Neighbourhood effects: spatial inequalities in tooth decay

Tom Broomhead

A thesis submitted in partial fulfilment of the requirement of the degree of Doctor of Philosophy

Unit of Dental Public Health
School of Clinical Dentistry
University of Sheffield

September 2017

In memory of Betty and Stan Broomhead.

Abstract

Objectives: Little theoretical work has been conducted on the topic of neighbourhood effects on health outcomes, let alone within dentistry. Previous work has often quantified and described outcomes without proper investigation of potential causal mechanisms and pathways. Therefore, the aim of this exploratory research was to investigate features of neighbourhood environments that may influence tooth decay in adults.

Methods: Relevant literature was mapped onto a neighbourhood based theoretical framework to create numerous pathways by which neighbourhoods influence decay. Spatial microsimulation was used to combine data from the Adult Dental Health Survey (2009) with Census data to create a synthetic dataset of individuals at the small area level for the city of Sheffield (UK), including associated socio-economic, demographic and dental characteristics. This data formed the basis of the agent-based models which were used to test the theoretical pathways in two contrasting study areas in Sheffield, as well as a hypothetical scenario involving an extra shop being added to each location.

Results: The trends of the agent-based models indicated that the same pathway (the interaction between shops, diet and sugar intake) had the largest impact in both study areas, leading to statistically significant increases in decay in both cases ($p < 0.05$). The results of the hypothetical simulation involving an extra shop revealed a statistically significant decrease in decay in the more affluent study area ($p < 0.05$), while decay scores remained similar in the less affluent study area.

Conclusions: The findings suggest the interactions between shops, diet and sugar intake may be the most important neighbourhood based mechanisms for tooth decay, regardless of socio-economic status. However, additional simulations pointed to more opportunities to reduce decay in the more affluent study area through the local food environment. The implications of these findings are discussed in light of previous research and future work.

Acknowledgements

Firstly, I would like to express my profound thanks to my primary supervisor, Professor Sarah Baker. Her continued support, advice and enthusiasm over the course of the PhD has been invaluable. Her belief in my abilities motivated me to push on with work even at the most difficult of times. Thanks must also go to my second supervisor, Professor Dimitris Ballas, whose methodological knowledge has been instrumental, as has his general positivity and encouragement throughout.

Thanks also go to the Unit of Dental Public Health, and the Social and Spatial Inequalities research group. I am very grateful for the constructive comments given during my confirmation review by Professor Barry Gibson and Dr Mario Vettore. Special thanks must go to Professor Zoe Marshman for the clinical advice which was vital in shaping the design of the analysis; to Dr Nick Malleson at the University of Leeds, for the time he gave to discuss ideas and concepts that helped in the construction of my models; to Kate Jones of Public Health England for supplying the locations of dental surgeries in Sheffield; and to Dr Mike Croucher of Sheffield Research Software Engineering for his help in getting my simulation models running on Iceberg.

I want to thank and acknowledge my fellow PhD students who I've had the pleasure of working alongside in C14 in Dentistry and D5 in Geography. Special thanks go to Amber Wilson for her optimism and encouragement; To Becky Holloway for the helpful chats and rescuing of her bike; to Beth Kamunge for her entertaining 'dull conversations'; to Azeezat Johnson for the 'GIS advice' and humorous back and forths; to Eoin Gray for the beer breaks; and to Ninu Paul and Rizwana Lala who I started this journey with. To my PhD neighbour Phil Jones, I want to express my sincere gratitude for all the time you gave up to help discuss my work, and advise me when I felt lost with my code.

I cannot emphasise enough my gratitude for the support and kindness of Hannah (and Poppy), my Mum and Dad, Rachael, Sammy and Lawrence over the course of the PhD, as well as for the welcome distractions they've provided when things weren't going as well. I couldn't have finished this without you all. Thank you to all my family and friends who have shown an interest in my work, particularly my badminton friends, and most of all my housemate Steph, who has also had to put up with me over the last four years.

Finally, thank you to Medical Humanities Sheffield for funding my PhD and allowing me to undertake this research. My thanks also go to the estate of the late Professor David Locker, which provided the scholarship that made finishing this PhD possible.

Table of contents

List of figures

List of tables

List of equations

Abbreviations

ABM	Agent-based model
ADHS	Adult Dental Health Survey
ANOVA	Analysis of variance
BMI	Body mass index
CO	Combinatorial optimisation
CSV	Comma separated values
DMFS	Decayed, missing and filled surfaces
DMFT	Decayed, missing and filled teeth
ED	Enumeration district
FE	Further education
GA	Genetic algorithm
GIS	Geographical information system
GMU	George Mason University
GP	General practitioner
GUI	Graphical user interface
HIV	Human immunodeficiency virus
HSCIC	Health and Social Care Information Centre
IMD	Indices of Multiple Deprivation
IPF	Iterative proportional fitting
KISS	‘Keep it simple, stupid!’
LSOA	Lower layer super output area
MSOA	Middle layer super output area
NA	Not available/not applicable
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Service
NOMIS	National Online Manpower Information System
NS-SEC	National statistics socio-economic classification
OA	Output area
ONS	Office for National Statistics

PPM	Parts per million
QGIS	Quantum GIS
SA	Simulated annealing
SAE	Standardised absolute error
SEI	Standard error around the identity
SEP	Socio-economic position
SES	Socio-economic status
SHA	Strategic Health Authority
SMILE	Spatial microsimulation model for the Irish economy
SPSS	Statistical Package for the Social Sciences
TAE	Total absolute error
TRS	Truncate, replicate, sample
UK	United Kingdom
USA	United States of America
US	Understanding Society
WHO	World Health Organisation

Chapter 1 – Introduction

1.1. Background and rationale

Tooth decay is a major public health problem. Kassebaum et al. (2015) found that untreated dental caries in deciduous (milk) teeth was the 10^{th} most prevalent condition worldwide in 2010, affecting 9% of the global population. In the same year, this was found to be the most prevalent condition worldwide in permanent teeth as well, affecting 35% of the global population. Research has shown that inequalities in tooth decay exist in the UK in both children and adults (Watt and Sheiham, 1999), with children in more affluent areas experiencing much lower levels of tooth decay than those living in less affluent areas (Public Health England, 2015).

Inequalities in tooth decay in adults have been identified via a number of social determinants of health (Wilkinson and Marmot, 2003), including income (Costa et al, 2012), socio-economic position (Hobdell et al, 2003), employment (Roberts-Thomson and Stewart, 2008) and education (Brennan et al, 2007). However, little theoretical work has been conducted on the topic of neighbourhood effects on oral health outcomes, and how the environments people live in may affect their teeth. Often outcomes are simply quantified and described, without proper investigation of the potential causal mechanisms and pathways that may influence these interactions. The need for extensive theoretical work to underpin research is vital (Baker and Gibson, 2014), and has been called upon within Dental Public Health in order for the field to mature and develop (Watt, 2002). Analysis of neighbourhood effects on oral health has specifically been identified as a potentially useful line of enquiry (Newton and Bower, 2006). Theoretical frameworks such as those developed by Macintyre and colleagues (2002) offer an opportunity to address these issues.

In general, there has also been a lack of simulation modelling within Dental Public Health. However, simulation methods offer advantages over traditional statistical approcahes that analyse problems in a linear fashion (Speybroeck et al, 2013). Methods such as agent-based modelling include the ability to simulate interactions and inter-dependant feedback mechanisms that occur between individuals, groups and environmental factors over time (Auchincloss and Diez Roux, 2008). Other methods, such as spatial microsimulation, are capable of creating accurate and representative population datasets for a range of variables at different geographical scales (Ballas et al, 2006). Methods such as these have rarely been used in Dental Public Health, and never together. However, the potential for investigating oral health related outcomes using these methods, when guided by appropriate theory, represents an area of huge potential for developing our understanding and testing future public health interventions.

Given this, the rationale for the research conducted in this thesis is as follows. As tooth decay is still a major public health problem, with inequalities still present despite overall levels falling, the lack of clearly thought out geographical analyses within Dental Public Health means that understanding of the importance of place and neighbourhood environments to disparities in this disease are lacking. In looking to address the spatial inequalities that exist it is important to know which mechanisms may be influencing tooth decay in individuals the most. Frameworks such as those proposed by Macintyre et al. (2002) take a comprehensive theoretical approach to hypothesising potentially important neighbourhood level pathways, while accounting for both neighbourhood and individual level constructs. Crucially, simulation methods such as agent-based models offer the chance to test these neighbourhood mechanisms in an interactive and non-linear fashion, by analysing the effects of various theoretical scenarios on an agent population. This combined approach would add to the geographical literature within Dental Public Health by producing a comprehensive theoretical study of neighbourhood based features, as well as furthering understanding of the relationship between these features and tooth decay. Additionally, the application of novel simulation methods to a field where they have so far remained underutilised offers a new approach to studying population level oral health.

1.2. Aim and objectives

Based on the above summary and rationale, the overall aim of this research was to conduct an exploratory analysis to investigate how neighbourhood effects may influence patterns of spatial inequalities in tooth decay in adults. In order to achieve the above aim, the three objectives that need to be met are:

- To identify theoretical pathways by which neighbourhoods influence tooth decay

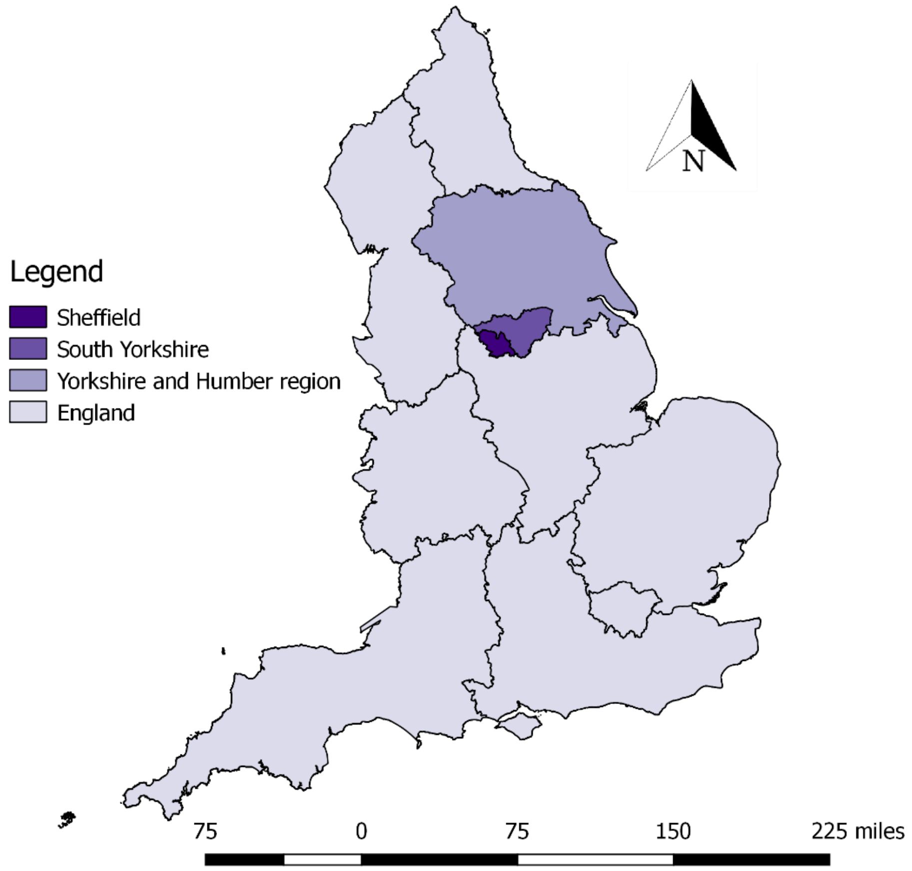
- To build simulation models capable of representing these theoretical pathways
- To use these simulation models to find the most influential theoretical pathways within different neighbourhoods within Sheffield

The first objective will be met through a thorough search of the literature surrounding inequalities in tooth decay, and health inequalities more generally, before being mapped into theoretically relevant pathways. The second objective will be met through the creation of a spatial microsimulation model which will provide a base population dataset of individuals, complete with dental related characteristics on which the theory will be tested. The creation of agent-based models of neighbourhood environments will allow the pathways to be tested in a dynamic and interactive nature, and through this the third objective will be met. This will involve running multiple iterations of these models, to analyse which theoretical pathways have the largest influence on tooth decay.

1.3. Study site

The city of Sheffield was chosen as the site for this research. Sheffield has been shown to have stark inequalities in tooth decay, patterned by social deprivation, in 5 and 12-year-old children and adults (Public Health England, 2015), and as such makes an interesting test site for this research. As of the 2011 Census, the city of Sheffield had a population of 552,698 residents living in 229,928 households (Sheffield City Council, 2011), making it the fifth most populous city in the UK after London, Birmingham, Leeds and Glasgow. This included 227,800 individuals of working age (16-74), while the total population of over 16s was 452,014. Ethnic groups made up 19.2% of Sheffield's population, while the city's population is further boosted by its near 67,000 university students. Historically the city's industry was based around steel and cutlery, which led to a divide between the factories and workers in the east of the city along the Don Valley, and the middle classes who moved to the west of Sheffield away from the heavy industry (Thomas et al, 2009). The steel industry collapsed in the 1980s, leading a shift in the employment structure of the city. Despite this, the historical divides between east and west Sheffield remain. Geographically, Sheffield is located centrally within England, at the south-west tip of the Yorkshire and Humber region (Figure 1). Sheffield is served to the east by the M1 motorway, while a third of the city lies within the Peak District National Park, to the west and north of the city. Sheffield is a city built on hills (sometimes optimistically compared with Rome) which constrains the urban area and provides it with its unique topography.

Figure 1: Sheffield's location within South Yorkshire, the Yorkshire and Humber region, and England



1.4. Key terms for tooth decay

Within dentistry there are numerous terms used to describe decay in teeth. While tooth decay is perhaps colloquially the best known term, clinically there are a far wider range, each with their own specific meanings. As such there are a number of different terms that will be included within this research, specifically when referencing previous studies that have analysed decay in some form. These have been defined below:

- Caries
    - A clinical term synonymous with tooth decay, this refers to the destruction of dental hard tissues through acid producing bacteria, in a process taking place in the oral biofilm (dental plaque) covering the

surfaces of teeth, leading to cavities. Saliva also plays an important role (Selwitz et al, 2007). As such caries can be considered a disease process, a continuum from subclinical subsurface changes, to clinically detectable subsurface changes, to advanced lesions with cavitation of the enamel (Pitts, 2004).

- Cavitated carious lesions
    - Formed as part of the caries process, the net loss of minerals leads to lesions on a tooth (Fejerskov, 1997), which if untreated (by not removing the oral biofilm for example – Kidd and Fejerskov, 2004) can further progress into cavities in the tooth.
- DMFT – decayed, missing and filled teeth
    - One of the most commonly used measures of decay in Dental Public Health, this index can be broken down into its constituent parts (D, M or F) for more specific analysis. Typically, the use of capital letters (DMFT) refers to the teeth of adults and adolescents, while letters left un-capitalised (dmft) refers to the teeth of children.
- DMFS – decayed, missing and filled surfaces
    - Similar to DMFT, this measure applies to surfaces (S), rather than teeth (T). As with the previous measure capitalisation refers to adult and adolescent oral health, with the un-capitalised version referring to children.

All of the above terms, while slightly differing in their definitions, represent tooth decay in one form or another. Thus while different definitions will be used when referring to previous studies, when referring to empirical work as part of this research, and the overall aims of this thesis, the term ‘tooth decay’ will be used throughout. This will act as a more general term to capture any type of decay affecting the teeth of individuals, including the terms defined above. This will also include the outcome variable used for this research, represented by the number of decayed or unsound teeth. Here the term ‘unsound’ refers to teeth with the ‘presence of unrestored dental caries’ (Bower et al, 2007 – p.120).

1.5. Thesis structure

The thesis is structured as follows:

Chapter two includes a review of relevant literature, which firstly covers key geographical concepts including space, place, and context, as well as a review of methods that can be used to study health at the small area level, such as spatial microsimulation and agent-based modelling. Secondly this chapter provides an overview of relevant literature on the determinants of inequalities in tooth decay, before moving on to discuss the current state of the Dental Public Health field with regard to spatial analyses of oral health, as well as simulation modelling of oral health outcomes.

Chapter three introduces the neighbourhood based theoretical framework that was used to guide and conceptualise this research. The Dental Public Health literature on the determinants of tooth decay was applied to this framework, in order to create theoretical pathways by which neighbourhoods may influence tooth decay. The operationalisation of the variables included in the pathways is also discussed.

Chapter four provides a detailed discussion of the workings of a spatial microsimulation model, and how this method was applied to a study of tooth decay. This includes details on how the technique was used to operationalise the data from the theoretical pathways, identified in Chapter three. The validation of these models is then discussed, before a demonstration of the benefits of using this method, and the potential different uses of the output data.

Chapter five provides a detailed discussion on the building of the agent-based models that were used to test the theoretical pathways, using the data from the spatial microsimulation model as their foundation. The first half of this chapter is written in the style of the ODD protocol (Grimm et al, 2006), which was designed to provide a standardised way of describing agent-based models. The second half of this chapter analyses the outputs of the agent-based models (which is not covered by the ODD protocol) for two separate simulation scenarios to determine the main results of this research.

Chapter six evaluates the findings from the agent-based models against the theoretical determinants identified in Chapter two, to see which themes were supported, and which were not, by the results of the agent-based models. A discussion of these findings is followed by a discussion of the trends in the two study areas used in this research, before reviewing the strengths and weakness of the approach taken. Finally, the assumptions used in the agent-based models are discussed.

Chapter seven summarises the findings from Chapter six, and discusses these in the context of the overall aim and objectives outlined in Section 1.2. The main conclusions from this work are then presented, followed by recommendations for future research, as well as potential policy implications.

Chapter 2 – Neighbourhoods, health and inequalities

2.1. Introduction

This literature review is split into two sections. The first will look at some key geographical concepts, before moving onto differing views of contextual studies of health. Methods for measuring health in contexts at the small area level will then also be discussed. The second section will focus on the oral health related elements of this research. Firstly, social and environmental determinants of oral health inequalities in tooth decay will be discussed in order to 'set the scene' for Chapter 3, and the creation of the theoretical pathways. Secondly, the previous uses of spatial analysis and simulation modelling in the Dental Public Health literature will be discussed, to give an idea of the current state of the field. The chapter concludes by presenting a summary of these findings.

2.2. Space, place and context

2.2.1. Space

Space is a fundamental concept to geography, and all work in the discipline contains it. The concept is relational, and only acquires meaning when related to other concepts (Mazur and Urbanek, 1983). Massey (1985) highlighted geography's difficult relationship with the notion of space, noting how before 1960 geography had a focus on places, difference and distinctiveness, and understanding such links, before the quantitative revolution of the 1960s led to notions of causality, scientific law and mathematics leading the way. Massey argues that this insistence on generalisability had shorn geography of one of its most important concerns, a position which was long argued to be untenable. Thrift (2003) attempted to define four main strands of 'space', these being: Empirical space – referring to processes that construct the mundane fabric of daily life; Block space – including processes where pathways of interactions are set up around commonly drawn boundaries; Image space – which refers to how the abundance of images in society create new visions and meaning for spaces; and finally place space, which refers to the ordering of spaces and particular rhythms that naturalise and confirm certain places existence.

2.2.2. Place

The last of Thrift's (2003) strands of space, place, has its own long and contested history. Dating back to Plato and Aristotle, the latter saw place as an essential starting point for the understanding of space, movement and change (Cresswell, 2009). Cresswell (2009) states it was in the 1970's that place began to become an important issue within human geography. Humanistic geographers (Yi-Fu Tuan, 1977) in particular critiqued approaches taken by the quantitative revolution within the discipline (Berry and Garrison, 1958) that occurred in the 1950's and 60's, stating that people needed to be thought of 'as knowing and feeling subjects, rather than either objects or simply rational beings' (Cresswell, 2009 – p.4). Place therefore offered a way of relating to the world, with meanings created based on peoples' experiences.

Cresswell (2009) further states that while humanistic geographers made the distinction between the abstract realm of space, and the felt, or experienced, world of place (Relph, 1976), this approach was criticised for not including notions of power, and its role in the construction, reproduction and contestation of places, and their subsequent meanings. Material structures of place were often the result of decisions made by those in the highest positions, with meanings likely to be assigned by people with the power to do so. Authors such as David Harvey (2000) explored the issue of the dark side of place such as gated communities guarding against perceived threats, while the dichotomy in the role of place has also been discussed by feminist geographers (Dyck, 1989; Walby, 1990), who saw 'home' as a place representing patriarchal authority, labour and even abuse, compared to the 'ideal place' as part of the humanistic approach, where feelings of security and safety dominate. Critical geographers discussed how place and its related meanings can be involved in processes of exclusion through the construction of normative places and values, where people can be seen as either 'in place' or 'out of place'. This idea has been explored across a variety of themes, including class (MacDonald et al, 2005), race (Garland and Chakraborti, 2006), sexuality (Binnie, 1997), gender (Massey, 1994) and disability (Kitchin, 1998). It has also been well documented that these constructions of place are constantly contested, with skateboarding on public furniture, graffiti, and loitering among a number of acts of resistance (Ferrell, 1995; Flusty, 2000).

Does place matter though? And if so, how much, and why? Dorling (2001) provides a persuasive, and relatable argument, stating that you do not need a degree in geography to know that place matters – we see that it matters in the everyday world around us. He cites the example of a subway under a dual-carriageway in his hometown of Oxford that divided housing estates, with life opportunities varying depending on the entrance taken. Walking through any urban area illustrates variation between different parts of a city, which affect residents of these areas in different ways. If all areas were the same then place as a differentiating concept would not matter. But they differ, and because of this the effects of living in different places has been analysed using both quantitative and qualitative methods, on outcomes including income (Galster et al, 2010), education (Roscigno et al, 2006), and health (Macintyre et al, 1993).

2.2.3. Context

The concept of place is similar to context in a number of ways. Gregory et al. (2009) define contextuality as 'the situated character of social life, involving coexistence, connections and 'togetherness' as a series of associations and entanglements in time-space' (p.111). They state that Torsten Hagerstrand (1974) translated the term context into geography as part of the ontological and epistemological basis of time-geography, while describing contextual approaches, where objects and events are treated in their immediate spatial and temporal setting, as having 'a property of 'togetherness' that must not be split' (p.111). Gregory et al. (2009) subsequently define contextual effects as 'the impact of local environments on individuals' attitudes and behaviour' (p.110). They state that much of social science is based on compositional approaches, where behaviours and attitudes are influenced by individuals' non-geographical positions within society, such as social class, where people of similar backgrounds are assumed to behave similarly wherever they live. However, as social interactions in places and neighbourhoods influence such learning, it may be that similar people living in different places act differently due to these interactions with neighbours. The contrast between compositional (individual level) and contextual effects strongly influenced geographical thinking in the 1980s. Thrift (1983) argued that being in the world involves inclusive (being a member of a meaningful world) and positional (a place in the world defined by personal characteristics) features, and that learning one's positional situation is structured contextually in locales.

2.2.4. Measuring places and contexts with regard to health

One contentious issue is how places and contexts should be measured and conceptualised when looking at population health. Cummins et al. (2007) distinguish between 'conventional' ways of measuring place, through specific, relatively static geographical boundaries where physical distance is a key measure, compared to 'relational' approaches, which views places as nodes in networks, where socio-relational distance, power relations and cultural meaning are important, and areas are fluid and capable of change. The authors make three recommendations:

- To recognise that context and place vary across both time and space
- To incorporate appropriate spatial scales to represent both the level at which contextual processes are occurring, and the level where the impact of this will be expressed
- To reject the 'false dualism' of contextual factors and composite factors - which may render traditional definitions too simple in nature

Such false dualisms can lead to outcomes of individuals in groups being treated as independent, a limitation in contextual analysis (Moyses et al, 2006). Macintyre et al. (2002) echo the criticism of treating context and compositional effects as mutually exclusive, introducing a third 'collective' category, focusing more on the cultural, social and historic features and norms of communities.

From a relational view, Cresswell (2009) states a mixture of materiality, meaning and practice are encountered in any place. This includes recognisable material structures, and material forms which pass through them (people, commodities, etc.), which evoke images, or identities of places. Places become places by having meanings attached to them, and the practices that take place in them help shape its meaning and identity. As Cresswell (2009) states, 'the sense we get of a place is heavily dependent on practice and, particularly, the reiteration of practice on a regular basis' (p.2). Such patterns arguably lead to normative dimensions of place.

Popay et al. (2003) argue that those removed from such norms face issues of isolation and humiliation, and associated negative health consequences, creating environments where the protective effects of collective action erode. Smith and Easterlow (2005) identified numerous sets of health trajectories related to housing, demonstrating how the purchasing of houses in different places has a huge impact when combined with a variety of different background contexts and health conditions. Bernard et al. (2007) attempted to understand place effects on health inequalities by focusing on the distribution of more general resources in society, identifying five domains that were key to this: physical proximity; economic price mechanisms; access to institutional resources; community organisations; and local social links. Plane and Klodawsky (2013) have also demonstrated the significance of local amenities such as parks and green space to mental health and anxiety. These spaces acted as social environments that allowed marginalised women to establish friendships and community activities which benefited their mental health, while the perception of these places changed to a more threatening nature at night. Clearly relational views offer a great deal in terms of understanding the effects of place and context on health. As conventional studies form the basis for the approaches taken in this research, these will be assessed more fully in the next section.

2.2.5. 'Conventional' studies of health inequalities

Conventional studies have perhaps not been helped by the number of different definitions presented within the literature. Terms such as 'area', 'place' and 'neighbourhood' are often used interchangeably, but have no consistent definition (Borenstein et al, 2013). Paradoxically, it becomes difficult to study this field without including and using a variety of different search terms, for fear of missing out on certain studies. Multiple definitions have therefore been researched in this section, despite the potential problems associated with this, but in general these studies have a 'conventional' approach (Cummins et al, 2007) due to their quantitative nature.

