical activity on mortality: results from the Whitehall II cohort study. *Am J Public Health* 2012;**102**:698–704. doi:10.2105/AJPH.2011.300257

195 Kirkwood B, Sterne J. *Essential Medical Statistics*. Oxford: Wiley-Blackwell 2003.

196 Stafford M, Marmot M. Neighbourhood deprivation and health: does it affect us all equally? *Int J Epidemiol* 2003;**32**:357–66.

197 Department for Communities and Local Government. Indices of deprivation 2004. 2005. http://webarchive.nationalarchives.gov.uk/+/http://www.communities.gov.uk/archived/general-content/communities/indicesofdeprivation/216309/ (accessed 25 Jul2012).

198 Annear MJ, Cushman G, Gidlow B. Leisure time physical activity differences among older adults from diverse socioeconomic neighborhoods. *Health Place* 2009;**15**:482–90. doi:10.1016/j.healthplace.2008.09.005

199 Fox KR, Hillsdon M, Sharp D, *et al.* Neighbourhood deprivation and physical activity in UK older adults. *Health Place* 2011;**17**:633–40. doi:10.1016/j.healthplace.2011.01.002

200 Watts P, Phillips G, Petticrew M, *et al.* Physical Activity in Deprived Communities in London: Examining Individual and Neighbourhood-Level Factors. *PLoS ONE* 2013;**8**:e69472. doi:10.1371/journal.pone.0069472

201 Mytton OT, Townsend N, Rutter H, *et al.* Green space and physical activity: An observational study using Health Survey for England data. *Health Place* 2012;**18**:1034–41. doi:10.1016/j.healthplace.2012.06.003

202 Eriksson U, Arvidsson D, Gebel K, *et al.* Walkability parameters, active transportation and objective physical activity: moderating and mediating effects of motor vehicle ownership in a cross-sectional study. *Int J Behav Nutr Phys Act* 2012;**9**:123. doi:10.1186/1479-5868-9-123

203 Lindwall M, Larsman P, Hagger MS. The reciprocal relationship between physical activity and depression in older European adults: a prospective cross-lagged panel design using SHARE data. *Health Psychol* 2011;**30**:453–62. doi:10.1037/a0023268

204 Lachowycz K, Jones AP. Does walking explain associations between access to greenspace and lower mortality? *Soc Sci Med* 2014;**107**:9–17. doi:10.1016/j.socscimed.2014.02.023

205 Maas J, Verheij RA, Spreeuwenberg P, *et al.* Physical activity as a possible mechanism behind the relationship between green space and health: A multilevel analysis. *BMC Public Health* 2008;**8**:206. doi:10.1186/1471-2458-8-206

206 ONS Geography ons geography@ons gsi gov uk. Administrative Geography. Off. Natl. Stat. 2010.http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/administrative/index.html (accessed 1 Jun2014).

207 Richardson EA, Mitchell R, Shortt NK, *et al.* Developing summary measures of health-related multiple physical environmental deprivation for epidemiological research. *Environ Plan A* 2010;**42**:1650–68. doi:10.1068/a42459

208 Department for Communities and Local Government. Place Survey 2008-09: Manual. 2008.http://www.communities.gov.uk/publications/localgovernment/placesurveymanual0809 (accessed 12 Jun2012).

209 Dunstan F, Fone DL, Glickman M, *et al.* Objectively Measured Residential Environment and Self-Reported Health: A Multilevel Analysis of UK Census Data. *PLoS ONE* 2013;**8**:e69045. doi:10.1371/journal.pone.0069045

210 Snijders TAB, Bosker RJ. *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. Los Angeles: : Sage 2012.

211 Fox KR. The influence of physical activity on mental well-being. *Public Health Nutr* 1999;**2**:411–8.

212 Mitchell, R, Richardson, E, Pearce, J, *et al.* Is socio-economic inequality in mortality explained by unequal exposure to physical environmental deprivation? A cross-sectional, ecological study. *Rev* 2011.

213 Adams EJ, Goodman A, Sahlqvist S, *et al.* Correlates of walking and cycling for transport and recreation: factor structure, reliability and behavioural associations of the perceptions of the environment in the neighbourhood scale (PENS). *Int J Behav Nutr Phys Act* 2013;**10**:87. doi:10.1186/1479-5868-10-87

214 Ezzati M, Hoorn SV, Lopez AD, *et al.* Comparative Quantification of Mortality and Burden of Disease Attributable to Selected Risk Factors. In: Lopez AD, Mathers CD, Ezzati M, *et al.*, eds. *Global Burden of Disease and Risk Factors*. Washington (DC): : World Bank 2006. http://www.ncbi.nlm.nih.gov/books/NBK11813/ (accessed 26 Dec2013).

215 Adams JGU. The Social Implications of Hypermobility. Speculations about the Social Consequences of the OECD Scenarios for Environmentally Sustainable Transport and Business-as-Usual Trend Projections. 1999.http://discovery.ucl.ac.uk/31687/ (accessed 23 Jun2012).

216 Hupkes G. The law of constant travel time and trip-rates. *Futures* 1982;**14**:38–46. doi:10.1016/0016-3287(82)90070-2

217 Owen N, Bauman A. The descriptive epidemiology of a sedentary lifestyle in adult Australians. *Int J Epidemiol* 1992;**21**:305–10.

218 Bassett DR Jr, Pucher J, Buehler R, *et al.* Walking, cycling, and obesity rates in Europe, North America, and Australia. *J Phys Act Health* 2008;**5**:795–814.