The analysis of geographical patterning of health dates back to the work of John Graunt in the 17th century, who found higher mortality rates in London than in rural locations, hypothesising environmental pollution as its cause (O'Reilly et al, 2007). Geographical analysis of neighbourhoods and their effects on health can be traced to the pioneering work of John Snow (1813-1858), which investigated cholera outbreaks in Soho, London. More recently, Curtis and Rees-Jones (1998) have argued for geography's place in understanding health inequalities, stating there is empirical evidence that context and places have power in explaining such patterns, independent of individual attributes, despite their review suggesting that individual characteristics 'explain more of the statistical variability in health' (p.667). These sentiments have been echoed in the work of Pickett and Pearl (2001), and Riva et al. (2007), whose reviews found evidence for sometimes modest neighbourhood effects on health, despite the methodological drawbacks of the studies involved. The issues identified by Riva et al. (2007) are detailed in Table 1. Pickett and Pearl (2001) have hypothesised that a perceived lack of contextual studies of health in the past could partly be due to the focus on individual behaviours and interventions by policy makers, combined with attempts to avoid violating the ecological fallacy.

Table 1 – Conceptual and methodological issues associated with neighbourhood effects on health – adapted from Riva et al. (2007)

Conceptual or methodological issue	Description
Conceptualising causal pathways	Lack of fully articulated theoretical frameworks with testable hypotheses
The ecological unit of analysis	Defined as ‘neighbourhoods’, ‘small areas’, ‘local areas’, ‘place’, these labels have been used interchangeably without concern, despite concerns over not appropriately defining the nature of the ecological units of analysis
Defining the spatial contours of ecological units of analysis	Usually defined by administrative and statistical spatial units, these may be of limited value as they lack intrinsic meaning in relation to health, may not correspond to spatial distributions of environmental features, and may not match how residents define their neighbourhoods
Defining ecological exposures	Research has tended to aggregate data from individuals to create neighbourhood based variables, giving less attention to ecological exposures
Controlling for individual-level variables	Lack of consistency in controlling for individual level variables, as well as whether these should be considered as confounders, moderators, or mediators of the association between neighbourhoods and health
Power, sample size, and representativeness	Considerations around statistical power and sample sizes are often neglected
Use of multi-level modelling techniques	These have not been applied regularly enough, potentially limiting novel perspectives from their application. Results

Conceptual or methodological issue	Description
	are also mainly reported for fixed effects, with the ramifications of random effects less well explored
Disentangling context from composition	There are arguments for disentangling the effects of where people live (context) from individual characteristics (composition), whereas some have argued this is a false issue as the two are inextricably intertwined
Dearth of longitudinal and experimental designs	Few studies have research designs beyond cross-sectional, which limits the duration and timing of ecological exposures and ascribing causality. Also, as people and neighbourhoods, and the relation between the two change over time, longitudinal designs take on extra importance

Boyle and Willms (1999) present their own definition of place effects as the combination of ‘contextual and environmental factors that influence individual susceptibility to disease’ (p. 577), however also found that place effects associated with administrative boundaries accounted for very little variance in health. Places defined by using spatial scales from administrative data could be a poor proxy for what ‘places’ or neighbourhoods are, and how people relate to them in a meaningful way (Diez Roux and Mair, 2010; Entwisle, 2007). Vallee et al. (2015) refer to the ‘constant size neighbourhood trap’, referencing the way spatial units fail to account for the ways residents experience their neighbourhoods, using custom GIS algorithms to demonstrate that perceptions of neighbourhood boundaries were often smaller than official, municipal boundaries when considering neighbourhood resources. These boundaries tended to be larger in more affluent areas, and left poorer populations less well served by healthcare services. Haynes et al. (2007) also used custom boundaries to demonstrate that residents in Bristol (UK) described conditions within very small areas from their home as their neighbourhood, contrary to larger administrative boundaries. This highlights the importance of considering what is being studied when deciding which boundaries may be most appropriate (Pickett and Pearl, 2001). Pampalon et al. (1999) echoed these sentiments, finding local level variation in health perception more frequently than at regional level. Despite the faults of administrative boundaries, Owen et al. (2016) state that geographical units such as output areas, created using automated zone design processes to aggregate postcodes, could be a combined to create more meaningful spatial units. Ellen et al. (2001) also leant their support to the idea of Census Tracts as a suitable proxy for neighbourhood conditions, contrary to much of the literature.

Entwisle (2007) states that agency needs to be considered in such studies, as choice can impact on overall patterns. This is problematic for cross-sectional studies, specifically when attempting to draw causal inference (Diez Roux and Mair, 2010). People are likely to change neighbourhoods at various points in their lives (Diez Roux, 2001), and health conditions may not be due entirely to living in the neighbourhood being studied, as people have life histories unrelated to their current place of residence (Pickett and Pearl, 2001). Neighbourhoods themselves are likely to change over time as well, potentially through collective action or patterns of behaviour of residents (Entwisle, 2007), highlighting the benefits of longitudinal approaches (Diez Roux, 2001; Curtis et al, 2004). Further, regional and national level decisions have important impacts on structural and economic factors at the local level. Neighbourhoods are not enclosed enough to withstand the often uneven effects of such decisions at higher geographies (Entwisle, 2007). The danger of focusing on neighbourhoods is that effects of policies or actions in other neighbourhoods could be missed or under accounted for (Owen et al, 2016). However, it has also been argued that local neighbourhood environments are more important to health than the wider regions the neighbourhoods are situated in (Blaxter, 1990).

A number of methodological consideration have also been identified. Macintyre et al, (2002) state that individual characteristics are likely constrained by ‘properties of the locality’ (p.129), and when controls are added to statistical models they may in fact be intervening variables ‘on the pathway between place and health’ (p.129). Pickett and Pearl (2001) have offered similar comments, stating that analyses need to consider how adjusting models may bias the outcome, as well as being explicit about the causal pathways involved. This lack of theorising of pathways and mechanisms that link where people live, and have lived, to their health has also hampered understanding. Place is sometimes seen as ‘some mythical single entity which directly influences health’ (Macintyre et al, 2002 - p.130), rather than a term summing up certain social, material and psychological exposures which affect health.

2.2.6. Empirical conventional studies of contexts and health

Pearson et al. (2013) state that there are three aspects of neighbourhoods that influence health – built, physical and social elements. Their study of health resilience in New Zealand found built features such as access to healthcare and proximity to gambling and alcohol outlets to be significant, while important physical factors included environmental deprivation, access to safe water, and air quality. Living in urban areas, social fragmentation and the churn of residents were the most important social factors. The authors hypothesised that a strong sense of community can be important for health resilience, an idea echoed by Ellen et al. (2001), who roughly concur regarding the three influential health related features of neighbourhoods identified by Pearson et al. (2013).

The trichotomy suggested by Pearson et al. (2013) mirrors the general themes within the health inequalities literature. Examples of relevant studies of built environments includes Macintyre et al. (2003), who found inequalities in access to favourable amenities by housing tenure types, demonstrating the importance of everyday structural elements and amenities. Their study acknowledges that effects are likely to vary by context, fitting with conclusions on health varying significantly at the local level (Pampalon et al, 1999) and broader regional levels (Curtis et al, 2004), supporting the idea that 'one size fits all' policies may not be appropriate. Further examples include Pearce et al. (2006) whose theory based study analysed resource availability in New Zealand, demonstrating strong urban-rural gradients in access to fresh food shops, as well as intra-urban inequalities in access. In a similar study however, despite vegetable consumption increasing with poorer accessibility to fast food outlets, Pearce et al. (2009) found little evidence of an association between fast food locations and poor health. Similarly, MacDonald et al. (2011) demonstrated that groups expected to be more sensitive to their immediate environments (females, the unemployed, and those without cars) were in fact not, when studying the link between supermarket access, diet and BMI. Access to, and increased intake of certain food types are associated with diseases such as diabetes, and Noble et al. (2012) present a novel way of visualising risk factors using various GIS methods, also demonstrating correlations between these risks and deprivation. The study's use of heat mapping presents a potential solution to issues with neighbourhood boundaries suggested by some authors (Pereira et al, 2010).

These studies are important for their inclusion of built environmental features, and not aggregated individual level data (Pickett and Pearl, 2001). Friche et al. (2012) attempted to create neighbourhood level variables through aggregating individual level survey responses on issues such as public services, neighbourhood aesthetics, participation, violence and social capital, and this approach sits somewhere between the two. The use of built environmental features is important, as Macintyre and colleagues (1993) differentiate between studies on the role of areas in influencing health, and those that merely use areas as ‘vehicles’ for exploring links between material deprivation in the causes of ill health.

Examples of the importance of the physical environment includes the work of Pearce et al. (2010), who cite the concept of environmental justice, or preventing groups being 'deprived of their environmental rights' (Cutter, 1995 - p.113), which may impact on health. Analysis of salutogenic and pathogenic variables showed clear associations with environmental deprivation (physical, chemical and biological factors) and income deprivation, with a clear north-south gradient. Mitchell and Popham (2008) highlight the importance of environments in promoting healthy behaviours and attitudes, using clearly defined theoretical pathways to demonstrate that access to green space increased as deprivation decreased, this being associated with lower rates of mortality and circulatory diseases. Inequalities in diseases were also lower in areas with greater access to green space. Ferguson et al. (2013) have shown that material resources such as transportation are also important for accessing recreational physical activities, which was higher by car than by bus among the most affluent quintile in Scotland, with bus access in general having more limited hours.

Social features of local environments also play a key role. These include collective actions, beliefs and behaviours, differentiating the concept from individual behaviours. McNeil et al. (2006) have commented that this influence occurs through creating norms, patterns of social control, varying opportunities to engage in certain behaviours, increased or decreased stress levels, and the constraining of individual choices. Their study highlighted the importance of social support, socio-economic position and inequalities, racial discrimination, and social capital and cohesion as key to differences in physical activity rates. A review by Pearce et al. (2012) came to similar conclusions, finding social practices, contagion effects, neighbourhood crime, disorder and stress, and social capital to be associated with increased smoking levels. Lochner et al. (2003) also demonstrated the importance of social capital, despite finding differences by race in the association between social capital and neighbourhood mortality rates in Chicago.

Within this literature there is a definite theme of neighbourhood level variables having an influence on health, despite some research finding that individual level factors were far more important than those of an ecological nature (Sloggett and Joshi, 1994), while others found only modest effects, based on sometimes flawed methodologies (Pickett and Pearl, 2001; Riva et al, 2007). Nevertheless, there seems enough evidence to make contextual variables worthy of studying to test for their importance to health. There is however no clear agreement on how to represent the relationship between neighbourhoods and health. This section has highlighted the many concerns involved in such studies, and both relational and conventional approaches highlight issues that need addressing. However, Diez Roux (2010) contends that it is not necessarily 'whether the definition of neighbourhoods used in the research is the right one, but whether the measure for the spatial context available is likely to be a reasonable proxy for, or highly correlated with, the true causally relevant spatial context' (p.134). Therefore, a more practical approach may be to ensure that the context of the research is appropriate for the outcomes it is studying. It may also be possible to combine elements of conventional and relational viewpoints, as they share some similar traits. Both point to contributions of individual level characteristics and social networks, as well as area level factors in determining local level health inequalities, and the use of longitudinal research design. Therefore, a theoretical approach using conventional statistics and data, and the attitudes, values and meanings of relational studies, at an appropriate spatial scale and over a certain time period would seem a sensible way to proceed.

2.3. Methods for the analysis of neighbourhoods – spatial microsimulation modelling

Guy Orcutt (1957) is considered the creator of microsimulation, which was used as a method 'to generate large synthetic population data to analyze the impacts of population changes on government policies and vice versa' (Koh et al, 2015 – p.20). Since Orcutt's work the method has been developed for use within geographical studies. These models become spatial when geographical information about the individuals involved is available, with Clarke et al. (1984) the first to integrate this information into a 'spatial' microsimulation model. Thus spatial microsimulation is used as a technique for creating simulated population microdata sets at the small area level (Ballas et al. 2006). The key procedure behind the method is the combining of Census data and survey data, by sampling or reweighting individual level microdata (often surveys) so that they match Census population totals. The 'small area level' at which the analysis is conducted will vary by country depending on data availability, and common variables present in both datasets.

There are a number of types of spatial microsimulation models, with Ballas et al. (2005a) offering a useful overview of the different approaches. The first distinction is between static and dynamic models, with the former based on 'a single snapshot' (p.8), or cross-sectional view of a population at one point in time. Dynamic models on the other hand are used to 'age the attributes of each micro unit' (p.8), constructing a dataset that projects the characteristics of these micro units into the future. There are a number of different types of static microsimulation models, including:

- Synthetic probabilistic reconstruction models - these involve random sampling to select records from microdata sources
- Probabilistic reweighting models – which reweight existing national microdata to fit geographical areas, using random sampling and optimisation techniques
- Deterministic reweighting models - which reweight microdata to fit small area characteristics, but without the use of random sampling, so results are consistent

There are also several dynamic microsimulation approaches:

- Probabilistic dynamic models - which use probabilities to project individuals into the future
- Implicitly dynamic models - which use small area projections, before applying static microsimulation methods to create microdata statistically

Far from being one singular method, there are numerous ways of approaching this technique. For example, iterative proportional fitting (IPF) is a deterministic reweighting method, which Norman (1999) describes as 'a weighting system whereby the original table values are gradually adjusted through repeated calculations to fit the row and column constraints' (p.2). Here, the term 'constraint' refers to variables that are known about an area, i.e. the number of males and females, which as mentioned previously are common to both datasets involved in the simulation. The calculations reweight one source of data to match the cross tabulated column and row totals of another, usually matching survey data to Census totals. A more detailed example of this will be given in Section 4.3, when discussing the models used in this research.

Other reweighting techniques include simulated annealing (SA), a probabilistic reweighting method, which works by iteratively selecting a random combination of individuals from a survey that sum to the same population totals as those present in the Census (Tanton, 2014). A new random sample is collected to address the error in the model, and when this error improves the changes are accepted, and the number of individuals being replaced falls (Morrissey et al, 2008). Another approach is Monte Carlo sampling, a probabilistic dynamic technique using a random sampling method based on probability distributions, which are then applied to individual characteristics (Birkin and Clarke, 1989), and can be used dynamically to age these characteristics. The choice of technique is likely to vary, and to some degree is subjective. Some authors have stated that simulated annealing can have more accurate results (Williamson, 2013), while IPF is notable for its consistency in producing the same results after each run. The method used in this research will be discussed in greater detail in Chapter 4.

Regardless of the technique employed, the method as a whole has several advantages, including overcoming the lack of spatially disaggregated data in surveys, due to the creation of the new datasets at either the individual or household level. This has the potential to help address some of the data issues in relation to neighbourhood based studies identified in Section 2.2.5 (Diez Roux and Mair, 2010), as the method allows researchers to work up from the individual level. Secondly, surveys may underestimate certain groups within a population (Smith et al, 2007) therefore reweighting or sampling survey data at a national level will not include such groups in the data, whereas the matching techniques used in spatial microsimulation means it is possible to include these groups (Morrissey et al, 2008).

It is also worth acknowledging what spatial microsimulation cannot do. Ballas et al. (2005a) have stated that the method is not suited to analysing long-term behavioural trends, particularly in relation to policy changes. This could be problematic for studying oral health, where behaviours, and potential changes in these play an important role (see Section 2.5.2). Additionally, scenarios involving ‘variables that are affected considerably by external and localised factors’ including public transport and transport networks, or disproportionately large industries and single employers are also not suited to such analysis. It should also be remembered that ‘spatial microsimulation outputs are model estimates (the reliability of which depends on a wide range of factors) and not findings' (Ballas et al, 2005a – p. 14). Lovelace and Dumont (2016) have also warned that the outputs of such models are not 'new' datasets as such, as 'the 'new' individuals are simply repeats of individuals we already knew about from the individual level data, albeit in a different order and in different combinations' (p.37). While this is true, it could equally be argued that these datasets did not exist before in anywhere near this much detail, so the outputs of such models are certainly providing something new for researchers to use in understanding population trends and dynamics.

2.3.1. Health related applications of spatial microsimulation

A further advantage of the 'bottom up' approach of spatial microsimulation is that outputs of such models can been used to study health inequalities at a variety of spatial scales. For example, work by Mitchell et al. (2002) has demonstrated the difficulty in using available Census data to study multiple factors associated with patterns of mortality, while also exhibiting the benefits of microsimulation, specifically IPF, in solving such problems. Table 2 contains a list of the studies which have employed spatial microsimulation in the study of health related outcomes. These have been conducted at the national level in numerous countries including Austria (Tomintz et al, 2016), England (Riva and Smith, 2012; Clark et al, 2014), Scotland (Campbell, 2011; Campbell and Ballas, 2013), The Netherlands (De Graaf-Ruizendaal and De Bakker, 2014), The Republic of Ireland (Morrissey et al, 2013a), New Zealand (Smith et al, 2011), Australia (McNamara et al, 2010), Sweden (Brouwers et al, 2010), and the USA (Koh et al, 2015). Fewer analyses have taken place at the regional level, with a study of adult obesity in England (Edwards et al, 2011), an estimate of regional disability in New South Wales, Australia (Lymer et al, 2008), and a study of GP access in the Republic of Ireland (Morrissey et al, 2008) among the examples of these. A higher number of studies have been conducted at the city level (or lower) however, and of these the majority have been conducted in England, and more specifically within the Yorkshire and Humber region. Given that the research within this thesis is concerned with neighbourhoods in a city in this region, it would be prudent to focus on previous analyses within this region and the scenarios these have been applied to, before also reviewing relevant regional and national level studies.

Table 2 – List of spatial microsimulation studies focused on health related outcomes

Authors	Country	Method	Scale	Outcome(s)
Tomintz et al. (2016)	Austria	Deterministic reweighting	State/ municipality	Smoking prevalence
Burden and Steel (2016)	Australia	Simulated annealing	State	Diabetes, angina
Campbell and Ballas (2016)	Scotland	Deterministic reweighting	City (OA)	Smoking prevalence, subjective well-being, alcohol consumption, obesity
Shulman et al. (2015)	England	Simulated annealing	Regional (MSOA)	Hospitalisation rates, heart disease
Koh et al. (2015)	USA	Deterministic reweighting	City (Census Tract)	Obesity prevalence
Morrissey et al. (2015)	Ireland	Simulated annealing	Country (ED)	Depression, access to services
Clark et al. (2014)	England	Combinatorial optimisation	National	Heart disease, stroke, diabetes, cancer, respiratory disease, arthritis
Cataife (2014)	Brazil	Combinatorial optimisation	City (Census Tract)	Obesity and dietary patterns

Authors	Country	Method	Scale	Outcome(s)
De Graaf-Ruizendaal and De Bakker (2014)	The Netherlands	Not stated	National	Provision of local healthcare services
Morrissey et al. (2013a)	Republic of Ireland	Simulated annealing	National/ regional	Hospital utilisation
Morrissey et al. (2013b)	Republic of Ireland	Simulated annealing	County (ED)	Long-term illness, GP access
Campbell and Ballas (2013)	Scotland	Deterministic reweighting	National	Long term illness
Riva and Smith (2012)	England	Deterministic reweighting	National	Psychological distress & alcohol consumption
Hermes and Poulson (2012)	England	Combinatorial optimisation	City (OA)	Smoking prevalence
Edwards et al. (2011)	England	Combinatorial optimisation	Regional (OA)	Obesity
Smith et al. (2011)	New Zealand	Deterministic reweighting	National	Smoking prevalence
Campbell (2011)	Scotland	Deterministic reweighting	National	Inequalities in ill health
Morrissey et al. (2010)	Republic of Ireland	Simulated annealing	National (ED)	Depression, access to services

Authors	Country	Method	Scale	Outcome(s)
McNamara et al. (2010)	Australia	Deterministic reweighting	National	Housing disadvantage
Brouwers et al. (2010)	Sweden	Dynamic	National	Smallpox transmission
Edwards et al. (2010)	England	Deterministic reweighting	City (Census Ward)	Obesogenic environments
Edwards and Clarke (2009)	England	Deterministic reweighting	City (LSOA)	Obesogenic environments
Lymer et al. (2009)	Australia	Deterministic reweighting	State	Need for aged care services
Proctor et al. (2008)	England	Deterministic reweighting	City (LSOA)	Obesogenic environments
Tomintz et al. (2008)	England	Deterministic reweighting	City (OA)	Smoking prevalence
Morrissey et al. (2008)	Republic of Ireland	Deterministic reweighting/ Probabilistic dynamic	Regional	GP access
Lymer et al. (2008)	Australia	Simulated annealing	State	Disability and need for care
Smith et al. (2007)	England	Deterministic reweighing	City (OA)	Obesity, Type 2 Diabetes
Ballas et al. (2006)	England	Implicitly dynamic	City (Census Ward)	Limiting long term illness, anxiety and depression, alcohol and drug problems
Ballas et al. (2005b)	Republic of Ireland	Deterministic reweighting/ Probabilistic dynamic	National	Fertility, mortality

2.3.2. Spatial microsimulation studies of the Yorkshire and Humber region

Spatial microsimulation has been applied at the city level in the Yorkshire and Humber region on numerous occasions. Ballas et al. (2006) used the British Household Panel Survey and small area Census data to simulate anxiety, depression, limiting long term illness, and drug and alcohol problems over a 30-year period in an exploratory analysis based on the population of York, UK. Through use of the 2005 General Household Survey, Tomintz et al. (2008) assessed the effectiveness of stop smoking services in Leeds by estimating individual smoking rates at the output area level. Further research in Leeds includes the work of Procter et al. (2008), who used the SimObesity model to investigate childhood obesity at the Lower Layer Super Output Area (LSOA) level, finding that obesity prevalence decreased in areas with higher incomes. Edwards and Clarke (2009) made further use of this model to investigate patterns in obesogenic environments across Leeds, demonstrating strong links between both social capital and poverty in patterns of childhood obesity. A Further study using the SimObesity model compared obesogenic environments in three Census wards of differing deprivation in Leeds (Edwards et al, 2010), and concluded that obesity levels were highest in the most and least deprived wards. Key local covariates associated with obesity differed however, as neighbourhood safety and fruit and vegetable consumption were more strongly associated with affluent areas, while food expenditure, purchasing school meals, TV ownership and internet access showed the strongest associations in deprived areas.

By optimising their SimHealth model (focusing on diabetes and obesity in Bradford), Smith et al. (2007) found a number of interesting conditions that improved the fit of their IPF models. These included: constraint order (with age proving the most effective first constraint); running the model in clusters of similar areas; and leaving weights as decimals rather than integers. Finally, Shulman et al. (2015) represents the only regional study of this area, analysing small area hospitalisation rates and heart disease in Kirklees, Wakefield, Leeds, Bradford, Harrogate and York at the Middle Layer Super Output Area (MSOA) level. Similar to Edwards et al. (2010), different variables proved important in different locations, with higher levels of heart disease being associated with low social class and the proportion of non-white residents in Bradford, elderly residents in semi-rural Harrogate and York, and a mixture of elderly, low affluent and non-white populations in Leeds.

2.3.3. National and regional level studies

Despite the intention of this study to use spatial microsimulation to study trends across one city, it would seem sensible to look at relevant studies that have been conducted at the national level, and in other regions. The Republic of Ireland has been the focus of an increasing number of studies, with Ballas et al. (2005b) constructing a dynamic spatial microsimulation model (SMILE) focusing on fertility, mortality and internal migration patterns within the country, finding their model compared favourably to other projections based solely on population. The SMILE model acted as the basis for the work of Morrissey et al. (2008), who focused on access to GP services in rural areas of County Galway. This demonstrated that the probability of attending a GP increased 1% with each additional unit increases in income, and that areas with the highest probability of attendance had the lowest levels of access. Morrissey et al. (2010) made further additions to the SMILE model to simulate depression rates in the Republic of Ireland at the Enumeration District (ED) level. The simulated data demonstrated higher rates of depression in the north-west and west of the country, while the addition of a spatial interaction model to the analysis demonstrated that the distribution of psychiatric services did not match the spatial distribution of depression.

Clark et al. (2014) used combinatorial optimisation techniques to analyse a range of morbidity estimates in elderly English populations. Strong regional patterns emerged from the model, with southern regions ranking higher in cancers, while heart disease and respiratory illness were more prevalent in northern areas. Finally, Campbell and Ballas (2016) used the SimAlba model to demonstrate inequalities in the spatial distribution of a variety of health outcomes (See Table 2) in Glasgow, Scotland. ‘Unhealthy’ areas were found to be clustered in the East End of the city, as well as Easterhouse and Castlemilk, compared to far lower frequencies of such areas in the more affluent western regions of the city.

These examples demonstrate the flexibility of the method with regard to the choice of outcome variables. Further flexibility is demonstrated by Ballas et al. (2006), who combined the data from their model with GIS software to visualise the spatial patterning of their output variables, while Morrissey et al. (2010) were able to incorporate a spatial interaction model into their research to test service coverage for individuals with higher levels of depression. This bodes well for innovative approaches to using the technique in the future. Other questions of the method remain however, such as the validity of the outputs. Validation forms a key part of the process, and will be discussed fully in Chapter 4, along with the creation of the models used in this research.

Spatial microsimulation can be used to ascertain more detailed data on individuals and their lives, as well as the spatial patterning of these variables. The models are to some degree limited by the data which is available to them however, and it is important that the data standing as a proxy for a particular variable is suitable. Its inability to accurately model behavioural trends means it may also lack some of the important mechanisms associated with relational and qualitative analyses of health inequalities. However, a microsimulation dataset for use in a dynamic model capable of representing such change, constructed around explicit theory to guide the research, would have huge potential. There is also a significant gap within the dental literature, with no ‘spatial’ microsimulation models having been conducted. Finally, while many studies have been conducted in the Yorkshire and Humber region, none have focused on the region’s second largest city, Sheffield, instead being concentrated primarily in Leeds, as well as York and Bradford.

2.4. Methods for the analysis of neighbourhoods - agent-based modelling

Auchincloss and Diez Roux (2008) define agent-based models as ‘computer representations of systems consisting of a collection of discrete microentities interacting and changing over discrete time steps that give rise to macrosystems’ (p.3). Macal and North (2010) explain that these emerging macro-systems, represented as patterns, structures and behaviours, are ‘not programmed into the models, but arise through the agent interactions’ (p.151). Further, they state that the focus on modelling heterogeneous agents across populations, and ‘the emergence of self-organization’ (p.151) are two key and distinguishing features that result in agent-based models comparing favourably with other simulation methods such as system dynamics and discrete event simulation.

Agent-based models can be traced back to cellular automata models in the 1970’s. These patch based simulations included Gardner’s (1970) Game of Life, while the work of Schelling (1971) represents one of the first, and most famous attempts to model human and societal behaviour. Many early agent-based models took the form of these models, gradually increasing in complexity over time. Sugarscape, a model created by Epstein and Axtell (1996), is a good example of this in the social sciences, with agents free to move from cell to cell, and cells containing spatially distributed resources that agents could acquire from the environment. Agent-based models have evolved as computational capacity has, and are now used across a wide range of academic subjects including molecular modelling, biology, ecology, epidemic and pandemic modelling, computational sciences, economic modelling, market analysis and various other 'real-world' systems including traffic, air traffic control, military manoeuvres and physical infrastructure such as electric power and energy markets (Macal and North, 2010).

Similar to spatial microsimulation, agent-based models have a 'bottom up' approach, concerned with behaviours and characteristics at the individual level. It is the sum of these characteristics and behaviours, and the interaction of these over time, that represents the system level model (Teweldemedhin et al, 2004). This is opposed to 'top-down' approaches, which specify global characteristics and system level behaviours. Systems are then divided into smaller parts, but generally ignore individual level characteristics, exploring these at a system level instead (Teweldemedhin et al, 2004). As such, bottom up approaches are more suited to analyses involving individual interactions in small area geographies. Features of such models are often simplified, due to the difficulty in modelling human processes exactly. While some features of these models can become oversimplified, this adheres to the 'KISS' principle ('Keep it simple, stupid!') introduced by Robert Axelrod (1997). This emphasises the importance of simplicity in design, so as not to make the model, or its output too confusing to interpret.

There is no agreed upon definition of what an agent is, but Crooks and Heppenstall (2012) outline three principles which each agent must adhere to, these being:

1. Autonomy – agents must be 'governed without the influence of centralised control' (p.87), and be able to absorb and exchange information with other agents, who they are free to interact with. This in turn informs decision making.
2. Heterogeneity – 'Agents permit the development of autonomous individuals' (p.87). Agents will have individual attributes regarding age, sex, occupations,

etc. Any groups of agents that exist are created from amalgamations of these autonomous individuals.

3. Activity – ‘Agents are active because they exert independent influence in a simulation’ (p.87). Agents should therefore be: pro-active and goal directed, reactive and perceptive, interactive and communicative, mobile, adaptive and capable of learning, and finally having ‘bounded rationality’ – it is assumed that the choices of agents are rational, and that by bounding this rationality agents can ‘make inductive, discrete, and adaptive choices that move them towards achieving goals’ (p.87).

Further, agent-based models contain rules that have direct effects on agent behaviour and interactions, with these typically based on theory, expert knowledge or data analysis. Rules can be applied across groups or individually, and are usually based around ‘what if’ scenarios, while the behaviour of agents can also be specified when interacting with other agents and their environment. These interactions are key to the running of the models, and Macal and North (2010) state that ‘the two primary issues of modelling agent interactions are specifying who is, or could be, connected to who, and the mechanisms of the dynamics of the interactions’ (p.154). Further, Torrens and McDaniels (2013) have stated that ‘agents’ dynamics in simulation are not scripted; rather, they are processed or computed from a model that determines their behavior given agent characteristics (states) and algorithms (rules) that feed on agents’ endogenous attributes’ (p.23).

Environments within agent-based models can be equally as important, and are demarcated areas in which agents operate, or a ‘miniature laboratory where the attributes and behaviours of agents, and the environment in which they are housed, can be altered and the repercussions observed’ (Crooks and Heppenstall, 2012 - p.90). The locations of agents in these environments may or may not be relevant, depending on the type of study being conducted. Agent-based models are extremely useful for observing events that, when in the field, only become obvious when the process has already started and is progressing (such as segregation), meaning taking records of the event becomes impossible (Batty et al, 2004). The method allows for the reconstruction of such processes and environments, which can be followed from start to finish.

As with spatial microsimulation, validation of agent-based models is crucial to their relevance. This is potentially more difficult given that these models may not have defined output in the same way a microsimulation model does, depending on what is being studied. Heppenstall et al. (2016) have stated that validation is an area that agent-based modelling can still improve in, and this will be discussed in more depth in Chapter 5, which details the agent-based models used in this research.

2.4.1. Health based applications of agent-based models

Agent-based models are well suited to studying health inequalities. Speybroeck et al. (2013) reviewed a number of simulation methods and concluded that 'among all approaches explored in this review, ABM is likely the most suitable tool for studying a complex health inequality situation as it integrates most of the characteristics of a complex system' (p.5762), including feedback loops to provide insights that are potentially beyond traditional statistical models, through complex non-linear interactions and relationships. Due to the highly interdisciplinary nature of the method there are numerous examples of health oriented agent-based models, covering a wide range of topics. There are too many to cover here, so instead this section will aim to give a flavour of the breadth of themes researched using these models in a public health context.

Auchincloss et al. (2011) used agent-based models alongside behavioural economics literature to investigate the effects of economic segregation on healthy food consumption, demonstrating that price was the key driver for preferences, even when low income families had a preference for healthier food. Yang et al. (2011) investigated urban walking patterns, using survey data to calibrate their model. Despite no explicit assumptions being made about socio-economic status, their models showed that those with lower SES walked further for transportation reasons, and less for leisure, demonstrating that patterns emerge 'even in the context of similar starting walking preferences by SES' (p.359). Gorman et al. (2006) demonstrated the flexibility of agent-based models by combining one with a system dynamics model when investigating patterns of drinking behaviour. Their approach showed that susceptible drinkers were likely to convert with the presence of even one current drinker, while also highlighting the importance of agent movement, speed, and resource positioning. Cerda et al. (2014) modelled the population of New York City to study the effects of collective efficacy on violent victimisation. The model was used to test Link and Phelan's (1996) fundamental causes theory, and tested both targeted and universal interventions, before demonstrating that despite universal interventions having the greatest influence on population health, ethnic and racial inequalities could not be addressed without sorting the causes of these inequalities first. Racial disparities were not explicitly programmed, but emerged from the model nonetheless, similar to the findings of Yang et al. (2011).

Agent-based models have also proven popular for modelling the spread of diseases and epidemics. Burke et al. (2006) used the method to analyse response scenarios for a smallpox attack, showing that contact tracing, wide spread vaccination along with isolation of infected individuals limited the effects of such outbreaks. Burke et al. (2006) commented that 'because of their explicit inclusion of physical space, local interactions, and individual heterogeneity, the agent models produce fundamentally different spatiotemporal epidemic dynamics than smoothed differential equation models' (p.1143). Potter et al. (2012) studied the dynamics of school closures on the spread of influenza. They demonstrated that duration of closure was more important at low virus transmissibility, however when transmissibility was high closure length had little effect, with longer closures seeing spikes of the virus upon re-opening. Agent-based models have also been used to investigate historic incidents, as O'Neil and Sattenspiel (2010) demonstrate in their simulation of the 1918-1919 flu epidemic in fur trading communities in central Manitoba, Canada. This analysis identified seasonal residence and mobility as key factors in the spread of the disease. A number of agent-based models have also investigated the dynamics of HIV infections, including the work of Teweldemedhin et al. (2004), which explored HIV prevalence rates in South Africa, and achieved more than 90% accuracy when compared to official statistics. Marshall et al. (2012) used the method to model HIV transmission in New York City, with their analysis allowing for greater analysis of sub groups trends. Further work on the subject of HIV transmission was conducted by Tully et al. (2013), showing that perceived population prevalence levels of HIV rose in line with actual HIV prevalence increases, and as this increase occurred HIV negative individuals were more likely to switch from unprotected to protected sex.

2.4.2. Agent-based models of urban environments

Agent-based models are also well suited to the study of urban environments, and neighbourhood level analysis. Heard et al. (2015) stated that 'when the behaviour of interest is driven by an ensemble of small-scale (often local) actions, agent-based modelling can be a powerful tool' (p.260). Heppenstall et al. (2016) propose three main aspects to consider when studying cities from the bottom up, namely, appropriate spatial scales, temporal scales, and behavioural processes of the agents involved.

There are many examples of studies that have applied this method in city contexts, sometimes with specific geographical contexts in mind. Crooks (2006) has commented that the 'ability to include different features and their attributes of different layers in a GIS, allows for a greater representation of the system of interest when modelling' (p.3), while also removing 'some of the restrictive assumptions imposed by cellular automata style models…while giving access to basic data for setting initial model conditions and parameters' (p.8). Further to this, Crooks (2008) highlights the importance of including not only geography but geometry too, stating this allows 'for different sizes of features such as houses and roads, for example, to be portrayed and how these features might affect the simulation outcomes depending on the processes being modelled' (p.2). This research built on the work of Schelling (1971) by adding geometric features and additional social groups to a model of London wards, and showed that geometry can act as a barrier to segregation. Other examples of agent-based models of residential patterns and processes include the work of Haase et al. (2010) who studied residential mobility in Leipzig, Germany. Their analysis based on federal and Census data supported the idea of Leipzig being a shrinking city, while also accommodating areas of urban and suburban growth. The changes in segregation resulting from this analysis had further implications for land use, ecosystems, and environmental resources.

Geometry has been included in other models of urban systems, including the work of Malleson et al. (2013), who incorporated GIS into their study of crime hotspots in East and South-East Leeds, while also testing future scenarios for crime reduction initiatives. Batty et al. (2003) demonstrate how such models can incorporate geometry to focus on specific events within neighbourhoods, such as the Notting Hill Carnival. Their research challenges traditional assumptions that the system in question has evolved through behaviours which are independent of extraneous controls, or will subsume these if not. The authors state though that it is impossible to think of an event such as a carnival as anything but controlled. The paper demonstrates how alternative routings for the carnival, including less circular routes, led to reduced crowd densities and higher perceived public safety. Agent-based models have also been used to analyse the behaviour of crowds in riotous situations within neighbourhoods (Torrens and

McDaniel, 2013), using literature from the field to inform parameters within the model. Perhaps unsurprisingly, agent-based models have often been used for the analysis of common urban concerns such as pedestrian modelling (Crooks et al, 2008; Johansson and Kretz, 2012), and traffic and transportation simulation (Chen and Cheng, 2010).

2.4.3. Spatial microsimulation and agent-based hybrid models

While Merler et al. (2013) combined a stochastic microsimulation model with an agent-based model in their study of a flu outbreak in the Netherlands, very few studies have utilised a similar approach. Wu et al. (2008) represents one such rare example, demonstrating the use of combining these bottom up approaches when investigating student migration patterns in Leeds. A hybrid model was shown to be very accurate in predicting the geographical spread of students across the city, compared to a microsimulation model which failed to capture the clustering of students in certain locations. It was noted that microsimulation is driven more by statistics and probabilities, whereas agent-based models can rely on their rule based (built in intelligence) nature. However, the tried and tested techniques of microsimulation provide ‘important statistical mechanisms that ensure the similarity of what it predicts and what is actually observed in the gathered data’ (p.446), which can aid in guiding patterns of population evolution.

Wu and Birken (2012) comment that the combination of the two methods allows large scale data to be processed using list processing power, as well as the consequences of behaviours and policies at a macro scale through predefined transition rates, while also modelling interactions and behaviours of individuals. Their mortality model (again set in Leeds) demonstrated the importance of personal histories, with previous places of residence influencing health regardless of the current residence of certain individuals, showing that ABMs can ‘compliment MSM [microsimulation] by retrieving personal histories with great ease’ (p.356). The potential shown by these models begs the question as to why this approach has not been used more often.

Cajka et al. (2010) further demonstrated the benefit of combining data from a microsimulation model with an agent-based model, commenting that ‘although these models [ABMs] can simulate the realistic propagation of epidemics, they require input data about the social networks that are part of the agents’ day-to-day activities’ (p.1).

Although not modelling the spread of diseases, the paper demonstrates how personal attribute data assigned to agents from a microsimulation dataset can be used ‘to create school, workplace, and public transit interactions and then code this information into the data’ (p.1), in attempting to build realistic social networks into disease modelling.

This section has demonstrated the wide range of applications that agent-based models have in both a health and urban studies context, for both longitudinal and historical analyses. The incorporation of interactions gives the method a different focus to spatial microsimulation modelling, while also demonstrating incredible adaptability in its setup. This includes connecting with other model types, and incorporating other datasets, allowing them to ‘borrow’ the strengths of these. The ability to base models around theory (e.g. Cerda et al, 2014) also offers the exciting prospect of simulating conceptually valid and relevant scenarios. This theoretical grounding is important if work is to avoid being designated as ‘blue skies research’. Despite all this, research combining multiple models within an urban environment backed by specific theory is rare. Many studies have used hypothetical data in their analyses, so the opportunity to impute a representative population into a model using spatial microsimulation has huge potential, particularly for studies of oral health at the small area level, which remain untouched in the Dental Public Health literature.

2.5. Determinants of oral health inequalities in tooth decay

Within all geographical areas there are a variety of determinants of oral health status. Although various authors have warned against simplistic dichotomies of individual (composite) and neighbourhood level (contextual) variables, for simplicities sake the determinants of tooth decay will be presented in these categories below. Firstly, individual and household level indicators will be presented, followed by behaviours that are important to tooth decay. The final section will cover neighbourhood level determinants of tooth decay.

2.5.1. Individual and household level indicators

Age

There has been a great deal of research that has examined the relationship between age and oral health (Stahlnacke et al, 2003; Sanders and Spencer, 2004). Such research has more often been concerned with self-rated or subjective oral health outcomes (Stahlnacke et al, 2003) as well as tooth loss (Sanders and Spencer, 2004). Steele et al. (2015) have demonstrated that age groupings are critical for oral health, with effects differing at certain ages. There is no income-related relationship with the number of missing teeth in young people (rich or poor) for example, as this group tend to have complete dentitions. With age this becomes more of an issue however, and the difference between the rich and poor was nearly 8 teeth (Steele et al, 2015). For caries, effects are significant up to the age of 50 (including the protective effects of income), but disappear for older groups for whom tooth loss is a more pertinent issue (Steele et al, 2015). Data from the National Institute of Dental and Craniofacial Research (2014) in the USA would seem to confirm such patterns (Table 3).

Table 3 – Decayed, missing and filled teeth by age grouping (NHANES, 1999-2004)

Characteristic	Decayed permanent teeth (DT)	Missing permanent teeth (MT)	Filled permanent teeth (FT)	Total decayed, missing, or filled teeth (DMFT)
Age				
20 to 34 years	0.93	0.62	4.61	6.16
35 to 49 years	0.75	2.39	7.78	10.91
50 to 64 years	0.55	5.30	9.20	15.05

The data presented in Table 3 shows that as participants of the National Health and Nutrition Examination Survey (1999-2004) age, the number of decayed teeth decreases, while the number of missing teeth increases. The number of filled teeth also increases with age, and is most likely due to the lifetime effects of tooth decay.

Gender

There have been numerous studies that have investigated the links between gender and tooth decay (Lukacs and Largaespadad, 2006; Ferraro and Vieira, 2010). Antunes et al. (2006) for example demonstrated that girls had higher odds of having untreated decayed teeth in their study of 12-year-old schoolchildren in Brazil. Antunes et al. (2003a) found comparable patterns in a similarly aged group of Brazilian schoolchildren. Antunes and colleagues (2003b) suggested that this pattern could be attributed to the earlier eruption of permanent teeth in females, which would leave them exposed to cariogenic environments for longer (Lukacs and Largaespada, 2006), although not all of the literature has confirmed this pattern (Ferraro and Vieira, 2010).

Ferraro and Vieira (2010) listed a number of other potential mechanisms through which tooth decay scores are higher in women. Alongside the early eruption of teeth, other biological factors such as variations in the Amelogenin gene which contributes to enamel formation, as well as differences in saliva flow rates between women and men were identified. Saliva can help with washing, buffering and remineralising, and hormonal fluctuations in women have been identified as one mechanism that might influence salivary flow rate. Pregnancy is also listed as a potentially important mechanism, again due to hormone fluctuations inducing changes in the oral cavity environment and salivary flow rate. Other mechanisms related to the historical role of women in society have also been suggested, with Lukacs (2008) suggesting that the development of agriculture may have led to increases in fertility, while Lukacs and Largaespada (2006) listed women's historical role in food preparation and access to food and snacking as another potential cause. How valid some of these theories are to oral health remains questionable, as some of these points are less relevant in today's society. The theme of gender, while appearing in many studies, is under researched with regard to more detailed causal mechanisms.

While the literature suggests that women are more vulnerable to tooth decay, women also seem more likely to engage in recommended oral health behaviours and dentist visits, while there are also connections between 'masculine behaviour' and dental related problems (Doyal and Naidoo, 2010). A number of studies have demonstrated that females tended to have more favourable behaviours including interest in their oral health (Ostberg et al, 1999), and higher levels of dental knowledge (Tada and Hanada, 2004). Bertea et al. (2007) found similar patterns and confirmed 'the paradox that although women take more care of their teeth, their dentition is not better than men's' (p.365). However, the link between these behaviours, and the previously suggest genetic and social hypotheses remain under researched within the dental literature.

Race and ethnicity

Race and ethnicity are subtly different concepts. Race can be seen to involve physical or biological commonality which is involuntary, while ethnicity is more concerned with cultural commonality (shared practices, views and values), and can be voluntary (Morning, 2005). Within the dental literature these terms have at times been treated as synonymous. Reid et al. (2004) used both race and ethnicity in their study, and demonstrated that non-Hispanic white populations experienced fewer carious teeth than both Mexican American and non-Hispanic black populations in the United States. Similarly, Hudson et al. (2007) used both ethnicity and race, and again found that the white population in their study had lower levels of decay than the Black and Mexican-American population samples. Drummond et al. (2015) focused specifically on race in their research, yet found similar results, showing that white populations experienced less decay than those of African, East Asian, indigenous or mixed race descent. These findings are not unanimous however, as work by Delgado-Angulo et al. (2015) on decay and ethnicity demonstrates. Their research showed that Asian and Black ethnicities experienced lower DMFT scores than those of British, European and other white ethnicities in East London.

These relationships are complicated however, with some studies finding that adjusting for certain socio-economic factors accounted for the difference between groups. Reid et al. (2004) for example found that the addition of material factors (i.e. income, education, employment, dental insurance, area of residence) explained most of the variation in decay between groups in the USA, while behavioural variables had little effect. Drummond et al. (2015) also found that education and income fully explained differences in tooth decay between white populations and populations of African descent in Brazil, while accounting for most of the variation between white populations and those of mixed race. Watt and Sheiham (1999) have questioned the relevance of including ethnicity as a variable in Dental Public Health studies, stating that 'it could divert attention from more important variables such as income and social class' (p.8).

Conversely, a study by Delgado-Angulo et al. (2015) in East London found that Asian and Black populations had better oral health than white populations, and that 'the association between ethnicity and caries experience was independent and not confounded by time lived in the UK or socioeconomic measures' (p.58). Hudson et al. (2007) used existing literature to test hypotheses based on pathways, and also found that ethnic and racial differences in decayed surfaces and missing teeth could not be fully explained by socio-economic differences, or variations in health-related behaviours or annual dental care. The context of the studies and measurements used may have played a part in the disparate findings.

Other trends within the dental literature point towards disadvantage for minority groups. Julihn et al. (2006) found that adolescents with foreign born parents were more likely to have high DMFT scores than adolescents with native born parents, indicating this may be a risk factor, despite the lack of clear mechanisms. Race has also been associated with dental treatment, as Black and Hispanic children in urban areas were more likely to receive dental treatment than white children, while the opposite pattern was seen for non-white children in rural areas (McKernan et al, 2015). Fonseca (2012) found that neighbourhoods where African American ethnicities resided contained only around half the supermarkets present in white populated areas, and that low income African American populations tended to have less nutritionally rich diets.

Importantly, ethnicity is difficult to evaluate due to the number of categories included under this term, more so than most variables. Butani et al. (2008) have outlined oral health-related cultural behaviours and their histories for some ethnic groups; however, with so many different ethnicities it is difficult to understand mechanisms regarding oral health for each one. This remains an under researched area within the dental literature.

Income

Relative income inequalities have been shown to be detrimental to health in general (Wilkinson and Pickett, 2006) as well as tooth decay (Costa et al, 2012; Schwendicke et al, 2015). As such, a number of oral health related studies have investigated income inequalities using the Gini coefficient (Gini, 1912), with Celeste et al. (2009) showing that higher municipal level income inequalities were associated with missing teeth, edentulism and dental caries in Brazil. Pattussi et al. (2001) have also used the Gini coefficient to demonstrate higher prevalence and severity of dental caries in children in areas with higher income inequality. This tallies with Wilkinson's (1996) theory that more egalitarian, rather than the richest societies, tend to have the best population health.

Work by Celeste et al. (2011) came to similar conclusions to Celeste et al. (2009), while Aida et al. (2008) showed that average income exerted a relatively large influence on variance in DMFT scores at community level in Japan. Geyer et al. (2010) also found levels of tooth decay to be higher within the lowest income brackets, while a systematic review by Costa et al. (2012) found individual's income to be significantly associated with greater caries occurrence. Bhandari et al. (2015) suggest three plausible pathways between income inequality and health: disinvestment in public services, with less spent on social goods in unequal societies; social capital erosion, which creates less supportive climates for policies that maintain health; and psychosocial effects of stress through social comparisons. The authors show service use to be negatively associated with income inequality, while countries with greater income inequality had lower spending on health services and fewer dentists per population, supporting the first suggested pathway.

Beyond the wider societal effects, income inequality has also been shown to influence individual and household access to various useful resources. A systematic review by Costa et al. (2012) revealed that higher income was associated with access to dental services, better access to fluoridated water, to products that maintain good oral health, and better levels of information about oral health. Patterns regarding tooth brushing frequency were however less consistent in a study by Peltzer and Pengpid (2014), where students from more affluent backgrounds were less likely to brush twice a day. Fonseca (2012) has commented on prejudices some low-income children may receive from dentists in the United States, as well as the risk of 'food instability' for those on low incomes, or below the poverty line. Children in low income families are also more at risk of missing more school days through dental related issues (Fonseca, 2012).

Income is also an important marker of material living standards. Galobardes et al. (2004) summarised housing tenure, conditions within homes, and household amenities as the main markers of material circumstances, stating that 'housing is generally the key component of most people's wealth, and accounts for a large proportion of the outgoings from income' (p.9), and can act as both a direct and indirect marker for exposures. Wilkinson (1997) included additional variables alongside housing in his article, including diet, inadequate heating, and air pollution. All of these issues are inextricably linked to income, through which material standards impact on oral health in a number of ways. Within the UK, O'Hanlon et al. (1997) showed graded increases in oral cancers between less and more materially deprived wards, as measured by the Townsend Deprivation Index (Townsend et al, 1988). Nicolau et al. (2005) further demonstrated the importance of such measures, showing that Brazilian adolescents who experienced poorer material conditions at birth and age 13 were more likely to experience high dmft scores. Section 2.5.2 will also demonstrate how income affects other material resources such as diet, as well as dental knowledge and resources.

Education

Perhaps unsurprisingly, numerous studies have investigated the links between education and tooth decay (Schwendicke et al, 2015), and levels of educational attainment have been shown to be related to a number of oral health outcomes at various stages of life. For example, Muirhead and Marcenes (2004) demonstrated that mean school scores in linguistic awareness, English and mathematics were strongly associated with school mean dmft scores in a sample of primary school children in Wandsworth, London. Conversely, caries has been shown to directly impact on children's education through time taken off school to deal with the disease, and through restricted activity within school (Sheiham, 2006), while sleep loss resulting from painful teeth can negatively affect mood, attention span and interactivity (Casamassimo et al, 2010; Alameda County Oral Health Needs Assessment, 2006).

The importance of education to future prospects has also been demonstrated (Kuh and Wadsworth, 1991), where lower educational attainment could see individuals set on trajectories that are detrimental to oral health. A number of studies have shown that educational disparities have a persisting influence on dental caries experience in adult and later life, including studies from Australia (Brennan et al, 2007), Greece (Mamai-Homata et al, 2012) and Germany (Geyer et al, 2010). A systematic review by Costa et al. (2010) also supported this trend, with statistically significant associations with caries being found in 6 out of the 9 studies included in the review.

Geyer et al. (2010) have stated that education may 'point to differences in knowledge on how to practice dental hygiene and how to prevent dental decay' (p.124). Schwendicke et al. (2015) have also commented that education may act as a mediating pathway between caries and socio-economic position, through impacting on income and related access to various preventative oral health measures. A second suggested pathway asserts that education may also affect characteristics such as health literacy and behaviours, which are key to oral health outcomes. Sabbah et al. (2015) found educational gradients in tooth decay in a Finnish population, with scores becoming successively higher each time education dropped a band, while educational gradients also existed in the oral health behaviours included in the study. Sabbah and colleagues (2015) state that 'this observation highlights the importance of the contribution of formal education as a marker of socioeconomic position to health and dental behaviours' (p.38). The authors suggest that better education may allow people to benefit more from health promoting messages and policies, as well as allowing more flexible lifestyles that allow for the adoption of certain behaviours. Similar educational gradients in health compromising behaviours were also found in a sample of British adults (Singh et al, 2013).