219 Giles-Corti B, Knuiman M, Timperio A, *et al.* Evaluation of the implementation of a state government community design policy aimed at increasing local walking: design issues and baseline results from RESIDE, Perth Western Australia. *Prev Med* 2008;**46**:46–54. doi:10.1016/j.ypmed.2007.08.002

220 Giles-Corti B, Timperio A, Cutt H, *et al.* Development of a reliable measure of walking within and outside the local neighborhood: RESIDE's Neighborhood Physical Activity Questionnaire. *Prev Med* 2006;**42**:455–9. doi:10.1016/j.ypmed.2006.01.019

221 Gibbs JW, Bumstead HA. *Scientific Papers of J. Willard Gibbs: Thermodynamics*. London: Longmans, Green and Company 1906.

222 Western Australian Land Information Authority. Landgate website. http://www.landgate.wa.gov.au/corporate.nsf/web/Spatial+Cadastral+Database (accessed 8 Dec2012).

223 Tobler WR. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ Geogr* 1970;**46**:234. doi:10.2307/143141

224 UK Data Service Census Home. http://census.ukdataservice.ac.uk/ (accessed 22 Nov2013).

225 UKMap – The GeoInformation Group. http://www.geoinformationgroup.co.uk/products/ukmap (accessed 19 Nov2012).

226 Haklay M. UKMap access. 2011, *pers comm*.

227 Ordnance Survey http://www ordnancesurvey co uk/oswebsite/support/contact-us html. OS MasterMap Integrated Transport Network (ITN) Layer - intelligent geographical data. 2013.zotero://attachment/713/ (accessed 28 Jul2014).

228 Ordnance Survey http://www ordnancesurvey co uk/oswebsite/support/contact-us html. OS MasterMap Integrated Transport Network (ITN) Layer - Urban Paths theme. 2013.zotero://attachment/715/ (accessed 28 Jul2014).

229 EDINA. http://edina.ac.uk/ (accessed 20 Nov2012).

230 Using ArcGIS Network Analyst with OS MasterMap ITN Data. http://www.liv.ac.uk/CSD/acuk_html/305.dir/305.docx (accessed 24 Nov2013).

231 Census Dissemination Unit Census Website - CP2. http://cdu.mimas.ac.uk/ (accessed 13 Dec2012).

232 Geoconvert. Geoconvert. 2006.http://geoconvert.mimas.ac.uk/ (accessed 25 Jul2012).

233 UKMap User Guide Version 4.3. 2011.

234 Leslie E, Coffee N, Frank L, *et al.* Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes. *Health Place* 2007;**13**:111–22. doi:10.1016/j.healthplace.2005.11.001

235 Handy SL, Boarnet MG, Ewing R, *et al.* How the built environment affects physical activity: views from urban planning. *Am J Prev Med* 2002;**23**:64–73.

236 Casweb Homepage CP1. http://casweb.mimas.ac.uk/ (accessed 13 Dec2012).

237 National Statistics Postcode Directory: 2006 User Guide. 2006.http://geoconvert1.ds.man.ac.uk/help/documentation/06may/User%20Guide/NSPD%20User%20Guide%202006%20v2.pdf (accessed 11 Feb2013).

238 Google Earth. http://www.google.com/earth/ (accessed 13 Mar2014).

239 Wegener M, Fuerst F. Land-Use Transport Interaction: State of the Art. Rochester, NY: : Social Science Research Network 2004. http://papers.ssrn.com/abstract=1434678 (accessed 15 Dec2013).

240 Demographia World Urban Areas: 9th Annual Edition. Belleville, Illinois 62222 USA: 2013. http://www.demographia.com/db-worldua.pdf (accessed 8 Nov2013).

241 Jefferis BJ, Sartini C, Lee I-M, *et al.* Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. *BMC Public Health* 2014;**14**:382. doi:10.1186/1471-2458-14-382

242 Chatterjee K, Hardin J, Chatterjee K, *et al.* The contribution of walking and cycling to achieving recommended levels of physical activity. Milton Keynes, UK: 2011. http://www.transport.uwe.ac.uk/reports/utsg2011_chatterjee_hardin.pdf (accessed 20 May2014).

243 Proper KI, Singh AS, van Mechelen W, *et al.* Sedentary behaviors and health outcomes among adults: a systematic review of prospective studies. *Am J Prev Med* 2011;**40**:174–82. doi:10.1016/j.amepre.2010.10.015

244 Grøntved A, Hu FB. Television viewing and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a meta-analysis. *JAMA J Am Med Assoc* 2011;**305**:2448–55. doi:10.1001/jama.2011.812

245 Astell-Burt T, Mitchell R, Hartig T. The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health* 2014;**68**:578–83. doi:10.1136/jech-2013-203767

246 Riva M, Apparicio P, Gauvin L, *et al.* Establishing the soundness of administrative spatial units for operationalising the active living potential of residential environments: an exemplar for designing optimal zones. *Int J Health Geogr* 2008;**7**:43. doi:10.1186/1476-072X-7-43

247 Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med* 1965;**58**:295–300.

248 Dishman RK, Washburn RA, Heath G. *Physical Activity Epidemiology*. Human Kinetics 2004.

249 Head J, Stansfeld SA, Ebmeier KP, *et al.* Use of self-administered instruments to assess psychiatric disorders in older people: validity of the General Health Questionnaire, the Center for Epidemiologic Studies Depression Scale and the self-completion version of the revised Clinical Interview Schedule. *Psychol Med* 2013;**43**:2649–56. doi:10.1017/S0033291713000342

250 De Jong K, Albin M, Skärbäck E, *et al.* Area-aggregated assessments of perceived environmental attributes may overcome single-source bias in studies of green environments and health: results from a cross-sectional survey in southern Sweden. *Environ Health Glob Access Sci Source* 2011;**10**:4. doi:10.1186/1476-069X-10-4