Employment

Similar to education, there have been a number of studies that have analysed links between employment and outcomes in tooth decay (Costa et al, 2012). The aforementioned systematic review by Costa et al. (2012) showed occupations of higher standing to be linked with lower severity of dental caries in adults. Further evidence of the importance of employment to caries development is provided by the work of Roberts-Thomson and Stewart (2008), who found that being unemployed or in receipt of government benefits were risk factors for decayed surfaces and total caries experience respectively. Similarly, Tellez et al. (2006) found that unemployed caregivers had 2.5 times more untreated decayed surfaces than those who were employed. The evidence is not unanimous however, as Julihn et al. (2006) demonstrated that parental occupation (as well as education level) was not found to be a significant risk factor among 19-year-olds in Sweden. This is in contrast to other literature, which has demonstrated the importance of parental occupation to children's oral health, specifically caries (Vanobberge et al, 2001; Gokhale and Nuvvula, 2016). It would seem though that unemployment, more-so than type of employment, is a key risk factor for tooth decay.

Unemployment has been associated with less beneficial oral health related behaviours (Al-Sudani et al, 2016), including less frequent use of dental services and xylitol, while income and education had a mediating effect on tooth-brushing frequency, sugar use and alcohol consumption. Similar patterns were demonstrated by Guiney et al. (2011), who found that employment, particularly for males, was associated with dental service utilisation in adults from the Republic of Ireland. Al-Sudani et al. (2016) also demonstrated that unemployment was associated with increases in smoking and drinking, which, while not as clear cut in their relationship with tooth decay, are important to oral health more generally.

Socio-economic position

Socio-economic position (SEP) refers to social and economic factors that influence the positions groups and individuals hold within society, and has been related to resources, exposures and susceptibilities that affect health (Galobardes et al, 2004). There is no best measure of socio-economic position, and variables that affect health may differ and also overlap. Socio-economic position is distinct from socio-economic status (SES), as the former is more concerned with actual resources, rather than the status, rank and prestige base of the latter (Krieger et al, 1997). Within the dental literature some studies have used social class classifications (Watt and Sheiham, 1999), while others represented SES using employment data (Thomson et al, 2004), and SEP through resource based variables (Hobdell et al, 2003). This shows the lack of consistency in representing measures of social stratification. Despite not all measuring socio-economic position in the way defined above, all of these studies have been included as they analysed concepts related to this measure in some way.

Socio-economic position plays a significant role in shaping an individual's opportunities and life trajectories. Socio-economic position, and particularly one's position on the social gradient, is also known to influence the health of individuals (Marmot et al, 2010), and the concept has been used as a central starting point for the World Health Organisation's Conceptual Framework on the Social Determinants of Health (WHO, 2010). As such, a number of studies within Dental Public Health have analysed the effect of socio-economic positions on tooth decay (Schwendicke et al, 2015). For example, Watt and Sheiham (1999) have shown that children from lower socio-economic backgrounds are more likely to suffer from high levels of caries, with inequalities in adult life present but less marked. A systematic review by Schwendicke et al. (2015) also showed lower socio-economic positions to be associated with a higher risk of caries experience, or having carious lesions. Contrary to expectations, the authors 'found caries experience to be more unequally distributed in highly developed countries with low income inequality than in underdeveloped, more unequal countries' (p.15).

Hobdell et al. (2003) found gradients in three oral health diseases, including dental caries, related to socio-economic status, confirming that 'poor socio-economic circumstances' (p.95) adversely affect oral health. The authors draw interesting conclusions, stating that it is not merely individual issues, but structural level factors that lead to such patterns, and need to be addressed. In contrast, a systematic review by Costa et al. (2012) found considerable variance in the results of SES related studies investigating caries, with the authors commenting that the way socio-economic status was classified was subjective and variable between the studies. The importance of socio-economic positions has been demonstrated by Dalstra et al. (2005) who state that 'research to date has generally concluded that socio-economic differences persist into the latest stages of life' (p.2047). The effects of this can be seen in the longitudinal research of Thomson et al. (2004), which found sizeable differences in socio-economic status emerging in the study's cohort when measuring tooth loss from dental caries. The authors suggest that 'early socioeconomic inequalities in a number of important oral health indicators do persist well into the third decade of life' (p.351).

Psychological factors

Dalstra et al. (2005) state that relative socio-economic position is known to have psychosocial consequences, which in turn can impact on oral health outcomes. Situations causing stress on a regular basis have been shown to be more prevalent among those in lower socio-economic positions (Taylor and Seeman, 1999), and can trigger the 'fight or flight' mechanism (Cannon, 1929) that suppresses a number of biological processes, including the immune system, which can be harmful if activated too often or for prolonged periods of time. Psychological stresses are known to be detrimental to general health in a direct way (Marmot et al, 1991), and have also been shown to negatively influence oral health. More often studies have demonstrated links between psychological stress and self-rated oral health. For example, Finlayson et al. (2010) demonstrated links between psychosocial stressors and resources, and reporting of poor self-reported oral health. Sanders and Spencer (2005) found associations between perceptions of self-rated oral health and psychosocial variables such as personal constraints and chronic stress in low income groups in Detroit, Michigan, hypothesising that poorer material and social conditions constrain access to dental services and health promoting messages. There have also been a number of studies investigating the links between psychological stress and periodontal conditions (Warren et al, 2014).

Fewer studies have made links between stress and tooth decay, however those that have support the link between the two. For example, stress has been associated with increased levels of dental caries through biological factors (Belstrom et al, 2014), for instance in noxious family environments (Lorber et al, 2014). Boyce et al. (2010) examined how these stresses were linked to caries development in childhood, and found that families from lower socio-economic positions were more likely to experience financial stresses. This was associated with increased levels of cortisol secretion in the mouths of affected children, negatively influencing both enamel thickness and density. The literature on psychosocial impacts is not unanimous however, demonstrated by Armfield et al. (2013), who found no significant relationships between psychosocial variables and oral health.

2.5.2 Behaviours

Diet

Given the obvious links between diet and oral health it is perhaps not surprising that there have been a multitude of studies on this subject (Moynihan and Petersen, 2004), as well as national level datasets, with many of these pointing to a social gradient in diet. For example, nutritional intake has been shown to differ between high and low income groups, with those on lower incomes tending to have poorer diets (National Diet and Nutrition Survey, 2014). A study by Barker et al. (2008) demonstrated that women in lower educational groups felt less able to change family diets due to financial pressures, and lack of support. A combination of a lack of confidence in cooking skills, and a limited range of food at a younger age led many to experience a perceived lack of control over their food choices. A review by Mobley et al. (2009) concluded that areas of lower socio-economic status were more likely to be served by stores 'offering high-energy, low-nutrient–dense foods' (p.412), which lack appropriate nutrients to aid oral health, demonstrating that environmental factors are also important to diet. In line with this, Fonseca (2012) states that lower income areas tended to be served by fewer large chain supermarkets which tend to offer a greater variety of produce, while high food prices and travel distances presented further barriers to such populations.

Sheiham (2006) has commented on the susceptibility of undernourished children to dental caries, while Moynihan and Petersen (2004) state that dental decay leading to tooth loss is associated with diets that are low in fruits, vegetables, non-starch polysaccharides and vitamins A and D. Increases in sugar consumption combined with under-nutrition can lead to larger than expected levels of caries. The link between sugar consumption and tooth decay is well established, with the amount and frequency of sugar intake playing an important role (Sheiham, 2001). Due to this there is considerable interest in the increase in soft and fizzy drink consumption. High sugar intake has been linked to increased caries incidence in low income adults with increased soft drink consumption (Burt et al, 2006), while Warren et al. (2009) demonstrated similar patterns in a longitudinal study of pre-school children from lower socio-economic backgrounds. The importance of a healthy diet is reinforced by a number of studies that have found that metabolic syndrome may be associated with caries and decayed teeth (Timonen et al, 2010; Ojima et al, 2015), particularly in the case of obesity, hypertension, dyslipidaemia and hyperglycaemia in the latter study.

Coping mechanisms

Continued psychosocial pressure can lead to some individuals employing coping mechanisms to alleviate this stress, and a number of studies have assessed the links between smoking and tobacco use and tooth decay (Reibel, 2003). Previous research has demonstrated that links exist between stress levels and smoking in adolescents (Kassel et al, 2003) and adults (Ng and Jeffery, 2003), as well as being associated with financial pressures (Siahpush and Carlin, 2005). While associations have been found between smoking and caries (Hudson et al, 2007; Bernabe et al, 2014), some studies have reported that no direct aetiological link can be found between the two (Reibel, 2003; Vellappally et al, 2007). Despite this, it would still seem that smoking is a risk factor for dental caries (Axelsson et al, 1998; Reibel, 2003), and that certain chemical processes resulting from smoking may make teeth more susceptible to the disease (Vellappally et al, 2007). Locker (1992) reports that smoking among older adults (aged 50 plus) was associated with edentulism, tooth loss, caries and periodontal disease, while Reibel's (2003) review of smoking and tobacco use showed that smokers had a significantly higher risk of developing oral cancers, and more severe periodontal disease than non-smokers. Smoking and excessive alcohol consumption in combination also had negative effects for oral health. Winn (2001) has also suggested that chewing tobacco may be harmful with regards to tooth decay, due to the high proportions of fermentable sugar found in these products.

In line with this, a study by Franceschi et al. (1999) demonstrated that alcohol consumption increased the risk of oral cancer 'in each stratum of smoking, never-, and ex-smokers' (p.2), with similar patterns noted for smoking increases in alcohol consumers. While smoking has been seen as the greater threat to oral health, less attention has been paid to alcohol consumption within the Dental Public Health literature. Harris et al. (1997) demonstrated that while previous oral health conditions showed no positive correlations with level of alcohol consumption, a third of the study participants exhibited tooth wear correlated with consumption, most likely due to gastric acid regurgitation associated with alcohol gastritis. Higher levels of dental trauma were also positively associated with alcohol consumption. However, Harris et al. (1996) report that oral health in a group of alcoholics was not compromised.

Dental health habits

Due to the importance of oral health related habits, particularly the use of fluoridated dentifrice, for tooth decay (Marinho et al, 2003; Zero, 2006), a number of studies have investigated behaviours associated with this, and how these are produced. For example, Poutanen et al. (2006) found associations between oral health knowledge of both parents and children, and the child's associated dental behaviours. While parental attitudes were not shown to be associated with their child's oral health practices, their behaviours were, supporting the idea that children learn from their parent's actions. Results from Adair et al. (2004) support this, the authors finding that parental attitudes and efficacy towards oral health behaviours influenced behaviours in their own children. More deprived families exhibited less confidence in being able to control such patterns. Sabbah et al. (2009) found similar links between socio-economic indicators and oral health related behaviours in a representative sample of adults from the United States, where behaviours were poorer among those with lower levels of income and education.

Williams et al. (2002) reinforced the idea of socio-economic disparities in attitudes towards oral health, finding that higher proportions of parents from non-deprived areas had higher levels of dental knowledge, while having a higher level of education was also shown to be associated with a higher level of dental knowledge and a better attitude towards dental care. This echoes conclusions from Singh et al. (2013), who found educational gradients throughout patterns of clustering of health behaviours. The work of Riley et al. (2006) shows that education is a key feature of such clustering, with better oral health and attitudes to attendance as a result of better education. However, the work of Sanders et al. (2006) 'disputes the notion that poorer adults care less about their oral health than the more affluent ones' (p.74), as while dental visiting patterns followed a social gradient, dental self-care trends did not, highlighting the wide variation in oral health behaviours.

Chu et al. (1999) demonstrated the benefit of positive attitudes towards oral health, showing that children of parents with higher dental knowledge experienced lower levels of dental caries, possibly due to the instillation of positive oral health behaviours. This highlights the importance of parental involvement in oral health promotion, and other studies have also emphasised the importance of parental characteristics and variables in the development (or lack thereof) of caries in children (Hooley et al, 2012). Additionally, Kumar et al. (2016) have demonstrated the importance of tooth brushing for good oral health, with infrequent brushers shown to have higher incidence and increment of carious lesions.

Attendance at services

There have been a large number of studies that have investigated the effects of dental attendance on oral health, as reported in a systematic review by Davenport and colleagues (2003). Despite the aforementioned systematic review finding no (high quality) evidence to refute or support the practice of six-monthly check-ups, longitudinal research has demonstrated that long term routine dental attendance is beneficial with regard to tooth decay, as well as tooth loss and self-rated oral health (Thomson et al, 2010). A number of other studies have assessed patterns and trends in dental attendance.

For example, Socio-economic position is known to affect patterns of attendance at dental services, and Lang et al. (2008) found that older adults living in more deprived areas only attempted to use dental services when symptomatic, compared to the more regular attendance of those in higher social classes, and/or those who were better educated. Irregular attendees have tended to experience worse oral health than regular attendees. Tickle et al. (1999) found significant differences between such groups with regard to dmft scores in a sample of 5-year-old children in the North-West of England. Irregular attenders were more likely to be from deprived backgrounds, and less likely to receive treatment to fix problems. Other research has shown that while regular attenders have lower levels of decay, they also have higher numbers of fillings and treatments (Sheiham et al, 1985).

Tickle et al. (2000) found a highly inverse relationship between ward level dmft scores and rates of 3-5 year olds in contact with primary dental care, suggesting that service usage increases with decreasing disease experience. Further research in this area by Eckersley and Blinkhorn (2001) demonstrated that children from deprived wards were less likely to have attended a dentist than those from less deprived wards, while also being more likely to attend solely when symptomatic, echoing the findings of Lang et al. (2008). These children were also more likely to start brushing later in life and on a less frequent bases, however parents from more deprived backgrounds were also more inclined to change oral health related habits and behaviours upon receiving advice than participants from less disadvantaged wards. Tickle et al. (2003) found that the treatment previously received by children was the most influential factor for healthcare preferences of parents, as they gained knowledge and confidence in certain treatments. This was particularly the case for parents from deprived backgrounds.

Listl (2012) demonstrates the importance of dental attendance from an early age, as ‘a considerable proportion of inequalities in regular dental attendance is already established in childhood and persists throughout the life-course’ (p.96), rendering future interventions potentially futile. Work by Wallace and Macentee (2012) shows that low-income groups can face considerable barriers to dental services through their opinions on dentists, and through some dentist's attitudes to low income groups regarding the limits of private practice and public dental benefits. Other studies have also found that favourable pre-existing socio-economic and demographic characteristics are generally associated with regular dental attendance (Guiney et al, 2011; Muirhead et al, 2009), while evidence from the Adult Dental Health Survey (University of Essex, 2009) suggests that attendance patterns are associated with other oral health preventive behaviours (Hill et al, 2013). Donaldson et al. (2008) summarise the importance of attendance, stating that 'the socio-economic gradient in the number of sound teeth in adults is partially explained by dental attendance, which in turn is determined by the effect of SES on barriers to regular dental attendance' (p.63).

2.5.3. Neighbourhood based indictors

Water fluoridation

While physical aspects of the local environment are under researched in the dental literature, none may be more influential than the presence of fluoridated water. This has been recognised by bodies such as the World Health Organisation (WHO, 2003) as having the potential to reduce levels of tooth decay in both children and adults. Desai et al. (2015) explain that in acidic environments fluoride promotes remineralisation of tooth surfaces via the use of phosphates and calcium ions in saliva, helping to reduce the loss of tooth substance. The antimicrobial effects of fluoride also reduce the cariogenic potential of bacteria.

Given the perceived importance of water fluoridation to oral health it is not surprising that a number of different studies and systematic reviews have been conducted on the subject. Indeed, analysis of systematic reviews concerning water fluoridation (excluding professionally applied fluorides) by Petersen and Lennon (2004) revealed that water fluoridation reduced levels of dental caries in both relative and absolute terms, while fluoridated toothpastes and mouthwashes reduced decayed, missing and filled surfaces (dmfs) by between 24% and 26%. Further systematic reviews include the work of McDonagh et al. (2000) whose review again found that fluoridated water led to reductions in dental caries, although this reduction was seemingly lower than had been reported in the past. Perhaps the most worrying conclusion from the review was the low overall quality of the evidence on the topic. Yeung's (2008) systematic review drew similar conclusions to previous studies, showing that areas with fluoridated water experienced lower levels of dental caries than areas without a supply. All three reviews found dental fluorosis to be associated with the presence of fluoridated water, although in all cases this was mild and led to no other health concerns. While the previous studies focused on children, Griffin et al. (2007) demonstrated the benefits of fluoridated water to adults, showing that caries was consistently higher in groups exposed to non-fluoridated water, and that ‘exposure to any mode of fluoride reduced caries by about 25%’ (p.414).

McGrady et al. (2012) have demonstrated how fluoridated water (along with fluoridated dentifrice) can decrease the social gradient in oral health inequalities, in their study of dmft scores in Newcastle and Manchester. Similar findings emerged from a study by Slade et al. (1996), who showed that inequalities in the experience of dental caries was greater in children who lacked exposure to fluoride in drinking water. In the UK fluoridated water is far from universally distributed, with The British Fluoridation Society (2012) estimating that only around 6,000,000 people were exposed to fluorides, of which only around 300,000 had access to naturally fluoridated water.

Social capital

Social capital has become an increasingly important concept when analysing health inequalities, and there have been a number of studies that have investigated links between this variable and oral health (Rouxel et al, 2014). Lida and Rozier (2013) found that social capital was associated with children’s use of dental care, and argued that social capital is important to health in four ways: through the diffusion of knowledge of health promotion; maintenance of healthy behavioural norms; promotion of access to local service and amenities; and psychosocial processes providing support, while building esteem and mutual respect. Turrell et al. (2007) have also hypothesised the effects of social capital on oral health, through extensive social webs leading to more active social engagement through community organisations, and through influencing health behaviours through norms and appropriate practices. Social capital has also been shown to be important for the oral health of elderly citizens, and Aida et al. (2011) used multi-level modelling to demonstrate a significant association between contextual level social capital and self-rated oral health.

Specific examples have also been found in relation to tooth decay, with Aida et al. (2008) finding that community context, particularly community centres and grocery stores were significantly associated with dmft, potentially through chances to increase social cohesion and solidarity. Similarly, Tellez et al. (2006) found relationships between social institutions such as churches and improved oral health. This study is notable for its consideration of multiple built neighbourhood environmental features, while also discussing area size and population as well as using theory for variable selection. In Brazil, Santiago et al. (2014) found variation in neighbourhood level caries experience to be related to perceived levels of empowerment in adults, while individual level social capital was not associated with the disease. Avlund et al. (2003) have also commented on associations between people who either lived alone, or were not happy with their social contacts and higher levels of caries.

Not all studies have found associations between oral health and social capital however (Mathur et al, 2016), and a commentary by Rouxel et al. (2014) found mixed results, although still concluded that despite the underdeveloped literature, social capital appears to be a potential determinant of oral health. The problem with the definitions of social capital that were used may not have helped this literature overall (Rouxel et al, 2014).

Access to services

Given the importance of attending dental services, it is perhaps not surprising that a number of studies have focused on patient access to such services (Hill et al, 2013). For example, Lang et al. (2008) have stated that differences in attendance patterns may not be helped due to certain areas being underserved by dental services, and historically there has been an uneven distribution of services within the UK (Cook and Walker, 1967). This sentiment has been echoed in other studies. Jones (2001) conducted statistical analysis on area level deprivation scores and NHS dental registration records in England, with their conclusions supporting ‘a working hypothesis that there is an inverse ‘dental’ care law for children in England’ (p.205). Landes and Jardin (2010) comment that up until the dental contract changes of 2006, practices were free to set up where they wished. They called for investment in new practices, or existing ones, after demonstrating that practices in deprived areas in County Durham tended to be smaller, yet still served higher proportions of the deprived community. However, Macintyre et al. (2008) included dental services in a list that ‘showed no clear pattern by deprivation’ (p.911) in their study of resource distribution in Glasgow.

The type of services available (i.e. NHS or private) is also an important consideration. Hancock et al. (1999) have demonstrated that residents in the south of England had more difficulty accessing NHS dentists, due to the rise in private dentistry in this region, leading to increased dissatisfaction with NHS services. Such geographical findings were echoed in the work of McGrath et al. (2001). Hancock and colleagues (1999) concluded this shift was not something the public had demanded, while McGrath and colleagues (2001) suggested some were paying for private treatment when they did not want it. Propper (2000) found that use of private healthcare in the UK was strongly determined by income, with richer individuals and households being more likely to use such services. Despite this, analysis of the 2009 Adult Dental Health Survey by Hill et al. (2013) demonstrated that the NHS was still the majority provider for dentate adults.

Area of residence and features of the built environment

Watt and Sheiham (1999) have stated that ‘area-based indicators are better predictors of oral health status than measures of socio-economic status’ (p.8). A number of studies have used area based measures to assess variations in dental health using the Townsend Deprivation Index (Prendergast et al, 1997), the Jarmin Index (Jones et al, 1997) as well as the Indices of Multiple Deprivation in England (Broomhead et al, 2014), Scotland (McMahon et al, 2010) and Wales (Richards et al, 2005). Such measures are perhaps more effective due to the number of different variables included within these composite indicators.

However, features of the built environment have in general not been covered in as much depth. Dental surgeries are an obvious example that have been given sufficient attention, however the literature on other neighbourhood features is at times lacking. Despite this, there are examples which have studied such features within the literature. As mentioned earlier, Mobley et al. (2009) and Fonseca (2012) both hypothesised the importance of shops and supermarkets for diet and oral health. Aida et al. (2008) found significant associations between the number of grocery stores per resident and dmft scores, while Tellez et al. (2006) demonstrated similar trends with cavitated carious lesions. Borenstein et al. (2013) did not find associations between supermarket locations and self-rated oral health, however this research represents a rare approach within the Dental Public Health literature, due to the inclusion of multiple features of the built environment, such as fast food restaurants, the number of social services, and the number of parks. Tellez et al. (2006) also included churches as context specific environmental features in their study, again finding significant associations with cavitated carious lesions, while Aida et al. (2008) found significant associations with community centres. Finlayson et al. (2010) included '7 different neighbourhood resources', yet curiously only made mention of public libraries, medical services, parks and banks in their paper. These variables related to individual responses on how close participants felt they lived to these services, as opposed to their exact locations. Although no significant relationships were found, the research found that living closer to an increased number of services was associated with lower levels of poor or fair oral health, although this was dealt with as a homogenous 'services' variable.

Antunes et al. (2002) included household overcrowding as a demographic variable in their study, while Moyses et al. (2008) included 'areas of substandard living conditions', defined as a 'gathering of poor dwellings without access to urban planning policies, characterized as of greater socio-environmental risk with regard to human health' (p.450). This slightly vague definition is hard to unpack exact relationships from though. Carvalho et al. (2010) included 'the number of urban social facilities in the Health District' (p.3), and while these represent built environmental features, this measure again grouped all potential facilities into one measure, rather than using exact geographical locations. Freire et al. (2013) also included the supply of fluoridated water in Brazilian towns in their research, a variable used numerous times in this context (Antunes et al, 2002; Antunes et al, 2004).

Despite the inclusion of these features, there has been a lack of attention paid to the built environmental features of neighbourhoods overall, and this is an area that requires further research. Some variables have used aggregated individual level data to create neighbourhood level variables, such as empowerment (Santiago et al, 2014), while others have used individual perceptions of environmental features (Finlayson et al, 2010). There is nothing wrong with these approaches, but as they are divisible to the individual level (Cummins et al, 2007), they may not be capturing the exact geographical nature of neighbourhoods. This process may be harder to judge using statistical models, and may be an area where a method such as agent-based modelling would be useful.

What becomes clear is that the determinants of tooth decay are multiple, interlinked and complicated in nature. Certain individuals, families and groups can become stuck in a vicious cycle due to their adverse circumstances, which have a number of knock-on effects which worsen the state of their oral health. The type of neighbourhoods that people live in may also be important in determining the severity of these effects, however while features such as fluoridated water may mitigate these effects to a certain degree, they are unlikely to fully overcome social and economic forces that lead to oral health inequalities. While the mechanisms and determinants of tooth decay have become that bit clearer, due to their complicated nature it would be hard to test these without some kind of framework to adhere to, and map variables onto in order to make sense of them and their associated pathways. This will be covered in more detail in Chapter 3, which will detail the theoretical side of this research.

2.6. Spatial inequalities in oral health

While Section 2.5 outlined some of the key determinants of tooth decay, a number of studies have attempted to measure the spatial impacts of these determinants on oral health outcomes. Given the geographical focus of this research, this section will therefore review previous uses of spatial analysis within Dental Public Health, to give an overview of the state of the field. This section will then move on to look at hierarchical spatial studies, before finally reviewing simulation methods used to analyse Dental Public Health related scenarios.

2.6.1. Non GIS based studies of oral health

Some studies of oral health have focused on specific neighbourhoods, areas within cities, and regions without the use of GIS analysis, often through the use of aggregated statistics. Antunes et al. (2003b) used this approach, along with K-means cluster analysis to demonstrate ethnic disparities in both dmft scores and service usage in schoolchildren in Sao Paulo, in addition to general income inequalities and private service availability. McDonald and Conde (2010) utilised aggregated data from the

Canadian Community Health Survey, along with basic urban-rural classifications, to analyse inter-regional and urban-rural patterns of health service usage among older Canadians, demonstrating a lower uptake of services in rural areas. Antunes et al. (2004) used 'space as an organising frame' (p.41) to demonstrate the positive effects of higher socio-economic status for caries in schoolchildren in Sao Paulo, while water fluoridation was shown to be effective in reducing the disease, although this resource was also unequal regarding access. The authors cite Hart's (1971) inverse care law, and acknowledge the potential for violating the ecological fallacy by applying their findings below the city scale. Moyses et al. (2006) combined a micro-level ecological study with cross-sectional individual level data to show that dental trauma rates were highest in the east of the Brazilian city of Curitiba, with the 'physical environment' and 'public social policies' both significantly negatively associated. This is a rare example within the place based oral health literature that goes into such analytical depth when analysing individuals and areas, and including geographical considerations.

2.6.2. Early use of spatial analysis in dentistry

Spatial analysis in dentistry began in the 1960s, when Walker et al. (1965a) used concentric circles to compare regional proportions of dentists and payments to dentists in the UK. Further work by Walker et al. (1965b) created descriptive classifications of areas based on dentist numbers, money spent on treatment, and payments per dentist. Little geographical patterning was shown, although the need for more dental practitioners in the UK overall was stressed. Cook and Walker's (1967) study highlighted similar trends to the previous studies, again using concentric circles, this time at the smaller district level. Unsurprisingly disparities between London and parts of the South and the rest of the UK were found, and higher socio-economic classes were shown to receive more treatment. However, the authors acknowledged the effects of district averages in potentially concealing local variation. Further geographical analysis came from Coates and Rawstron (1971) who used similar methods, and commented on the high numbers of councils with below average dentist-to-population ratios, the tendency for many dental students to stay in their region of study, and huge variations in the distribution of school dentists in England and Scotland.

Bradley et al. (1978) present the first study to go 'beyond visual appraisal of maps and tabulations at a rather gross scale' (p.529). Their study of 5-year-old children in

Newcastle used dental data from schools, indicators of service demand and access, and travel times to show that demand for school dental services was highest where dental health was poorest. Significant relationships between working class areas and access to school dental services were also found. Taylor and Carmichael (1980) furthered this work, introducing thiessen (voronoi) polygons for catchment areas for school dental services, representing the first small areas analysis in dentistry. Using similar data, social class was shown to positively correlate with dental health and treatment need, while demand for community services was found to be higher in working class areas.

2.6.3. More recent research on spatial analysis of access to dental services

GIS and spatial analysis techniques have increased in scope since Taylor and Carmichael's (1980) work, and now cover a worldwide dental literature. Themes such as access to dental services have emerged as popular topics within the literature, and in the Australian context in particular. McGuire et al. (2011) used concentric circles to estimate patient distances to dental clinics in Victoria, while also demonstrating positive relationships between increased emergency service usage and areas of lower socio-economic status. Rocha et al. (2013) identified similar patterns in Melbourne, using geocoded patient data to show that those living over 10 kilometres from the Royal Dental Hospital experienced poorer access, while deprivation increased with distance from the hospital. Similar patterns were found in a study of public and private dentist locations in New South Wales (Willie-Stephens et al, 2014), demonstrating per capita income to be a good indicator of private dental service location, with comparable trends also seen in New Zealand (Kruger et al, 2012).

Kruger et al. (2011) used various geographical scales overlaid with graticule squares, combined with concentric circles, to show that 345,000 Western Australians fell outside 2 kilometre service buffers added to dental practices, with large disparities between urban and rural areas in practices per graticule. When compared to Queensland and Victoria, Western Australia was shown to have the fewest Census districts within 200 kilometres of a public dental clinic (Perera et al, 2010), while deprivation scores were also higher in these excluded populations. In line with this, Madan et al. (2010) demonstrated a trend for higher rates of public sector general anaesthetic in remote areas, compared to higher private rates in more accessible areas of the state which also correlated with 'better' socio-economic profiles. Again the findings suggested the need to focus on more disadvantaged and remote communities. Similar methods to those above have also been used in Australia to examine water fluoridation coverage among schoolchildren (Desai et al, 2015). The use of buffers linked to socio-economic data demonstrated links between socio-economic deciles and the proportion of children in non-optimally fluoridated schools, as well as disparities between urban and rural areas in fluoride coverage. Given the vast space services in this country must cover, it is unsurprising that research has taken this focus. However, these techniques are often descriptive and offer little explanation as to why certain patterns occur.

Transportation can also act as a significant barrier to services. Borrell et al. (2006) used subway lines to demonstrate disparities in the locations of dental care for elderly residents of ethnic minorities in New York City. This showed that residents in North Manhattan may have been forced to use Columbia University's dental clinics due to being negatively affected by transport limitations. Rail transportation formed the basis of work by Zainab et al. (2015), which used buffers, geo-coded locations and socio-economic data to demonstrate that, despite the city's excellent urban transport planning, retirees and elders in Sydney had lower accessibility to train networks capable of getting them to the Sydney Dental Hospital. This study is typical of many within this literature, in that buffers are added to geo-located points, then compared with measures of socio-economic or demographic status usually taken from administrative data or surveys.

Contrary to the results of previous studies measuring distance to services, Dumas and Polk (2015) demonstrated that distance to dental providers was not sufficient in explaining barriers to services in urban areas when investigating dental clinic utilisation among Medicaid insured children in Pittsburgh, with children often travelling further than needed rather than visiting their closest dentist. A state-wide study from Iowa however suggests that Medicaid insured children faced more serious travel burdens than those with private insurance (McKernan et al, 2016), living further from their nearest dentist, and having to drive further to access their current service. Lower rates of 'bypass' were also seen in Medicaid and rural populations, most likely due to travel practicalities and resources. McKernan et al. (2015) identified dental service areas in the same state using small area analysis of patient origins and destination data, using this to create weighted dental visits per zip code for each dentist. Their analysis using these service areas again demonstrated urban-rural disparities, as well as racial disparities in geographical accessibility. The urban-rural disparities demonstrated in this study are typical of the North American context (Emami et al, 2016). Nasseh et al. (2017) used road networks instead of straight line distance in combination with travel times, and a two-step floating catchment area method to define counties with dental shortages. This demonstrated that higher proportions of publicly insured children lived in service deficient areas in Missouri than Wisconsin, despite significant variation in both states.

Despite the descriptive nature of some of the literature, novel geographical methods have also been used in spatial studies to analyse dental related trends. Feng et al. (2016) used geographically weighted regression, local indicators of spatial association, and spatial autocorrelation analysis to demonstrate a lack of overall association between the size of the dental workforce with the varying utilisation rates seen across the Appalachia region of the United States. Similar spatial statistics were applied in Jones et al. (2016a), with the addition of a three step floating catchment area method to calculate neighbourhood level access to dental care and family physicians. This demonstrated that service types were concentrated close to other similar services, with dental services being more highly concentrated and located in the urban centre and important commercial areas to the east of the centre, areas of greater socio-economic advantage. Spatial autocorrelation was also used by Meyer (2014) to analyse 'hot spots' of dental clustering in major urban areas of Ontario. Hot-spots of dental offices tended to be located in downtown areas, with 'high-sale offices' typically in suburban areas adjacent to neighbourhoods with deficits in dental services, also being characterised by higher income, population, growth and younger residents. Jager et al. (2016) used a novel approach involving ordinary least squares regression and geographically weighted regression of demographic data and data on dental workforce losses to estimate the dental workforce in 2030 in Northern Germany. Many areas were shown to be over-served, while some rural regions had no dental services at all. Use of Moran's *I* also showed there was no compensation from over-served neighbouring areas.

Saman et al. (2011) utilised hot-spot analysis to identify clusters based on the percentage of adults with 6 or more teeth removed due to tooth decay or gum disease, along with cartogram analysis to demonstrate oral health disparities by population and the ratio of dentists per 10,000. Counties with lower populations and dentist ratios were shown to have worse oral health. A fuller explanation of cartogram-based methods will be provided in Section 4.10. Horner et al. (2007) used GIS to identify zip-codes falling outside a 10-mile radius from dental practices, before using a 'location set covering problem' (p.114) to allocate services to these areas. Some new practices were able to cover multiple zip-codes, as under-served areas tended to be clustered spatially. This is a rare example of using GIS for potential practical solutions, rather than descriptive analysis. Krause et al. (2012) highlight some of the general challenges in using GIS to analyse oral health disparities, in particular the collection of data at aggregated scales, making small area analysis difficult.

2.6.4. Spatial analysis and oral health outcomes

Many studies have used spatial analysis to investigate specific oral health outcomes, with a large proportion from Brazil. Antunes et al. (2001) used K-means clustering to identify groups and patterning of oral cancers in Sao Paulo. Oral cancer mortality rates were positively correlated with income inequality, unemployment, overcrowding and illiteracy, while decreases in disease showed positive correlations with higher socio-economic status. Central, northern and eastern regions of the city showed the highest rates. In a similar study Antunes et al. (2008) compared inner city rates of oral and pharyngeal cancer between male populations in Barcelona and Sao Paulo. Mapping both cities showed disease tended to be higher in more deprived areas, with a steeper social gradient and greater polarisation in Barcelona. The smaller geographical size of the areas in Barcelona may have influenced this though. Moyses et al. (2008) found distinct patterns of dental trauma incidence in the Brazilian city of Curitiba. This included associations between areas of 'substandard living' and the number of cases of dental trauma in the south of the city, demonstrating the need for both social and environmental variables in such research. Carvalho et al. (2010) found incidence of dental trauma to be highest in southern and western Curitiba, however environmental variables were not statistically significant. In total 28.7% of trauma cases fell outside health coverage buffers added to health centres. These two studies are notable for the inclusion of varying neighbourhood level indicators, and while not as conclusive as some (Macintyre et al, 2003), it demonstrates a different approach to analysing small areas.

Caries has also been studied from a spatial perspective on a number of occasions. Antunes et al. (2002) mapped clusters of caries data for 5 and 12-year-old children to demonstrate spatial disparities in the disease in Sao Paulo, the worst affected areas also being correlated with lower income and employment, overcrowding and income inequality. Work by Freire et al. (2013) used survey data for all state capitals and municipalities in Brazil to demonstrate spatial inequalities in caries nationwide (particularly between the north-east and south-east of the country), with higher levels in low income areas with poor access to fluoridated water. This study also found contextual variables to be as important as individual level indictors. Pereira et al. (2010) utilised GIS alongside multi-level regression, and showed associations between higher DMFT scores and lower income, overcrowding, infrequent dental attendance and lower education at the individual level. However, at the conglomerate level deprivation was associated with neither DMFT nor the care index (‘the percentage of teeth with decay experience that have been treated by filling’ – Public Health England, 2016 – p.12). This work highlights the benefits of considering mapping and statistics concurrently.

2.6.5. Multilevel analysis

A number of studies have taken a similar approach to Pereira et al. (2010) in using multi-level models to analyse spatial inequalities in oral health. Such approaches are used where data is hierarchical, and allows for it to be represented at more than one level. Gelman (2005) illustrates how these models estimate the predictive effects of an independent variable separately from others, while also being able to control for correlation and dependence between the individual and area level variables. Santiago et al. (2014) have also stated that ‘multilevel models allow the estimation of the contextual effect of a variable measured at the area level, considering the spatial distribution of individuals within the areas’ (p.20).

Pereira et al. (2014) used this method to study the gingival status of schoolchildren in Sao Paulo. Individual level variables related to gingival bleeding included caries experience, while the percentage of illiterate heads of families was the only statistically significant contextual level variable. Better oral health was found in downtown Sao Paulo, an area containing both higher incomes and levels of education. This study highlights an unfortunate trend within some of the Dental Public Health literature, which involves using aggregated individual level data as a contextual variable, rather than actual features of local environments. Antunes et al. (2006) used multilevel modelling to demonstrate not only spatial disparities in oral health among schoolchildren, but also inequalities based around educational provision, ethnicity and gender. Water fluoridation was also shown to have a significant effect on improving oral health profiles. Celeste et al. (2009) investigated the link between income inequality and oral health, and demonstrated that neither social capital nor public health services had the desired effect of reducing the Gini coefficient, with more services actually increasing the coefficient. Again, lower income was associated with poorer oral health, leading to a potential double burden for certain groups. A strength of the paper, and a rarity among the literature, is the attempt to control variables in order to represent different theoretical mechanisms and pathways throughout the analysis, based on 'a priori postulated pathways' (p.1471), such as those in Kawachi and Kennedy (1999), setting out a number of mechanisms by which health is affected by income inequality. Further work from Brazil includes a study by Pattussi et al. (2006), who showed that levels of empowerment reduced as DMFT scores increased, independent of social variables at both scales.

Borenstein et al. (2013) applied multilevel modelling in a study of neighbourhood effects on oral health in Toronto, and represents one of the few dental studies to include numerous built features of local environments. Dentist visits were more likely from those on low incomes, but less so for those without insurance. Social services were significantly associated with dental visits, while social services and park space were significantly associated with self-rated oral health. Santiago et al. (2014) used multi-level modelling to demonstrate that neighbourhood empowerment levels were negatively associated with caries levels, while individual social capital variables had little impact on the model. Lopez et al. (2009) have also demonstrated that the size of geographical areas used in such studies matters for the outcome, and Celeste et al. (2011) have urged similar caution with regard to the size of study areas. Finally, Turrell et al. (2007) showed that the 'neighbourhood' was associated with self-reported oral health, independent of socio-economic traits, in Adelaide, Australia. Disadvantaged areas were associated with negative oral health outcomes, although the authors do acknowledge the lack of theoretical grounding in their study.

2.6.6. Simulation modelling in Dental Public Health

Several studies have applied top-down simulation approaches such as system dynamics modelling to oral health related scenarios, including Hirsch et al. (2012). This study simulated the impact of interventions on childhood caries in Colorado, and found fluoride varnishes for children, treating mothers with xylitol, and motivational interviewing to have significant effects on caries reduction, even more so when the interventions were combined. Saman et al. (2010) used system dynamics modelling to demonstrate that the rate of retiring dentists would not be matched by incoming dentists in the US state of Kentucky, leading to recommendations for higher recruitment of rural students into dentistry.

Individual level simulation models are equally rare, and to date only one study has applied microsimulation to an oral health related study. Brown et al. (1995) investigated fee and insurance changes in the United States, finding that spending per capita decreased upon decreases in dental insurance coverage, while tooth decay subsequently increased. However, as this study took no consideration of small area geographical variation it could not be considered a 'spatial' microsimulation study as such.

The work of Metcalf et al. (2013) is a rare example of an agent-based model applied to oral health. This was combined with a system dynamics model, and demonstrated the importance of the spread of word of mouth among an elderly population in New York City, which lead to increases in care seeking and preventative screening. A similar study using system dynamics and agent-based modelling again demonstrated the potentially important influences of social networks on dental clinic visits, and subsequent oral health outcomes (Wang et al, 2016). Agent-based models have great potential within Dental Public Health research (Roudsari et al, 2016), with further examples provided by Sadeghipour et al. (2017). This research into friendship networks and tooth brushing habits demonstrated that behaviours diffused through developed friendship networks, with agents who were closer in these networks becoming more similar in their brushing habits. More popular agents were also shown to have better brushing habits, which may encourage others to improve theirs. Sadeghipour et al. (2016) have also used this method to investigate demand for dental visits, demonstrating that the oscillatory nature of dental demand is dependent on the social network structure of individuals, and the number of effective connections within these. Adegoke (2016) also submitted a thesis based on agent-based models of behavioural change with regard to oral care and hygiene; however, the details of this thesis are not publicly available.

2.7. Summary

This chapter has demonstrated the complicated nature of geographical studies, as well as potential methods to address these and analyse small area concerns surrounding health inequalities. The complicated nature of the determinants of tooth decay have also been outlined, demonstrating the difficulty in analysing such a multifactorial disease. Currently the spatial analysis within the Dental Public Health field, while increasingly including more interesting and novel GIS methods, has not focused on neighbourhood effects, or small area analysis in much detail, let alone enough to consider all of the relevant determinants for tooth decay. Many outcomes have also been descriptive in nature, without addressing the underlying determinants of the patterns observed. It seems that the general health inequalities literature also tends to consider a wider variety of neighbourhood level variables and theoretical considerations, and treats neighbourhoods as more than aggregated socio-economic variables. While not always the case within the dental literature, it is more common to see the use of deprivation indices and contextual level socio-economic variables to represent neighbourhoods in this field. While little theoretical work has been conducted on neighbourhood influences on tooth decay, the literature, and the results from these studies, are still potentially useful for informing future theoretical work.

Despite the current state of the field with regard to small area analysis, this chapter has presented a number of potential approaches. An approach that considers conventional trends associated with socio-economic and material issues, as well as relational representations of feelings, attitudes and perceptions would be a suitable approach for oral health related studies. Section 2.3 demonstrated how spatial microsimulation can be used to create new variables in a representative dataset including those of both a conventional and relational nature. The compatibility of these outputs with agent-based approaches adds further to their value, and allows for the potential dynamic simulation of these populations, giving more insight into the important mechanisms for tooth decay at the small area level. This has obvious advantages over traditional statistical methods, and is a combination not seen in the dental literature before. If such methods were driven by the existing theory within the field to aid variable choice, and organised around a framework designed for the analysis of health in neighbourhoods (Macintyre et al, 2002), it would have huge potential for investigating relevant pathways and mechanisms. Such an approach could add to theoretical understandings of neighbourhood environments in an oral health context, while also introducing a new combination of methods to the field that has the potential to be useful in other oral health related scenarios beyond tooth decay. Through an exploratory approach such as this it may be possible to identify what the most important features of neighbourhood environments are (if at all) for tooth decay, and how these differ depending on location. In the rest of this analysis the term ‘neighbourhood’ will primarily be used to describe analyses of the local environments in which individuals live, as an overarching term to encompass other concepts discussed in this chapter including place, context and area.

Chapter 3 – The theory behind neighbourhood effects on tooth decay

'He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast' (Leonardo da Vinci - 1452–1519)

3.1. Introduction

This chapter will present the main theoretical work undertaken for this research, and will start with a review of current theories into why health inequalities persist, before introducing the main theoretical framework that will guide this research, specifically the simulation modelling that will be covered in Chapters 4 and 5. Following the introduction of this framework, the theory identified in Section 2.5 of the literature review will be used to populate each element of the framework, along with details of the operationalisation of all of these constructs.

3.2. Theories behind health inequalities

Recent decades have seen an increased interest in the study of the social determinants of health and their role in shaping health inequalities. The World Health Organisation (WHO) have defined the social determinants of health as the interplay between 'context, structural mechanisms and the resultant socioeconomic position of individuals' (WHO, 2010 - p.6), which operate through intermediary determinants of health. The main intermediary categories include psychological factors, such as stressful living conditions and relationships, and social support; material circumstances in which individuals live, including housing, neighbourhood environments, financial constraints and work opportunities; and behavioural and biological factors, such as nutrition, activity, excess consumption, and genetics (WHO, 2010). Marmot (2005) describes tackling the social determinants of health inequalities as a key 'thrust' in attempting to reduce health inequalities. A basic premise behind the social determinants of health is not the management or treatment of health conditions once they have occurred (the downstream effects), but to look at the ‘causes of the causes’ (Marmot, 2005 - p.1102) (the upstream effects) that lead to these health outcomes in the first place. While this dichotomy may be a slight oversimplification, it nonetheless highlights the importance of tackling relevant social issues before they lead to disease.

Marmot states that ‘if health of a population suffers it is an indicator that the set of social arrangements needs to change’ (p.2). Watt (2007) suggests a number of strategies that could be effective in successfully tackling these upstream causes in relation to oral health, including the empowerment of individuals at risk, effective partnerships across all relevant agencies/departments, and the use of existing knowledge of effectiveness and good practice among others. The effects of the social determinants are not evenly distributed among populations however. Work by Wilkinson and Marmot (2003) highlights the steep social gradient present in many societies, whereby those further down the ‘social ladder’ are more susceptible to a variety of health conditions and diseases. Additionally, Wilkinson and Marmot (2003) state that this effect is not confined to the poorest in society, as those in the middle of the gradient are likely to have worse health outcomes than those at the top. Wilkinson and Marmot’s report identified nine variables that are key to producing health inequalities, each of which has a clear social gradient, these being: stress; conditions in early life; social exclusion; work; unemployment; social support; addiction; diet; and transport.

In 2005 the WHO established the Commission on the Social Determinants of Health, with the goal of addressing social factors that lead to ill health and inequalities in these illnesses. In 2010, the commission published a conceptual framework, focusing on the role social position plays in such inequalities (WHO, 2010). The framework is notable for attempting to distinguish between different levels of causation, and the mechanisms that cause the social hierarchy. While the social determinants of health have received increased attention in recent years, it is one of a number of theories as to how health inequalities are created and persist. Some of these theories have been largely discredited, while others are more widely accepted within the health inequalities literature. These theories are discussed below.

3.2.1. Existing theories on health inequalities within the literature

Artefact theory: Widely disregarded, this theory implies health inequalities are merely a statistical artefact related to the way social status has been classified over time (McCartney et al, 2013). The theory is undermined by the sheer scale of widespread evidence for health inequalities, even when different statistical measures of social status are used. The theory also implies increased levels of relative inequality are inevitable when the overall level of a particular outcome falls, due to it being easier for 'determinants to produce a relatively high risk' (Mackenbach, 2012 - p.764).

Health selection theory: Effectively reverse causation, it is implied that poor health (rather than socio-economic status) leads to a social selection process, or 'social slide', leading to observed associations between social status and health (McCartney et al, 2013). This theory has been disproved by evidence reaching as far back as the Black Report (1980), as well as by more recent accounts (McCartney et al, 2013). Evidence has shown that social determinants are the cause of poor oral health, particularly where social gradients in dental caries have been identified through socio-economic status (Hobdell et al, 2003). McCartney et al. (2013) have also stated that longitudinal work can be used to test such theories, as 'pre-morbid social status' (p.222) explains the majority of 'the concentration of ill-health in lower social groups' (p.222) as opposed to any social slide. With regard to tooth decay, Thomson et al. (2004) found significant differences between socio-economic groups persisting over time when measuring tooth loss from caries, while Poulton et al. (2002) also demonstrated the lifelong effects of low socio-economic status on decayed surfaces.

Cultural and behavioural factors: This suggests that the link between social strata and health outcomes is a result of the differences between social classes in terms of their health related behaviour (Bambra, 2011), with those in lower socio-economic groups more likely to engage in health damaging behaviours (Sisson, 2007). However, McCartney et al. (2013) noted that this would mean socio-economic status was only an effect modifier. They state that when populations from differing social groups have equal exposure to behavioural risk factors, mortality is still higher in the lower socio-economic groups. Secondly, focusing solely on behaviours ignores how and why individuals adopt certain behaviours in the first place. Finally, the link between health behaviours and lower social standing has almost disappeared in some populations, but the link between lower social status and mortality remains. Sisson (2007) comments that human behaviour is extremely complicated, and therefore unlikely to be influenced purely by behavioural and cultural factors. Regarding oral health, Section 2.5 of the literature review demonstrated that lower socio-economic groups exhibited less favourable behaviours with regards to diet (National Diet and Nutrition Survey, 2014), oral health behaviours (Sabbah et al, 2009; Singh et al, 2013), and dental attendance (Eckersley and Blinkhorn, 2001; Lang et al, 2008). However, in line with the comments made by Sisson (2007) and McCartney et al. (2013), it seems unlikely that these factors alone explain oral health outcomes, and would leave an incomplete picture as these are known to be influenced by factors such as income in the case of diet (Fonseca, 2012), deprivation in the case of attendance (Lang et al, 2008), and education with regard to oral health habits (Sabbah et al, 2009).

Cultural capital: Influenced by the previous theory and originating from Bourdieu (1984), this concept posits that health inequalities can be explained by differences in attitudes, competency and knowledge between groups that have intergenerational cycles. These differences arise from the need for 'social distinction' with regard to being able to 'show off' one's social position, which requires a lot of cultural capital (non-financial social assets). Opportunities to distinguish oneself through health related behaviours may have increased due to fewer opportunities to do this through material means (Mackenbach, 2012). While there are no explicit studies on oral health related examples of showing off, the idea of intergenerational inequalities due to differences in attitudes and the cycles of these could exist through the influence of parental attitudes and knowledge on those of their children and their subsequent oral health behaviours (Adair et al, 2004; Poutanen et al, 2006), as well as the importance of parental education and occupation on their children's oral health (Saldunaite et al, 2014; Vanobberge et al, 2001). Some studies have also shown that maternal oral health when children are young is a good predictor of oral health status in adulthood with regard to caries (Shearer et al, 2010). Chaffee et al. (2014) also demonstrated that maternal salivary bacteria was longitudinally associated with childhood caries from birth to 36 months, while Bedos et al. (2015) have shown intergenerational differences in tooth decay in children based on maternal eduntulousness.

Psychosocial factors: A popular theory within health inequalities research, Mackenbach (2012) describes the key components of this theory as 'psychosocial stress, lack of social support and sense of control' (p.765). The premise is that some people are worried, anxious and feel situations are out on their control, and that continued stresses throughout life can increase the risk of poor mental health and death. A particular feature of this is known as the ‘fight or flight mechanism’, where hormones and the nervous system prepare individuals to deal with threats. For short periods this process has no lasting effect, but as it diverts the body’s resources form physiologically important processes it can have negative health effects if triggered continually over longer periods (Wilkinson and Marmot, 2003). Those further down the social scale are more likely to experience sustained stresses, in relation to unstable employment or lack of income to provide adequate material resources for example. Within the oral health literature there is support for this theory, with factors such as stressful environments which have biological impacts (Boyce et al, 2010), lack of resources (Finlayson et al, 2010) and personal constraints (Sanders and Spencer, 2005) all contributing.

The life course perspective: The premise behind this theory is that the life a person has lived up until a given point has consequences for their current health, something cross-sectional analyses often fail to account for. Sisson (2007) highlights two different models associated with this theory – the accumulation model, and the critical periods model. The first asserts that exposures to advantage or disadvantage at different stages of life have a cumulative effect, which decreases or increases the chance of developing chronic diseases. Circumstances in childhood set an individual on their own ‘health trajectory’, with no one factor being more important than another. The second suggests that some outcomes (e.g. heart disease, stroke) have origins ‘during critical periods of development’ (p.2) that determine future health outcomes independent of intervening factors. Ben-Shlomo and Kuh (2002) define these periods as ‘a limited time window in which an exposure can have adverse effects on development and subsequent disease outcome. Outside this window, this developmental mechanism for mediating exposure and disease risk is no longer available’ (p.288).

This theory has a number of strengths, as it provides a history of, and explanation for the development of diseases in people, as well as the persistence of inequalities over time. There has been support for the validity of both the critical periods model (Graham, 2002), and the accumulation model (Poulton et al, 2002) with regard to caries and periodontal disease. Thomson et al. (2004) also demonstrated the importance of the life course to a number of oral health outcomes including tooth decay. While some studies found no longitudinal associations with variables such as tooth retention (Heilmann et al, 2015), Nicolau et al (2007) found associations between low paternal education in childhood and adult levels of periodontal disease, supporting both the accumulation and critical periods model. All of these oral health studies showed significant associations between childhood socio-economic status and adverse oral health outcomes. Mackenbach (2012) has stated however that mortality and morbidity rates can often respond quickly and dramatically to changing social conditions, not always with the delays implied by the theory.

Neo-materialist theories: According to Peacock and Bissell (2011) neo-materialist theories comprise a combination of negative exposures due to a lack of resources, and a lack of investment in human, cultural and political-economic processes. This can, for example, focus on issues of income and the subsequent (lack of) access to goods, services and resources that may reduce physical and psychological risks (Bambra, 2011). The theory also encompasses structural elements, including issues such as taxation, legislation, regulation, policy, and changes in the broader distributions of income and power within society, moving away from focusing purely on the individual. This theory has particular relevance to oral health as those lacking in income (Costa et al), as well as resources such as food (Mobley et al, 2009; Fonseca, 2012) and oral health related knowledge (Williams et al, 2002) have all been shown to have higher levels of tooth decay. Wider structural issues, although not as well examined within the dental literature, also appear to have impacts, including making dentists harder to access for certain populations (Landes and Jardin, 2010), while the impact of income inequality measured using the Gini coefficient has also shown positive associations with tooth decay (Celeste et al, 2009).

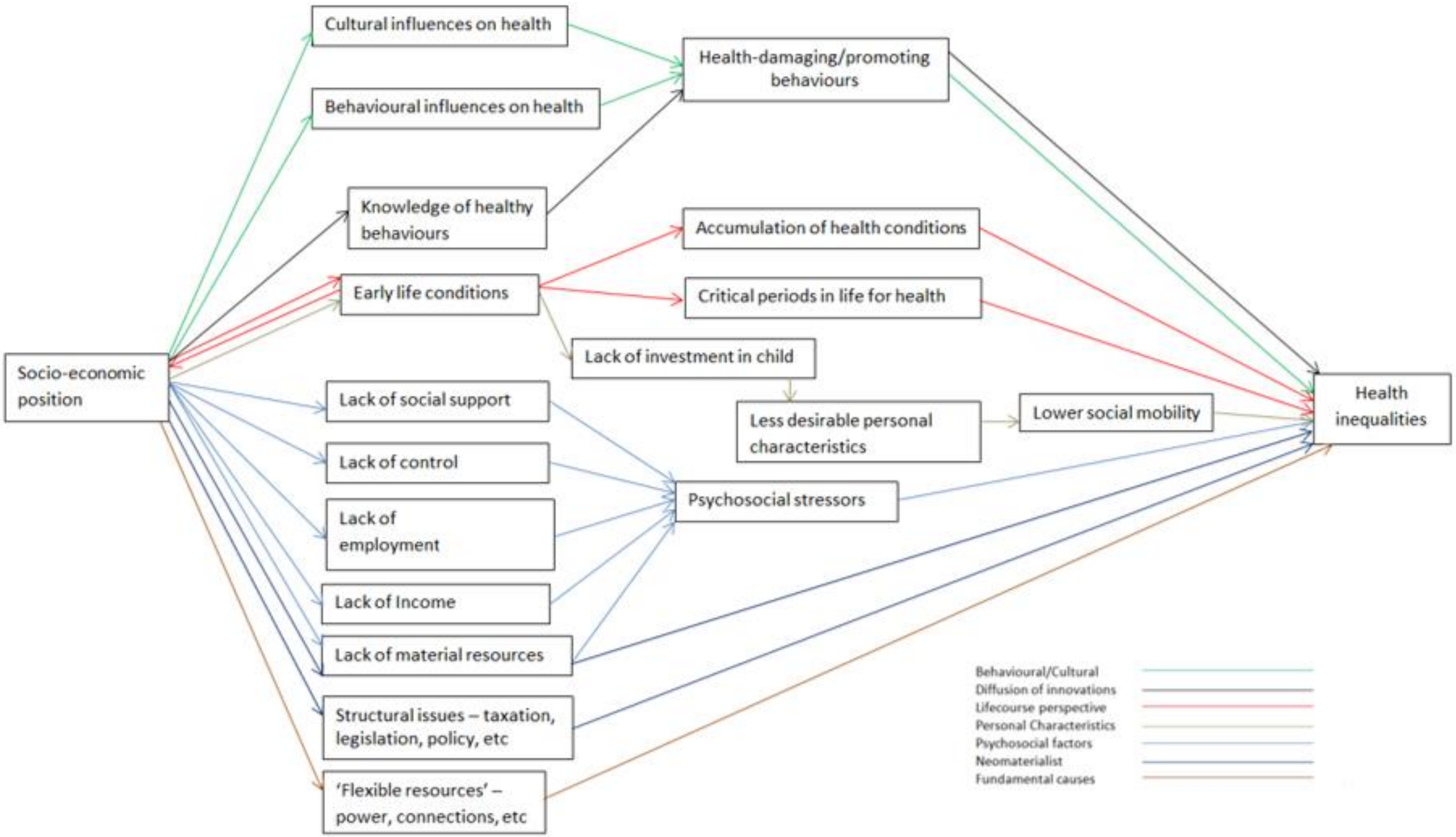
Mackenbach (2012) introduced three additional theories. The first of these, ‘fundamental causes’, was conceptualised by Link and Phelan (1995), and emphasises social forces underlying social stratification as the key to health inequalities. The persistence of such inequalities (over different time periods, and different national conditions) is due to social standing affording some individuals ‘flexible resources’ (p.764), including money, power and social connections that can be used to avoid, or minimise the consequences of disease. Similar to the neo-materialist theory, this theory would seem to have some credibility within Dental Public Health, as it has been shown that those of higher social standing have more flexible resources with regard to access to dental services (Landes and Jardin, 2010), as well as choice of services (McKernan et al, 2016; Nasseh et al, 2017). Income differences have also been shown to offer different access to food (Fonseca, 2012), as well as a variety of resources that are beneficial to oral health including access to products that maintain good oral health, and information that promotes better oral health (Costa et al, 2012).

The second theory, 'diffusion of innovations', posits that those in higher socio-economic groups tend to adapt to new health behaviours earlier than those of lower socio-economic status, providing a plausible explanation as to how health inequalities may widen when 'major improvements in population health are mediated by behavioural change' (p. 765). While there are no explicit examples of this within the dental literature, it could be argued that those who have higher education qualifications are more likely to take on more favourable oral health related habits (Sabbah et al, 2009), with those from less deprived areas having more regular attendance patterns which allow them to take advantage of professional advice (Lang et al, 2008), thus taking advantage of their position in society to better their health.

Finally, the 'personal characteristics' theory posits that upward mobility is more likely in people who have sets of personal characteristics that are 'conducive to good health' (p.764). These characteristics include cognitive ability and personality profiles, including the 'big five' personality traits (openness to experience, conscientiousness, extraversion, agreeableness and neuroticism). This has the potential to reduce intergenerational transmissions of conditions, diluting 'the health effects of social stratification' (p.675). Socio-economic inequalities leading to stressful family conditions and differences in early life may hamper investment in children, which can lead to inequalities in these characteristics. Health inequalities could have increased based on these personality characteristics, as they affect health related behaviours and social position. Certainly within the dental literature education (as a marker of cognitive ability) has been shown to be associated with tooth decay (Muirhead and Marcenes, 2004; Brennan et al, 2007), while stressful events have negatively impacted self-reported oral health (Sanders and Spencer, 2005; Finlayson et al, 2010), as well as caries in children (Boyce et al, 2010) through financial stress in the home.

The basic structures of the above theories are visualised in Figure 2, which was created as part of this research. The artefact and health selection theories have not been included, due to their widespread dismissal within the literature.

Figure 2 – Basic pathways of the theories of health inequalities



3.3. A new approach to theorising the role of neighbourhoods in oral health inequalities

As can be seen from the work presented in Section 2.6, the approaches taken within the dental literature using spatial analysis which either includes neighbourhoods within studies, or represents aspects of these, has more often than not done so by using deprivation indices or through the use of an aggregated social indicator (socio-economic status or deprivation for example) to act as an area-based score (Rocha et al, 2013; Pereira et al, 2014). Fewer studies have attempted to represent the built, physical or collective characteristics of different areas, specifically indicators that are not divisible to individual level characteristics (Pickett and Pearl, 2001; Cummins et al, 2005). As such, it would appear that there has been a lack of theorising of some potentially important features of place which may influence inequalities in tooth decay.

Section 3.2 presented an overview of the main theories that have been presented in an attempt to explain continuing patterns of health inequalities. Many of these theories are not mutually exclusive, and have common elements to them. Neo-materialist factors are thought to have significant psychosocial impacts for example, while the behavioural and diffusion of innovations theories share similar traits. However, these theories may not collectively capture enough of what it is about where we live that affects our health. Relatively speaking the theories are quite general, and tend not to theorise which features of the local environment are most important to health inequalities at the small area level.

Therefore, a new approach is required which can offer a potential solution to these aforementioned issues. This new approach can be guided by the theoretical work of Macintyre et al. (2002). Macintyre and her colleagues believe that many of the theories that explain health inequalities place more focus on the individual (with the exception of neo-materialist theories), and have often not taken into account social processes and the surrounding environments to which individuals are exposed concurrently, commenting that ‘within this research field, contextual material and institutional resources and opportunity structures have been less the focus of research than collective psychosocial characteristics’ (p.130).

As a result, Macintyre and colleagues conceptualised an organising framework by which hypotheses about the role of place may be tested, specifically concerning contextual aspects of the built, social and physical environment that may affect health. The features of the framework are listed below.

1. ‘Physical features of the environment shared by all residents in a locality’

This includes features such as the quality of air and water, latitude and climate, all of which are likely to be shared across wider areas by a number of different neighbourhoods. Macintyre et al. (2002) cite the example of Glasgow’s water supply, which all comes from the same loch, stating that the city’s large inequalities in death rates cannot therefore be explained by variations in drinking water.

2. ‘Availability of healthy environments at home, work and play’

The features of this pathway include elements such as the provision of acceptable housing, non-harmful employment sites and safe areas for children to play. Levels of these will vary by area, and may not affect all residents in the same way. Fewer safe play areas are more likely to affect families with young children for example.

3. ‘Services provided, publicly or privately to support people in their daily lives’

Services such as education, transport, policing, welfare and street maintenance are given as examples, with the effects potentially varying based on peoples’ circumstances. For example, poor public transportation may not matter as much if individuals own a car.

4. ‘Socio-cultural features of a neighbourhood’

This pathway includes features such as the economic, political, ethnic and religious traits/history of a community, as well as the values (and subsequent norms) associated with these. This can also include factors such as levels of crime, feelings of personal safety, and community support networks.

5. ‘The reputation of an area’

The final pathway considers how the way an area is perceived by a number of different groups (residents, service planners, banks/investors, etc.) can affect the infrastructure of these places. Reputation can also influence who moves in and out of said areas, as well as the self-esteem of those who reside there.

The first three measures, focusing on material and infrastructural resources, are labelled ‘opportunity structures’, which are socially constructed, and socially patterned features of the physical and social environment, promoting or damaging health both directly and indirectly through the possibilities they provide. The last two measures are referred to as ‘collective social functioning and practices’. Examples of these given by the authors include ‘people’s use of their local area, perceived social cohesion, aggregate and neighbourhood level measures of social participation such as the number of local voluntary groups, and symbolic and actual representations of the areas via newspaper and other accounts of the area, including photographs’ (p.132), and are designed to represent ‘non-material culture’ (p.135). Thus the authors have considered both social and physical (both built and natural) features of the local environment, as well as those that affect individual practices, and those of wider communities as a whole, bridging the gap between these aspects.

This conceptual work represents a comprehensive approach to studying the various aspects of the local environment in a health inequalities context, and this is why it has been chosen to guide this research. There have been other attempts at conceptualising what it is about place that influences health, however most have focused on a particular approach (i.e. relational or conventional), and not considered a wide variety of factors or scenarios. Smith and Easterlow (2005) for example focused entirely on housing from a relational viewpoint, while Popay et al. (2003) concentrated their research on the relational aspects of normative dimensions of place. On the other hand, studies such as Carvalho et al. (2010) attempted to include features of the local environment, but featured only a small variety of these. This study also focused mainly on conventional approaches, without going into the relational depth of the previous two studies. Macintyre et al. (2003) may represent one of the more comprehensive studies in this area, including data on individual circumstances, opinions on neighbourhood amenities, and neighbourhood based service scores.

The conceptual pathways of the organising framework suggested by Macintyre and colleagues (2002) span both conventional and relational viewpoints, covering themes such as social interactions and practices, political and structural factors, naturally occurring elements of the physical environment, local infrastructure, and traditional socio-economic and material indicators. So far, no other framework that lends itself to the study of neighbourhood environments has been identified that covers such a wide variety of themes from a number of different viewpoints in this depth. The framework also incorporates elements of all seven of the health inequalities pathways shown in

Figure 2 (Section 3.2), which, combined with its comprehensiveness, makes it potentially suitable for the study of neighbourhood effects in relation to tooth decay.

Macintyre and colleagues do acknowledge, however, that their framework is still limited in that it does not specify exactly what needs to be studied in these broad pathways, and that ‘a starting place for conceptualising and measuring area influences on health is to consider what humans need in order to live a healthy life’ (p.130). The authors also comment that a legitimate question to ask when addressing the framework would be, within a given society, what the geographical distribution of resources to meet the needs for healthy lives is, and whether this distribution is associated with geographical distributions of health. As such, it would make sense to operationalise measures that are representative of particular societies in specific historic periods in particular places. Finally, they state that they ‘are not suggesting that every study of area and health try to cover every single measure of the physical and social context which might influence human health. Rather, we wish to suggest that investigators should attempt to hypothesise and test specific pathways by which area might influence health’ (p.133). This could involve for instance, identifying major risk factors of a disease or outcome, and linking these with ‘features of the material and social environments which promote such risk factors’ (p.133).

The acknowledged lack of specificity in the conceptual pathways presents an opportunity for those studying place effects to employ variables relevant to their own research, as the framework could be applied to any health related field with relevant theory. In this research the pathways were created using the theory from Section 2.5, as well as theory from the wider health inequalities field where suitable. The pathways are therefore purely driven by theory, and not by data availability. Data and relevant variables are still an important part of this research though, and as such the operationalisation of each pathway will also be discussed.

Each of the five pathways represents a particular way in which place affects health. Variables that theoretically fit within the remit of each of these pathways, that represent an element of the neighbourhood environment, and that are relevant to inequalities in tooth decay will form the first step(s) of the pathways. For example, in the third pathway education is given as an example of a place based service to support people in their daily lives. As education is also relevant to inequalities in tooth decay, it forms one of the first steps in the third pathway. From here the theory gathered in Section 2.5 (and from the wider health inequalities field) was used to fill out the rest of the pathways, creating a sequence of events whereby neighbourhood environments influence levels of tooth decay in individuals. This approach is beneficial in that it allows for the inclusion of both individual and household level variables, as well as neighbourhood level indicators. Jamieson and Thomson (2006) have commented that individual and household level variables can fail to capture contextual influences involved in health inequalities, while a danger of contextual and area based measures is their assumption of homogeneity among the population in given geographical areas. The authors further add that there is likely great benefit for Dental Public Health research in ‘using a dual socioeconomic measurement approach to population research’ (p.108). The approach taken in this chapter allows for this to happen.

The following sections will first present the data sources considered for operationalising the theoretical pathways (Section 3.4), before presenting the justification for the pathways that have been constructed for studying inequalities in tooth decay (Section 3.5). Each pathway will have its own section, and will be accompanied by a visual representation of the relevant pathway (Figures 3-7), as well as data tables outlining the process of operationalisation (Tables 4-7).

3.4. Operationalising the oral health inequalities pathways

The Adult Dental Health Survey (2009) was the main source of survey data used for individual level responses in this research. The Adult Dental Health survey is a decennial representative national survey that was first conducted in 1968. The 2009 version, commissioned by the NHS Information Centre for Health and Social Care, consisted of interviews with 11,380 individuals, collected from 12 strategic health authorities across England, Wales and Northern Ireland. Of the original sample, 6,469 dentate adults underwent dental examinations. While other surveys include questions relating to dental health (e.g. number of teeth) or dental visiting patterns, none match the Adult Dental Health Survey for the sheer number of clinical variables. As well as a vast amount of clinical data the survey also includes information on socio-demographics, behaviours and material circumstances, as well as information on attitudes and service use. Upon consultation with Professor Zoe Marshman, from the Unit of Dental Public Health in the School of Clinical Dentistry at the University of

Sheffield, the ‘numdu98’ variable was chosen as the outcome variable for this research. This represents the number ‘of decayed or unsound teeth’, not including those with fillings or those that have been extracted, using the 1998 classification. This variable essentially acts as the endpoint of the pathways, as the outcome of this analysis will always be how the theoretical pathways influence levels of tooth decay. As the processes that occur along the pathways influence different people in different ways depending on a number of characteristics, the final point in the pathways will also represent how such processes lead to inequalities in tooth decay between different groups.

In order to test any theoretical pathways relevant data needs to be available that can be assigned to each of the pathway’s constructs. It is important to consider the type of data that is relevant to each section of the pathways as well. For example, as has already been mentioned, the first ‘row’ of boxes in the pathways represent a feature of the neighbourhood environment from which the impacts on tooth decay begin (i.e. not individual characteristics but those of the built, structural, or collective environment) while the boxes further down the pathways represent individual characteristics and behaviours. Thus it is important that variables are assigned to represent the right types of construct. Along with the ADHS, additional surveys were considered as data sources, including:

- The Health Survey for England - 2009
- Understanding Society – 2009-2013
- The Poverty and Social Exclusion Survey - 2012
- The British Social Attitudes Survey - 2013
- The Living Costs and Food Survey – 2013
- The National Diet and Nutrition Survey – 2014

While these surveys all contained data that would be useful to this research, and could supplement data from the ADHS where it may be lacking, the complications of including multiple surveys, and the population demographics associated with these, meant that this was beyond the scope of the current research. The next section will present the theoretical justification for the creation of each pathway, followed by a visual representation, a discussion of the available data, and tables showing the sources of this data for the separate elements of the pathways. Within the visual representations

(Figures 3-7) each box contains the name of a construct used in this research, with a more accurate description of how this construct is interpreted in the context of this research shown beneath in brackets and italics.

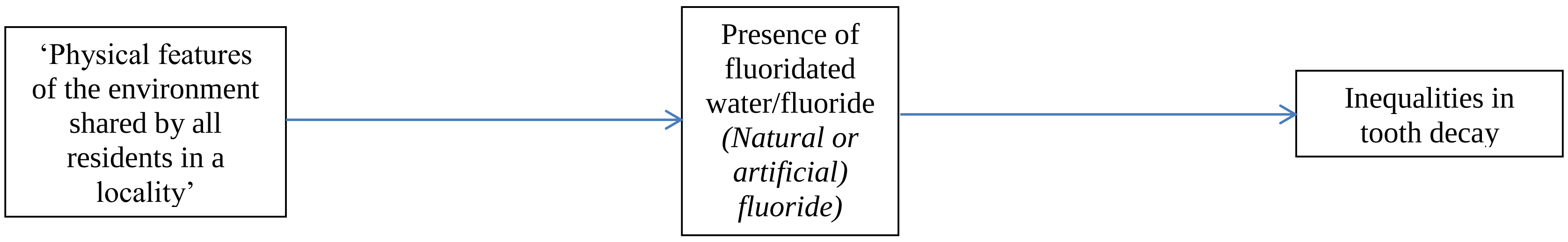
3.5. Creating the oral health inequalities pathways

Pathway one - Physical features of the environment shared by all residents in a locality

Macintyre and colleagues (2002) describe the physical features of the local environment as variables such as air quality, water quality, latitude and climate. Thus in this sense the word 'physical' refers to the natural environment, rather than physical structures within it. A further tenet of this pathway is that the variables in question should affect all residents that are exposed to it in much the same way. With regard to inequalities in tooth decay there appears to be only one such feature that is theoretically relevant, this being the presence of fluoridated water. Here both naturally and artificially fluoridated areas could be considered.

As demonstrated in Section 2.5, optimal levels of either type of fluoride supply have been shown to reduce levels of dental decay (McDonagh et al, 2000; Griffin et al, 2007) in populations, albeit with the possible side effect of mild fluorosis in some cases. A further benefit of water fluoridation appears to be the reduction of the socio-economic gradient in oral health inequalities in certain areas (McGrady et al, 2012). Environments that receive either kind of fluoride supply therefore would seem to aid in maintaining healthier teeth, or at the very least have some sort of protective effect, if only to a small degree. The British Fluoridation Society (2012) estimates that 330,000 people in the UK receive optimal levels of natural fluoride in their water, while a further 5,797,000 receive artificial fluoride to the optimal level. Thus, the distribution of this resource is far from universal within the UK, with the result being that certain areas will reap the benefits of this protective resource, while others are not so fortunate. This pathway is visualised in Figure 3.

Figure 3 – Pathway 1: pathways for the 'Physical features of the environment shared by all residents in a locality'



With regard to the first pathway, the Adult Dental Health Survey has two variables concerned with fluoride levels. However, these are related to practices such as the use of fluoride toothpastes and mouthwashes, rather than the levels of fluoride in the local water supply. It may be possible to obtain such levels from either the local water board (e.g. Yorkshire Water), or through oral health needs assessments for relevant areas. However, the city of Sheffield is not fluoridated to the optimal level of one part per million (ppm), either naturally or artificially, and only certain areas to the south-east of the city receive naturally fluoridated water, at a level of 0.4 parts per million (Public Health England, 2015). Thus this is not a practical variable for the context of this research. Nevertheless, water fluoridation is a theoretically important aspect of the local physical environment, and is certainly worth consideration in future research on neighbourhood effects in other locations in the UK and abroad where relevant. Certainly the methods used in this research would be suitable for testing the effects of this particular intervention for reducing inequalities in tooth decay.

Pathway two - Availability of healthy environments at home, work and play

The primary examples given for this pathway by Macintyre and colleagues are the provision of 'decent housing, secure and non-hazardous employment, safe play areas for children, etc.' (p.131). However, unlike the previous pathway, these features may not necessarily affect all residents in the same way.

While no literature suggesting that housing type influences tooth decay has been identified, the material conditions associated with household ownership do appear to have an impact. Nicolau et al. (2005) have highlighted the importance of material circumstances to oral health, including variables such as inside toilets, household construction material, the availability of piped water, and ownership. Dalstra et al. (2005) have stated that home ownership can have a psychosocial impact, through feelings of being in control of life, and being able to take pride in one's home. Thus, a lack of feeling of control and pride could lead to a number of psychological stresses. Such stresses are also known to be detrimental to general health (Marmot et al, 1991), as well as oral health related outcomes (Sanders and Spencer, 2005; Finlayson et al, 2010). Stressful situations have also been associated with increased tooth decay, with biological factors shown to be important (Belstrom et al, 2014), through the stress associated with low socioeconomic status (Boyce et al, 2010).

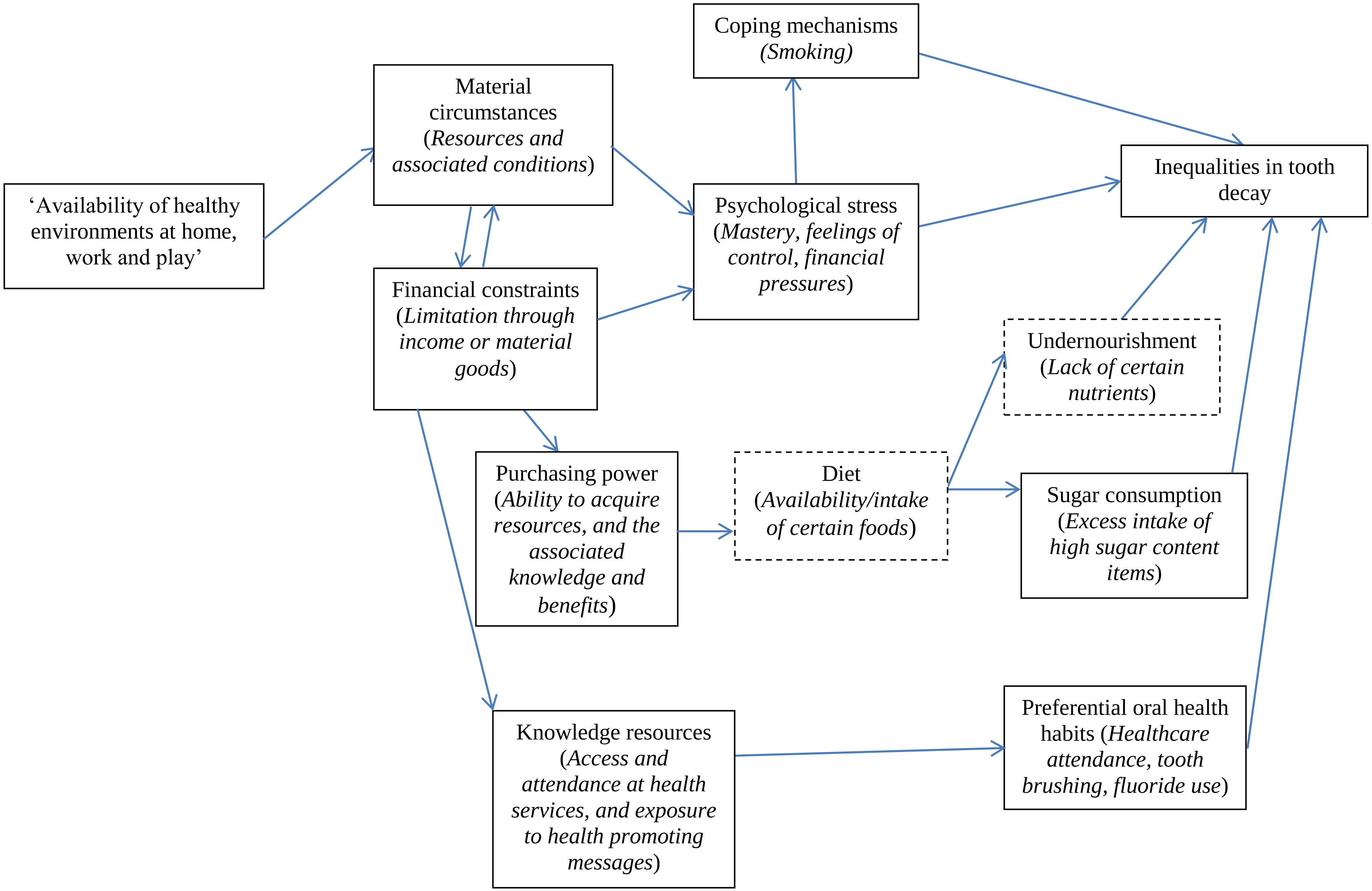
The stress of dealing with these issues can lead some to employ certain coping mechanisms (Popay et al, 2003), to either alleviate the stress or take them away from their problems temporarily. One coping mechanism is smoking, which has been shown to be associated with increased stress levels in both adolescents and adults (Kassel et al, 2003; Ng and Jeffrey, 2003), with financial pressures also showing strong associations (Siahpush and Carlin, 2005). Bernabe et al. (2014) found associations between smoking and caries; however, these conclusions differ to studies that found no direct links (Reibel, 2003; Vellappally et al, 2007). In contrast, other evidence suggests smoking could still be considered a risk factor for decay (Axelsson et al, 1998; Reibel, 2003), and can lead to teeth being more susceptible to decay (Vellappally et al, 2007). There do not seem to be direct links between excess alcohol consumption and tooth decay however (Harris et al, 1997).

Material possessions and circumstances are also likely to reflect financial constraints impacting on an individual or a family, and a study conducted by Barker et al. (2008) demonstrated that this can have a direct impact on the diet of certain groups and individuals, limiting the types of foods that can be purchased (Fonseca, 2012). This includes fresher, healthier and ultimately more perishable items, which are sometimes avoided due to the risk of wastage. Frozen and ready meals often provide a more cost-effective alternative. The diet an individual has can have an impact upon their oral health through the excess intake of certain foods. Diets rich in sugar, in particular fizzy drinks, are shown to have a detrimental effect on teeth, with increasing caries levels being the most common concern (Burt et al, 2006). Such patterns are known to be more prominent among lower income groups (Burt et al, 2006; Warren et al, 2009). A lack of healthy and nutritious food is also detrimental to oral health, and lower intake of essential vitamins has also been shown to be associated with low income groups (National Diet and Nutrition Survey, 2014). Research has shown that undernourishment can lead to children in particular being more susceptible to tooth decay (Moynihan and Petersen, 2004; Sheiham, 2006).

Financial constraints also have the potential to indirectly impact on dental knowledge. As highlighted in Section 2.5, those living under greater financial constraints or in deprived areas are likely to have a lower overall dental knowledge than those from less deprived backgrounds. Williams et al. (2002) demonstrated this in their study when confirming that ‘parents from Asian families, those living in deprived areas and parents who have had no further education all had less chance of having high levels of dental knowledge and positive dental attitudes' (p.653). Higher levels of dental knowledge have been shown to be beneficial with regard to caries levels in children, as 'parents with a positive dental attitude and better dental knowledge will probably build up better oral health habits in their children and look after the children's oral health' (Chu et al, 1999 - p.619). Work by Sabbah et al. (2009) and Singh et al. (2013) has also shown socio-economic and educational gradients in oral health related behaviours in adults in the USA and UK.

No literature on the play environment and its potential impact on tooth decay or inequalities in decay was found. It may be that these environments are more likely to be associated with other oral health outcomes such as dental trauma (Moyses et al, 2006: Harris et al, 2007; Moyses et al, 2008; Carvalho et al, 2010), but is an area that remains relatively underexplored. Studies of the work environment are also less common, with many including socio-economic status or an occupation based variable to represent an individual's 'work', rather than the setting in which the work takes place. One study focusing on work related stress found this to be related to periodontal disease rather than tooth decay (Marcenes and Sheiham, 1992). The above literature was mapped onto the second pathway which is shown in Figure 4. The variables in the boxes marked with dashed lines are those that either had to be removed due to a lack of data, or changed due to available data.

Figure 4 – Pathway 2: pathways for the 'Availability of healthy environments at home, work and play'



Firstly, it is important to address the two variables in the dashed boxes, 'diet' and 'undernourishment'. Diet was originally intended to be represented by data on the number of servings of fruit and vegetables eaten each day from the Understanding Society dataset, however as previously stated, technical difficulties in combining two survey datasets meant this was not possible. Diet was therefore replaced with the 'ncakes' variable from the ADHS, representing the number of cakes eaten per week. While perhaps not as good an indicator of diet as fruit and vegetable consumption it nevertheless acts as a proxy for diet. Unfortunately, dietary based variables are relatively limited within the ADHS, and no suitable proxy for undernourishment was found. This part of the pathway was therefore removed as it could not be operationalised. Diet was therefore still included in the simulation, just in a more truncated form. This represents the type of compromise that has to be met when undertaking an analytical approach to theoretical pathways such as in this research.

Within the data tables for each pathway the neighbourhood level variable will be presented first (or in the first few rows where there are multiple of these), before moving on to the data from the ADHS. Where possible the neighbourhood level variables were not meant to be divisible to individual level characteristics (Pickett and Pearl, 2001; Cummins et al, 2005), or essentially the whole pathway would represent individual characteristics in some form, and not those of the neighbourhoods they live in. The neighbourhood based indicator for the second pathway was the median house price for MSOAs in Sheffield (c8,500 people), which was used to act as a neighbourhood level indicator of material circumstances (Table 4). Previous studies have indicated that house prices may act as a good indicator of material circumstances in different areas (Nkosi et al, 2011). The Joseph Rowntree Foundation have also commented that 'among those who have housing and are not homeless, housing costs constitute the most important and most direct impact of housing on poverty and material deprivation' (Tunstall et al, 2013 - p.33).

The Adult Dental Health Survey provided adequate data coverage for all eight of the individual based variables in the second pathway, or variables that can at least act as proxies. Financial constraints were represented by whether an individual had to delay treatment due to cost, while knowledge resources was represented by whether an individual used other dental hygiene products beyond tooth brushing (i.e. other sources of fluoride, or flossing). Purchasing power, in a similar fashion to financial constraints, was represented by whether the cost of a certain type of treatment affected an individual’s choice to undergo treatment. Psychological issues were represented by how often in general an individual had felt psychologically tense, and while ideally a psychological variable related to resources or income would have been preferable, this variable represented one of the few stress based questions in the ADHS.

Coping mechanisms were represented by a variable enquiring as to whether the survey participant had ever been given advice on quitting smoking. While a variable on the frequency of smoking would have been preferable, the lack of smoking variables in the ADHS meant this was not possible. The variable used to represent smoking was thought to be preferable to the ‘DVSMOKE’ indicator (‘smoking status’), as simply being a current smoker would not necessarily indicate a problem with smoking explicitly, whereas having to get advice about quitting would seem more likely to. While this choice is to some degree based around assumptions, it is also based around finding variables that are as logical as possible to represent the pathway variables from the data available. Sugar consumption was far easier to represent through the use of the ‘highsug’ variable, classified as ‘high’ or ‘not high’ sugar intake. While more scales in this variable would have been preferable, it is a highly relevant variable to map onto this construct. Finally, preferential health habits were represented by general dental attendance, which related not to exact frequencies of visiting (i.e. ‘every six months’, ‘once a year’, ‘once every two years’), but rather the more general approach taken by individuals to attendance (i.e. ‘regularly’, ‘only when symptomatic’). These variables can be seen in Table 4.

Table 4 – Variables for the 'availability of healthy environments at home, work and play' pathway

Pathway	Pathway variable	Data/source	Present vs missing values
2	Material circumstances	'House Price Statistics for Small Areas, 2013' – Neighbourhood Statistics (MSOA)	N/A
	Financial constraints	'Whether had to delay dental care/treatment because of cost' (ADHS)	11370-10
	Knowledge resources	'Use other dental hygiene products' (ADHS)	11289-91
	Purchasing power	'Whether cost affected type of dental care/treatment' (ADHS)	11337-43
	Psychosocial issues	'Psychological discomfort – self tense' (ADHS)	11378-2
	Coping mechanisms	'Ever been given advice on giving up smoking' (ADHS)	10774-606
	Diet	'How often eat cakes' (ADHS)	11378-1
	Sugar consumption	'(D) high sugar intake – version 5' (ADHS)	11377-3
	Preferential health habits	'General dental attendance' (ADHS)	11342-38

Pathway three – Services provided, publicly or privately to support people in their daily lives

Education, transportation systems, street cleaning and lighting, police services, health services and welfare services are all given as primary examples of the main features of this pathway by Macintyre and colleagues. From the literature surrounding inequalities in tooth decay so far it would seem that there are four main services that are relevant to this pathway that influence oral health inequalities – education, employment opportunities, shops and supermarkets, and dental services.

Even in childhood, education levels have been shown to be associated with dmft levels through linguistic, English and maths scores (Muirhead and Marcenes, 2004), and this effect has been shown to continue into adulthood (Brennan et al, 2007; Mamai-Homata et al, 2012). Conversely, inequalities in tooth decay can also have an adverse effect on educational levels, as Sheiham (2006) reported that time spent away from the classroom dealing with oral health issues and restricted activity can reduce a child's ability to learn. Education is likely to have a large influence on future prospects, in particular employment opportunities available to individuals, and as a result associated factors such as income earned in adulthood (Kuh and Wadsworth, 1991). Such processes will likely impact upon a range of lifestyle factors, and opportunities.

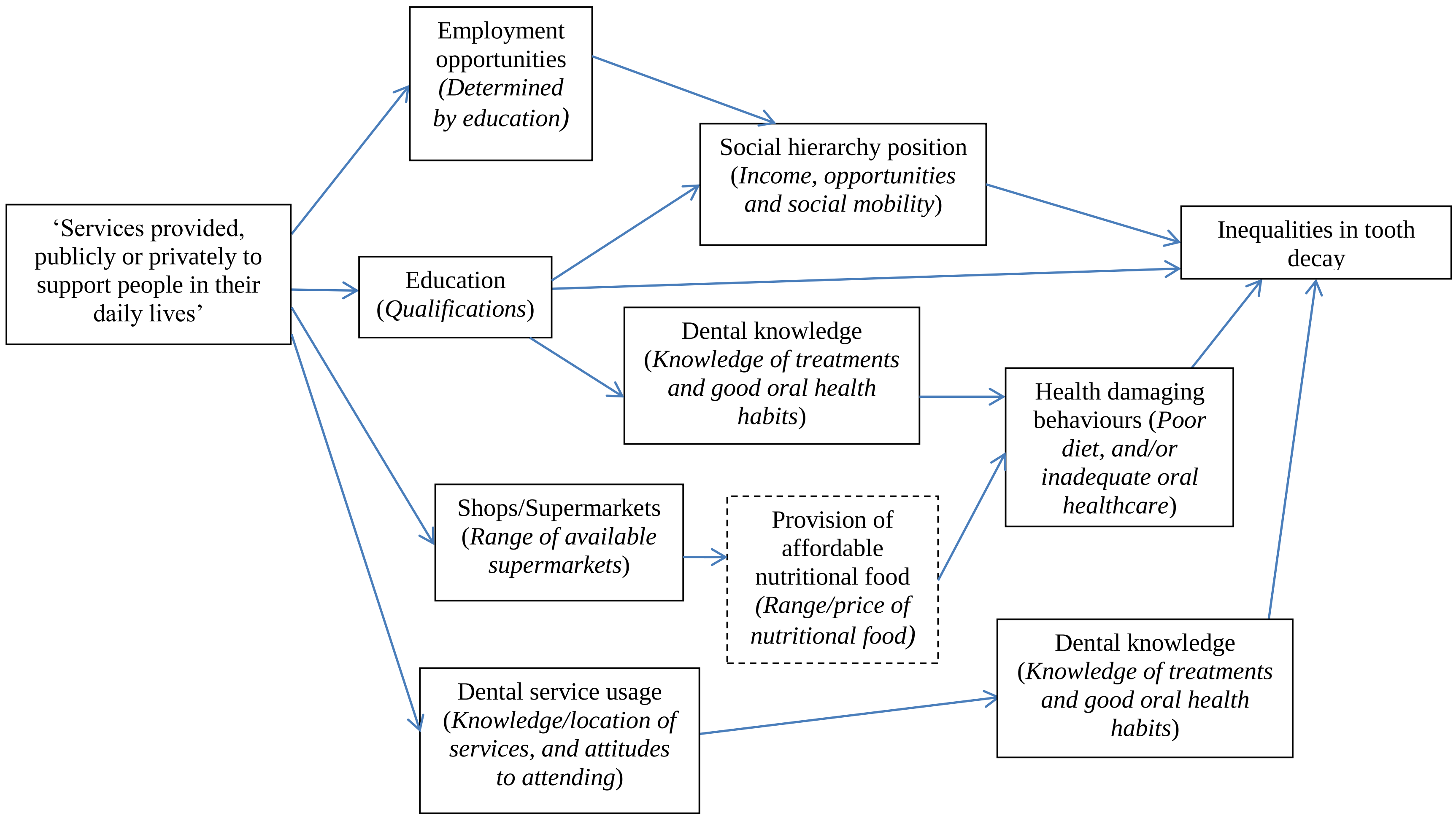
Indeed, a systematic review by Costa et al. (2012) revealed that occupations of higher social standing were linked to lower levels of decay. The importance of employment to tooth decay is echoed by Roberts-Thomson and Stewart (2008), who demonstrated that unemployment was a risk factor for decayed surfaces, while welfare receipt was a risk factor for total caries experience. Tellez et al. (2006) have also found a link between unemployment and higher levels of tooth decay. Through processes such as education and employment, individuals and families can be set on detrimental life course trajectories. The importance of life trajectories in relation to oral health have been well established by studies such as Thomson et al. (2004), and Poulton et al. (2002). Both education and employment are also highly correlated with income levels, which in turn have been shown to be important for tooth decay levels (Celeste et al, 2009; Geyer et al, 2010). Unemployment has also associated with less beneficial oral health related behaviours (Al-Sudani et al, 2016), but due to this being included in Pathway 2 it was not repeated, as it may have given undue influence to this interaction in the simulation modelling.

Education has also been shown to affect oral health through dental knowledge and subsequent behaviours associated with this. Williams et al. (2002) demonstrated that parents from deprived backgrounds who had lower levels of education were less likely to have high levels of dental knowledge. Similar patterns were found by Chu et al. (1999) in Hong Kong. Education has subsequently been linked to oral health related behaviours, with poorer oral health habits being more common among those with lower levels of education, and those from more deprived backgrounds (Sabbah et al, 2009).

A third relevant theme is the presence of food types in neighbourhoods. The work of Barker et al. (2008) demonstrates the importance of food availability to individuals and families in particular, while oral health focused studies have drawn similar conclusions, and emphasised the importance of the availability of such items within the local environment. Mobley et al. (2009) suggest that areas classified as being of lower socio-economic status are more likely to be served by stores ‘offering high-energy, low-nutrient–dense foods’ (p.412), rather than healthy produce such as fruits and vegetables. Lower nutrient intake has been shown to be detrimental to oral health (Moynihan and Petersen, 2004). This point is echoed by Fonseca’s (2012) review of the evidence surrounding childhood development and oral health, which stated that lower income areas tended to have fewer chain supermarkets. This limits the amount, and types of produce available in these areas, which tend to be located further from shops with more varied resources. Compounding matters, these resources are inevitably more expensive.

A fourth consideration is the presence of dental services. While dental services may be present within neighbourhoods, the literature has shown differences in the usage of such services, often differentiated by socio-economic status. Eckersley and Blinkhorn (2001) demonstrated that parents of children from more deprived backgrounds were less likely to take their child to a dentist regularly, and more likely to attend when symptomatic. Evidence also points to a potential uneven distribution of dental services between areas, pertaining to Hart’s inverse care law (1971). Jones (2001) supports this claim through their study of child dentist registrations and deprivation scores, while older accounts point to historic disparities in service distribution (Cook and Walker, 1967). Landes and Jardin (2010) note that up until the dental contract changes of 2006, practices could set up where they wished. Their study in County Durham demonstrated that a higher proportion of the deprived community accessed practices in deprived areas, which also tended to be smaller in size. However, the work of Macintyre and colleagues (2008) found contrary evidence in their study of resource and facility distribution in Glasgow, finding that dental practices were not patterned by deprivation. Aida et al. (2008) also showed that the number of dentists was not significantly associated with dmft scores. Thus while the evidence for the importance of the presence of these services is not universal, it would seem that either a lack of access to services, or negative attitudes towards regular attendance can contribute to unequal attendance rates between different groups. Figure 5 visually represents the pathways associated with this literature.

Figure 5 – Pathway 3: pathways for the 'Services provided, publicly or privately to support people in their daily lives'



As with the previous pathway the ‘provision of affordable nutritional food’ variable was to be represented by the number of servings of fruit and vegetable each day, however again this was not possible due to difficulties in combining the samples from the ADHS and US. This was instead replaced with a dietary variable based on sugar consumption, as this tied in with the variable for the ‘health damaging behaviours’ construct (sweet consumption) that finished the pathway.

This set of pathways contained the most neighbourhood based indicators of the five covered by the framework. Employment was represented by income estimates at the MSOA level. While technically an aggregated individual level variable, this data is not available at the individual level within the UK, and thus the estimates represent the average weekly income for each MSOA. Income is also far too important a variable to leave out given its influence, which was demonstrated in the literature. Income was chosen over the location of employment opportunities as these would likely be too numerous in a large urban area, which may bias the simulation modelling that tested the pathways. The location of education facilities and those offering apprenticeships for young adults were used to represent education, while shop and supermarket locations were geocoded to represent their locations within the city. A similar approach was taken for all of the dental surgeries in the city, the locations of which were kindly provided by Kate Jones, National Consultant in Dental Public Health at Public Health England.

Adequate data coverage was provided by the ADHS for the remaining three individual level variables. Social hierarchy position was represented using data from the NS-SEC classification (National Statistic Socio-economic Classification - Rose and O’Reilly, 1998), the primary social classification used in the UK. Dental knowledge was represented by the level of fluoride intake through toothpastes or other oral dentifrice, while health damaging behaviours was represented by the frequency at which the survey participants ate sweets. All of the above variables are shown in Table 5.

Table 5 – Variables for the ‘services provided, publicly or privately to support people in their daily lives’ pathway

Pathway	Pathway variable	Data/source	Present vs missing values
3	Employment	Model based income estimates (MSOA) – Neighbourhood Statistics	N/A
	Education	Location of further education providers in Sheffield - Sheffield City Council	N/A
	Shops/supermarkets	‘Supermarket locations data’ – GeoLytix website	N/A
	Accessibility of dentists	Location of all dentists in Sheffield - Public Health England	N/A
	Social hierarchy position	‘NS-SEC – Eight-class version’ (ADHS)	11380-0
	Dental knowledge	‘Fluoride level’ (in ppm - ADHS)	11108-272
	Diet/sugar intake	‘(D) high sugar intake – version 5’ (ADHS)	11377-3
	Health damaging behaviours	‘How often eat sweets?’ (ADHS)	11378-2

Pathway four - Socio-cultural features of a neighbourhood

Macintyre and colleagues (2002) described the features of the fourth pathway as including ‘the political, economic, ethnic and religious history of a community: norms and values, the degree of community integration, levels of crime, incivilities and other threats to personal safety, and networks of community support’ (p.131), with a particular positive being that it ‘adds an anthropological perspective to the socioeconomic, psychological, and epidemiological perspectives often used to examine area effects on health’ (p.130).

Of the factors listed above, two have so far been identified as being relevant within the oral health inequalities literature. The first of these is health behaviours, and associated attitudes towards such behaviours, which include a number of aspects such as diet and disease prevention techniques, and represent norms and values within groups. As mentioned in Pathway 2, lower income groups are more likely to have higher sugar intake (Warren et al, 2009) and lower levels of nutrient intake which have a protective effect on teeth (Moynihan and Peterson, 2004), leading to more fertile conditions for the development of dental caries. These health behaviours are not restricted to diet and consumption however, as they also relate to practices favourable to good oral health such as tooth brushing. Kumar et al. (2016) have shown the importance of tooth brushing, with higher numbers of carious lesions associated with infrequent brushing. Poutanen et al. (2006) have also demonstrated that parental behaviours tend to influence those of their child, particularly with regard to frequent tooth brushing and the use of other dentifrice, and that this varied based on the parent’s occupation. Adair et al. (2004) found that parental attitudes had a strong influence on oral health behaviours, interestingly also reporting that families from more deprived backgrounds felt less able to ensure that their child brushed twice a day, or stuck to recommended levels of sugar intake.

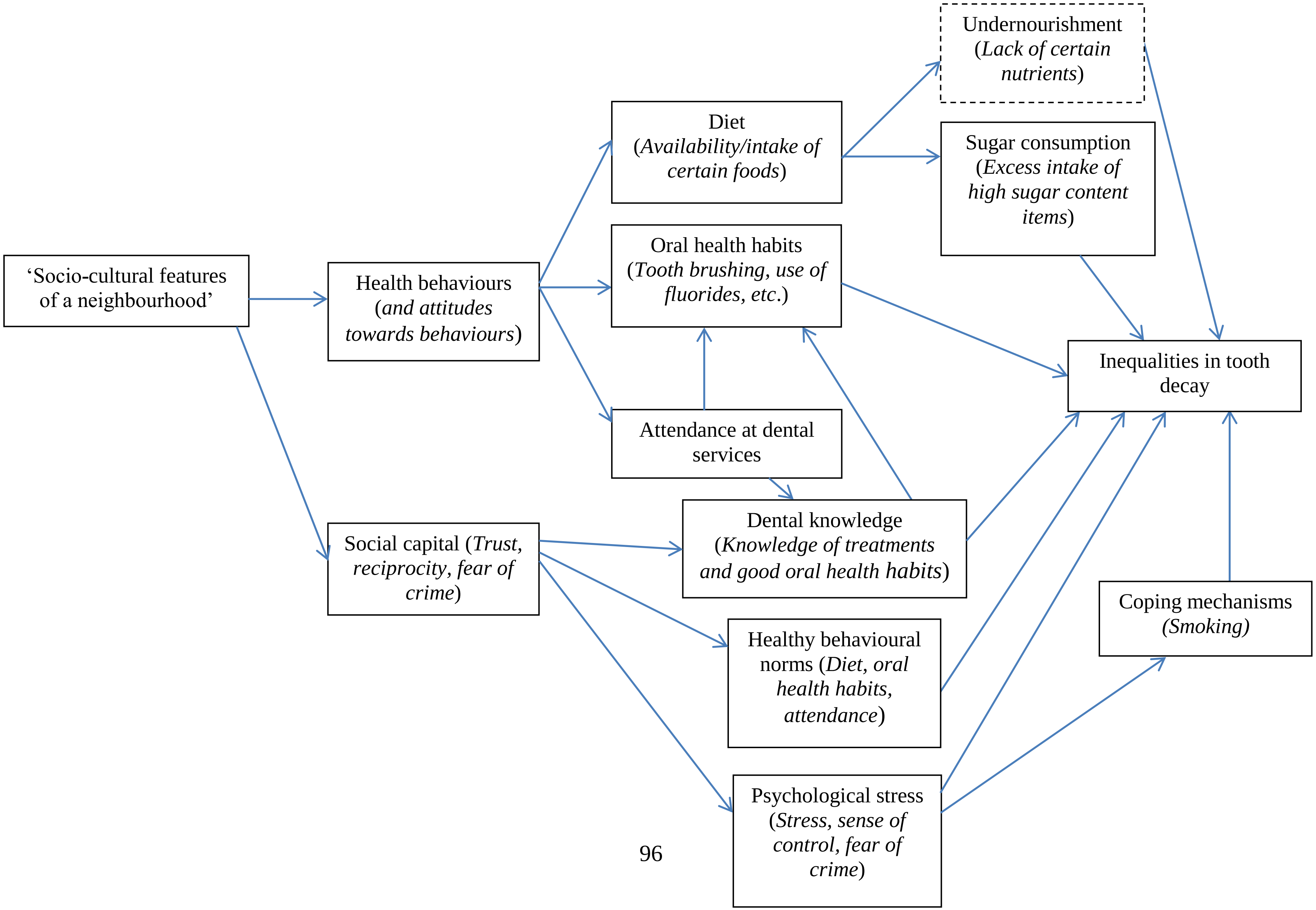
These attitudes constitute an important part of the role that health behaviours play in producing oral health inequalities. Adair et al. (2004) demonstrated that parental attitudes can influence oral health behaviours, and similar findings come from the work of Williams et al. (2002) and Chu et al. (1999), who both found that positive dental attitudes, as well as an increased level of dental knowledge, were associated with lower levels of disease. Interestingly, higher levels of education, and socio-economic status were associated with both positive dental attitudes and increased dental knowledge in both studies. Singh et al. (2013) also found education to be important, as social gradients in this were present in the clustering of health behaviours. This sentiment is echoed in the work of Riley et al. (2006), where education is shown to be key to such clustering, with better attitudes to attendance and oral health associated with higher levels of education. Sanders et al. (2006) however dispute the idea that poorer adults care less about their teeth than those from less deprived backgrounds. While dental visiting patterns do follow a social gradient, this work demonstrated that dental self-care trends did not, and highlighted wide variations in oral health behaviours.

With regard to preventative habits, Eckersley and Blinkhorn (2001) found that parents from more deprived backgrounds were more likely to change their oral health related habits upon receiving advice than those from less deprived backgrounds. The authors also found that attitudes played a significant role in dental attendance, demonstrating that, as mentioned in relation to Pathway 3, parents from more deprived backgrounds were more likely to attend only if symptomatic. Studies such as those carried out by Tickle et al. (2003) may suggest that these levels of attendance are important in determining other decisions. Their study demonstrated that patients, or those close to them, use their previous experiences of dental services to make decisions, based on the fact that they had gained confidence in a regularly used treatment. Residents of more deprived areas were more likely to opt for treatment based on this idea, while those from less deprived areas were happy to rely on the dentist's opinion (Tickle et al, 2003). Dental attendance from an early age has been demonstrated to be important (Listl, 2012), as it was found that inequalities in regular attendance are established in childhood, and persist throughout the life course. As a result, Listl (2012) states that, 'inequalities in later life-years may, thus, be relatively unresponsive to contemporaneous health policy interventions' (p.695).

Low income groups can face considerable barriers to dental services however (Wallace and Macentee, 2012), through their own opinion of dentists, but also through dentist's attitudes to low income groups, private practice and public dental benefits. Additionally, numerous studies have found that certain pre-existing socio-economic and demographic characteristics, deemed 'favourable', are associated with regular dental attendance (Guiney et al, 2011), while being male, aged 25-34, or having lone parent or immigrant status have all been shown to be predictors of less regular or symptomatic attendance (Muirhead et al, 2009). Donaldson et al. (2008) demonstrated the importance of attendance, stating that the socio-economic gradient in the number of sound teeth can partially be explained by attendance, which can be influenced by the effects that socio-economic position has on barriers to such attendance.

The second factor relevant to oral health inequalities within neighbourhoods is social capital, similar to the idea of networks of community support suggested by Macintyre and colleagues. Associations between social capital and related concepts have been found by a number of authors through social institutions (Tellez et al, 2006), non-volunteering and mistrust (Aida et al, 2011) and neighbourhood empowerment (Santiago et al, 2014). Self-rated oral health has also been associated with personal constraints and mastery (similar to personal control and self-efficacy) in those with lower incomes (Sanders and Spencer, 2005). Lida and Rozier (2013) have argued that social capital is beneficial to oral health as it can help spread knowledge of health promotion, as well as the maintenance of healthy behavioural norms. The authors also argue that it promotes access to local services and amenities, citing the example of shared resources to facilitate children's visits to the dentist. Finally, they state that social capital can also have psychosocial ramifications. As seen in Section 2.5, such psychosocial issues can have negative effects through the stress they cause (Boyce et al, 2010; Finlayson et al, 2010), and through coping mechanisms which may be employed to deal with life stresses, such as smoking (Reibel, 2003; Bernabe et al, 2014). Figure 6 is a visual representation of the pathways associated with the above literature.

Figure 6 – Pathway 4: pathways for the 'Socio-cultural features of a neighbourhood'



As with the previous two pathways, despite being initially included, the Understanding Society data on the amount of fruit and vegetables eaten per day had to be excluded again for technical reasons. This goes to highlight the potential importance of the variable given its appearance in three pathways, and future research should look at ways to combine survey samples so as to include this variable.

Health behaviours were represented by the 'years of lost life score' from the IMD Health Domain (2015). While referring more to general health than oral health, this variable was chosen as poor health choices are likely to shorten life expectancy, and the idea of co-morbidities posits that poor oral health may occur alongside the presence of other diseases and conditions (Bailey et al, 2004; Chroinin et al, 2016). One issue with this variable is that it could be considered aggregated individual level data, however data limitations meant that there were no suitable alternatives that were as relevant to the concept. Social capital was included through the use of additional IMD data, this time the score for the Crime Domain, including information on violence, burglary, theft and criminal damage per 1000 residents at the LSOA level (c1,500 people). While there are many possible variables that could be used to represent the wide ranging concept of social capital, few are publicly available at the small area level. Crime was judged appropriate for use as a proxy for community cohesion, and similar measures of social disorder such as homicide rate have been used in previous oral health related studies of social capital (Pattussi et al, 2001). While social capital can be seen as an individual level variable, Santiago et al. (2014) have argued for its use as a neighbourhood level indicator, stating that 'contextual or collective social capital emphasizes the resources that can be built collectively by individuals who are socially interconnected aiming to achieve collective goals, and has been evaluated and studied both in local levels of aggregation, such as neighborhoods, Census tracts or neighborhoods, and in broader levels, such as municipalities, states or countries' (p.17).

Data coverage from the ADHS for the fourth pathway was adequate to cover the remaining individual level constructs. Diet was represented by cake consumption, which while perhaps not the most accurate indicator of diet, was selected due to the exclusion of the Understanding Society data, leaving a smaller range of diet related variables to choose from in the ADHS. This was therefore the most appropriate variable available. Tooth brushing habits were represented using data on the frequency at which participants brushed their teeth each day, while attendance at dental services was included through a variable on the frequency of visits to the dentist. This differed to the dental attendance variable in Pathway 2, as it related to frequency in months and years, rather than general habits (i.e. regular check-up or only when symptomatic). Acquired dental knowledge was represented by an individual's fluoride intake, measured in parts per million (ppm), with a higher score being preferable. Healthy behavioural norms were represented by the 'general dental attendance' variable, and is the only time a variable was used twice for different concepts, having also been used for the 'preferential health habits' in Pathway 2. This goes to show the difficulties of populating a framework from one data source. It should also be pointed out that this differs to the attendance variable used for the 'attendance at dental services' construct (mentioned above), and that the same variable has not been used twice in the same pathway. As with Pathway 2, psychological issues were represented by how often in general an individual had felt psychologically tense, while sugar consumption was again measured using the 'highsug' sugar consumption variable. Finally, coping mechanisms was again represented using data on advice about giving up smoking. Details of these variables are shown in Table 6.

Table 6 – Variables for the 'socio-cultural features of a neighbourhood' pathway

Pathway	Pathway variable	Data/source	Present vs missing values
4	Health behaviours	IMD Health Domain (2015) – Years of potential life lost (LSOA)	N/A
	Social capital	IMD Crime Domain (2015) – contains violence, burglary, theft and criminal damage per 1000 (LSOA)	N/A
	Diet	'How often eat cakes' (ADHS)	11378-1
	Tooth brushing habits	'(D) number of times brush teeth per day - 3 groups' (ADHS)	10538-842
	Attendance at dental services	'How often do you go to the dentist' (ADHS)	11172-208
	Dental knowledge	'Fluoride level' (in ppm - ADHS)	11108-272
	Healthy behavioural norms	'General dental attendance' (ADHS)	11342-38
	Psychological issues	'Psychological discomfort – self tense' (ADHS)	11378-2
	Sugar consumption	'(D) high sugar intake – version 5' (ADHS)	11377-3
	Coping mechanisms	'Ever been given advice on giving up smoking' (ADHS)	10774-606

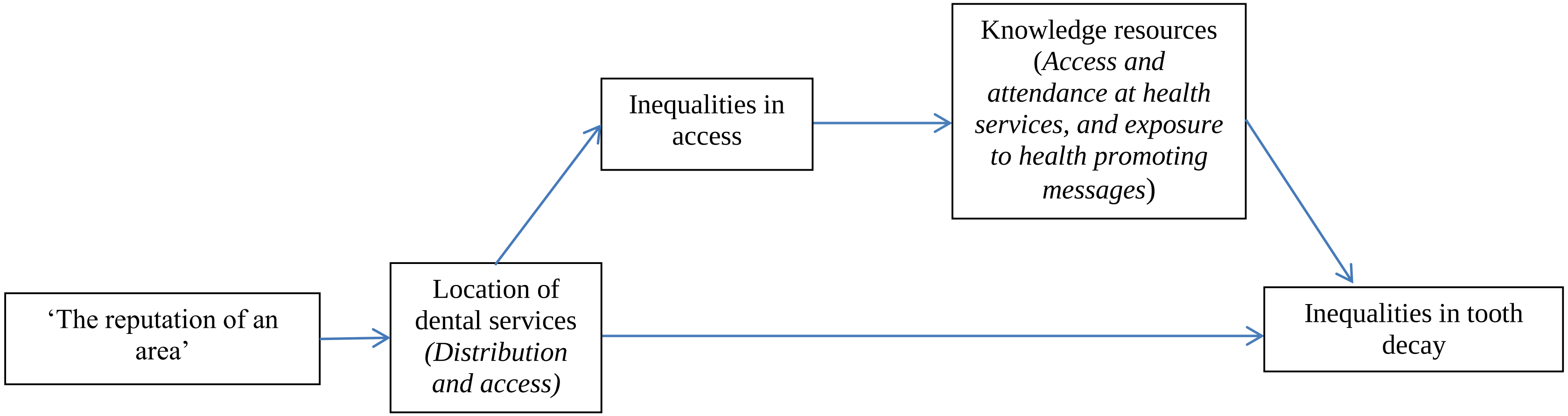
Pathway five – The reputation of an area

The fifth pathway looked at how an area's reputation can influence the decisions of planners, amenity providers and potential investors to locate in an area (or not), thus affecting local infrastructure. An area's reputation may also influence the morale of those who live there, as well as who moves in or out.

The initial set up of services in an area is an important consideration in this regard, and in relation to oral health perhaps the most important of these theoretically is the location, or lack thereof, of dentists (both NHS and private) within a given area. Lang et al. (2008) for example have stated that 'capitation systems may make deprived neighbourhoods unattractive settings for dentists' (p.477). Landes and Jardin (2010) would seem to add to the evidence of inequalities in the distribution of dental services by area deprivation, additionally stating that up until 2006 dental practices could locate wherever they wished. Other studies have pointed to the uneven distribution of dental services (Cook and Walker, 1967; Jones, 2001) based on socio-economic status, although it should be noted that not all studies found this to be the case (Macintyre et al, 2008). It would seem, however, that the presence of dental services within an area could provide more people with the opportunity to improve their oral health, perhaps through acquiring increased knowledge on attendance and preferential oral health behaviours (Tickle et al, 2003; Hill et al, 2013). As mentioned in relation to Pathway 3, there also seem to be inequalities in the distribution of different types of dental services (e.g. NHS or private), as several studies have identified problems in registering for NHS dental care in the south of England, possibly due to the rise in private dentistry in this region (Hancock et al, 1999; McGrath et al, 2001).

The dynamics of neighbourhood residential mobility and turnover, and the effects of neighbourhood based self-esteem in relation to dental outcomes do not appear to be covered within the literature, while there is also little theory as to how the actions of planners and amenity providers impact on oral health beyond the distribution of dental services. These are areas that would benefit from future research. Figure 7 represents the literature in this pathway visually.

Figure 7 – Pathway 5: pathways for 'The reputation of an area'



This pathway only contained three variables in total, one of which was the locational data of the dental surgeries used in Pathway 3. Data coverage from the ADHS was not an issue for the remaining two variables. Data for inequalities in access was represented by a variable which questioned how participants rated access to their local dental surgeries, while acquired dental knowledge was again represented by the level of fluoride intake per participant in parts per million (ppm). These data are presented in Table 7.

Table 7 – Variables for 'The reputation of an area' pathway

Pathway	Pathway variable	Data/source	Present vs missing values
5	Accessibility of dentists	Location of all dentists in Sheffield - Public Health England	N/A
	Inequalities in access	'Rate access includes dk's' (ADHS)	11195-185
	Acquired dental knowledge	'Ever been given advice about frequency of visits' (ADHS)	11153-227

3.6. Conclusions

The operationalising of the pathways inspired by the framework of Macintyre and colleagues (2002) was a key process, as the data attached to these were used to help represent the pathways in the agent-based models used in this research. These, in turn, tested the theory presented in this chapter. As Chapter 4 will demonstrate, the individual level data from the Adult Dental Health Survey were included in a spatial microsimulation model, which created representative individual level data for Sheffield, complete with associated characteristics and variables from the pathways. This, along with the neighbourhood based data, formed the basis of the agent-based models, representing the types of people that live in particular neighbourhoods, as well as the features and characteristics they would encounter and live with in everyday life. Chapter 4 will outline the process behind the spatial microsimulation modelling, before Chapter 5 describes the creation of the agent-based models.

This chapter, and the introduction of the main theoretical framework used in this research, helped to address the first of the research objectives for this thesis, set out in Section 1.2.

- To identify theoretical pathways by which neighbourhoods influence tooth decay

The coupling of the literature on tooth decay with that from the wider oral health inequalities field, and health inequalities more generally, with an organising framework allowed for a clearer idea of the main pathways associated with the effects of neighbourhood environments on tooth decay. The operationalisation of these pathways also aided the empirical work that will be covered in Chapters 4 and 5, providing real world data which could be used in the simulation modelling.

Overall, this chapter has demonstrated how the application of a neighbourhood based theoretical framework to the literature on tooth decay can help to create testable pathways by which the neighbourhoods people live in may influence this outcome. Despite some concerns around the operationalisation of some of the theoretical constructs, this process created a clearer picture of the potentially important interactions taking place. The flexibility of the theoretical framework is a positive as it means that such work can be replicated for other oral health outcomes, and health outcomes in general, while its wide ranging concepts allow for a comprehensive approach to neighbourhood based interactions between individuals, groups and their environments. There were a number of limitations noted throughout the chapter, including the construction of some of the pathways, and issues with operationalisation. These will be discussed in further detail in Chapter 6.

Chapter 4 – Spatial microsimulation modelling

‘Inequality is the cause of all local movements’ (Leonardo Da Vinci – 1452-1519)

4.1. Introduction

As mentioned in Section 2.3, spatial microsimulation is concerned with the creation of simulated population microdata sets at the small area level. The need for a technique such as this arises due to population microdata not being widely available in the UK. Even where such datasets do exist, such as the Sample of Anonymised Records (UK Data Service, 2011a) which consist of Census microdata, these are restricted to small samples and are not available at the small area level. Issues of cost and confidentiality are the key reasons for this. National level representative surveys could be considered a form of microdata, however such data is often only collected or available at large geographical scales. This lack of available data makes spatial microsimulation a valuable method for this research, where a representative synthetic population is required to test the effects of theoretical pathways, and form the basis of the agent-based models that will do this.

4.2. Which spatial microsimulation method?

Also highlighted in Section 2.3 was the number of different types of spatial microsimulation methods, which included probabilistic techniques such as simulated annealing (SA) and combinatorial optimisation (CO), as well as deterministic techniques such as iterative proportional fitting (IPF). Tanton (2014) has stated that spatial microsimulation techniques could broadly be categorised as those that select individuals from microdata to fill geographical areas, and those that adjust the original weights of the microdata files. Probabilistic methods such as combinatorial optimisation are an example of the former, and work iteratively by selecting random samples of households or individuals from the microdata, before assessing the effects of replacing one individual or household on the overall error. If this move improves the error the change is accepted and the next iteration begins, otherwise another household or individual is selected to replace it. This process repeats iteratively with the aim of improving the overall fit. An example of this is the 'hill climbing' approach, which, despite its speed, also has drawbacks in that it 'can quickly become trapped in suboptimal solutions' (Williamson et al, 1998 - p.794). Simulated annealing is another version of the CO approach, which (unlike the hill climbing approach) allows the algorithm at the heart of the method 'to climb down from sub-optimal solutions by allowing changes to the combination being optimised even if they make the solution worse' (Tanton, 2014 - p.10), essentially allowing it to take backward steps in the optimisation of the overall model fit.

Deterministic approaches such as IPF are examples of the latter type of model depicted by Tanton (2014), and are typically used where there is no access to suitable small area microdata (Birkin and Clarke, 1988). Norman (1999) describes IPF as a reweighting mechanism that adjusts (or re-weights) a table of data to fit row and column constraints through iterative calculations. This method will be described in greater detail later in this chapter. Dynamic spatial microsimulation models form a third category, which can take the form of probabilistic models (such as Monte Carlo simulations) which project individuals in the simulated dataset into the future using event probabilities, or implicit models which make use of small area projections before applying static spatial microsimulation methods to create the future microdata (Ballas et al, 2005a). However, as agent-based modelling was used to dynamically simulate populations in this research dynamic spatial microsimulation models were not considered for use.

Due to the different approaches of the above methods, there has been debate about which of the static spatial microsimulation approaches is the more preferable technique. For example, Williamson et al. (1998) have demonstrated that SA showed greater accuracy than hill climbing approaches. Further, Williamson (2013) has shown that probabilistic techniques including CO have tended to produce estimates with less bias and variance than deterministic models, while Harland et al. (2012) found SA constrained outcomes more accurately than deterministic approaches. However, Tanton et al. (2014) found no appreciable differences between CO and deterministic approaches. Harland et al. (2012) have also acknowledged that each approach has different strengths, and can be useful depending on the nature of the project at hand. With this in mind, IPF was chosen as the preferred method for this research. Despite the potential advantages of the other approaches with regards to accuracy, the deterministic nature of IPF means that consistent results are produced each time the model is run (Ballas et al, 2005c), as opposed to probabilistic models where the results of different runs can vary, and sometimes require optimisation techniques.

IPF is essentially a reweighting process which is carried out through the use of ‘constraint variables’. Constraints are variables that the existing microdata (more than likely national level survey data) are reweighted against. These usually come in the form of aggregated Census totals for small areas. The IPF technique uses the constraint variables as a guide when reweighting microdata, in a process where individuals and their associated characteristics are ‘weighted up’, or replicated, depending on how representative they are of the characteristics of a given area. A basic example of the principle behind the IPF process is demonstrated below in Figure 8 (Hunsinger, 2008).

The table on the left represents the microdata before the IPF procedure, and the table on the right after. The column and row totals (or marginal totals) around the edge of the table represent the Census population totals, while the data in the table represent individuals. It can be seen that the individual in the top left corner of the first table (represented with a 1) has been reweighted to 1.51 in the second table, indicating this individual is more representative of the characteristics of the area in question, and has been ‘weighted up’. Conversely the individual in the bottom left corner of the table has been weighted down (from 6 to 5.28), indicating that their characteristics do not match up as well to those of the area as a whole. This process continues iteratively, with the goal of bringing the individual level data further in line with the Census constraints after each iteration. Essentially, the technique is allocating, or replicating, individuals from the survey data to geographical areas where their characteristics are a good match. A worked example is provided in the next section.

Figure 8 – The iterative proportional fitting (IPF) process (Hunsinger, 2008)

		Marginal (column) totals		
		11	**9**	**8**
Marginal (row) totals	**5**	1	2	1
	15	3	5	5
	8	6	2	2

		Marginal (column) totals		
		11	**9**	**8**
Marginal (row) totals	**5**	1.51	2.31	1.18
	15	4.2	5.35	5.45
	8	5.28	1.34	1.37

4.3. A worked example of the IPF technique

The equation presented below from Ballas et al. (2005a) is provided to give a basic demonstration of the IPF process (see Equation 1).

Equation 1 - IPF formula

$$ni = wi \times sij/mij$$

As stated in Section 4.2, IPF is a reweighting technique which adjusts survey data to match Census population totals. Equation 1 represents the basic mathematics behind this, where *ni* is the new weight created for an individual, *wi* is the initial weight of an individual, *sij* is element *ij* of table *s* (small area Census data), while *mij* is element *ij* of table *m* (the survey data).

Table 8 – Hypothetical survey data

Individual	Education	Cars/vans owned	Initial weight
Individual 1	Other qualification	2 cars or vans in household	1
Individual 2	Degree or higher	2 cars or vans in household	1
Individual 3	Degree or higher	1 car or van in household	1
Individual 4	Other qualification	2 cars or vans in household	1
Individual 5	Other qualification	1 car or van in household	1

Table 9 – Hypothetical small area Census data

	Degree or higher	Other qualification
1 car or van in household	5	1
2 cars or vans in household	6	4

Table 10 – Survey data in cross-tabulated format

	Degree or higher	Other qualification
1 car or van in household	1	1
2 cars or vans in household	1	2

Table 8 represents the survey data, while Table 9 represents the small area data from the Census, and Table 10 the survey data in a cross-tabulated format. Looking at the population without degrees and with two cars (or vans), it can be seen that there are two such individuals in the survey data (individuals 1 and 4 - see Table 10), while there are four individuals with these characteristics in the Census data (See Table 9). Thus, by applying the formula listed above (see Equation 1), the original (or initial) weight (1) is multiplied by the total from the Census (4), over the total from the survey (2). This can be seen in the first line of Table 11, and leaves those without degrees and with two cars with a new weighting of 2. Other combinations of education and car ownership have also been calculated as further examples.

Table 11 – Reweighting the survey data

Individual	Education	Car ownership	Weight	New weight
Individual 1	Other qualification	2 cars or vans in household	1	1 x 4/2 = 2
Individual 2	Degree or higher	2 cars or vans in household	1	1 x 6/1 = 6
Individual 3	Degree or higher	1 car or van in household	1	1 x 5/1 = 5
Individual 4	Other qualification	2 cars or vans in household	1	1 x 4/2 = 2
Individual 5	Other qualification	1 car or van in household	1	1 x 1/1 = 1

This is a simplified example of the process that IPF uses to reweight individuals from the survey data. Essentially the process could be thought of as weighting individuals up or down depending on how representative their characteristics are for a given area.

Those without degrees and with two cars or vans have a new weight of two, indicating that they are more representative than their initial weighting of '1' suggested, while also being less representative of this area's population characteristics than those with degrees who own two cars, whose weight increased to six. This process then loops through the rest of the constraints in order, continuing the process and multiplying the new weight by the one produced in the previous stage. The process then loops back to the first constraint and begins the IPF process again (depending on the number of assigned iterations). The above data represent a hypothetical scenario with a small number of individuals. Details and discussion on how to run a full IPF model on a larger sample of individuals, with additional constraint variables will be covered in Section 4.5.

4.4. Data sources

IPF requires two sources of data, the first being an existing source of microdata that can then be reweighted to fit the characteristics of small areas geographies. In line with the operationalising of the theoretical pathways, the Adult Dental Health Survey (2009) was chosen as the main source of microdata used in this analysis. The comprehensiveness of this survey has been described in Section 3.4, leaving it as the obvious choice for this research. The second source of data (used for the constraint variables) was the 2011 Census of Population for England and Wales (Office for National Statistics, 2011a). The Census of Population is a decennial survey of the whole of England and Wales that was first conducted in 1801, and is the most complete source of demographic information available about these countries (separate Census' are conducted for Scotland and Northern Ireland). Census data is readily and freely available online (Neighbourhood Statistics, 2011; NOMIS, 2011), and at a variety of spatial scales, making it an invaluable and flexible tool for research. It should be noted that the Neighbourhood Statistics (2011) website has been shut down since this thesis was finished.

For the purposes of this research, the constraint data were collected at the Lower Layer Super Output Area level (LSOA). LSOAs are statistical geographical units (Office for National Statistics, 2011b) designed to improve research at the small area level through their consistent unit size, and have an average population of 1,614 per LSOA (an average of 672 households). As of the 2011 Census, the city of Sheffield was made up of 345 of these spatial units. This geographical level was chosen based on ideas taken from the theoretical work presented in Chapter 2. Certain studies made reference to the way that people saw their neighbourhoods as their immediate surroundings (Haynes et al, 2007), rather than larger administrative areas, which were shown to be less predictive of neighbourhood based trends (Pampalon et al, 1999). These areas were therefore seen to represent smaller neighbourhood boundaries, while also being large enough to potentially include a number of local services. Middle Layer Super Output Areas (MSOAs) were also considered, but with an average of 7,787 people per unit (3,245 households), and with 76 of these areas in Sheffield, these geographical units were considered too large. Conversely, output areas (the smallest geographical scale at which Census data is published), which are aggregated to create LSOAs, contain on average 309 people (129 households) per OA. With Sheffield consisting of 1,817 of these units, these were considered too small for the purposes of this research. A visual comparison of the boundaries of these Census geographies is provided in Figures 9 (OAs), 10 (LSOAs) and 11 (MSOAs). These data were collected from the UK Data Archive's Census boundary data collection (UK Data Service, 2011b).

Figure 9: Output Areas (OAs) within Sheffield

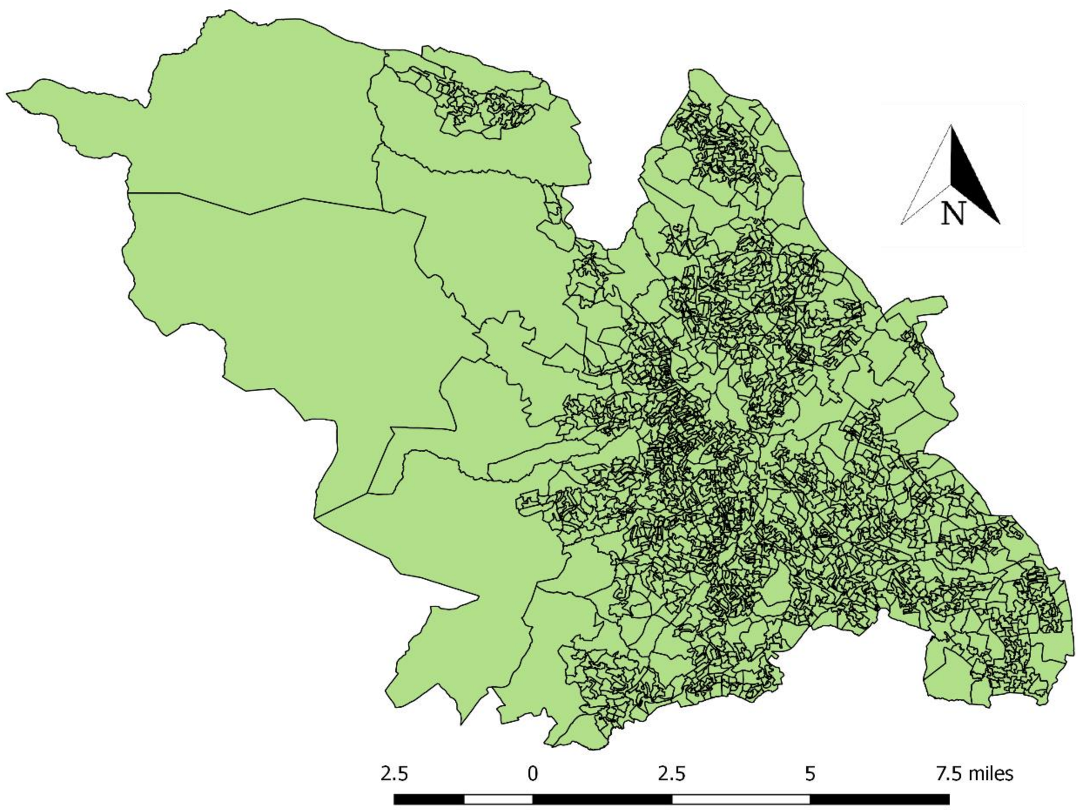
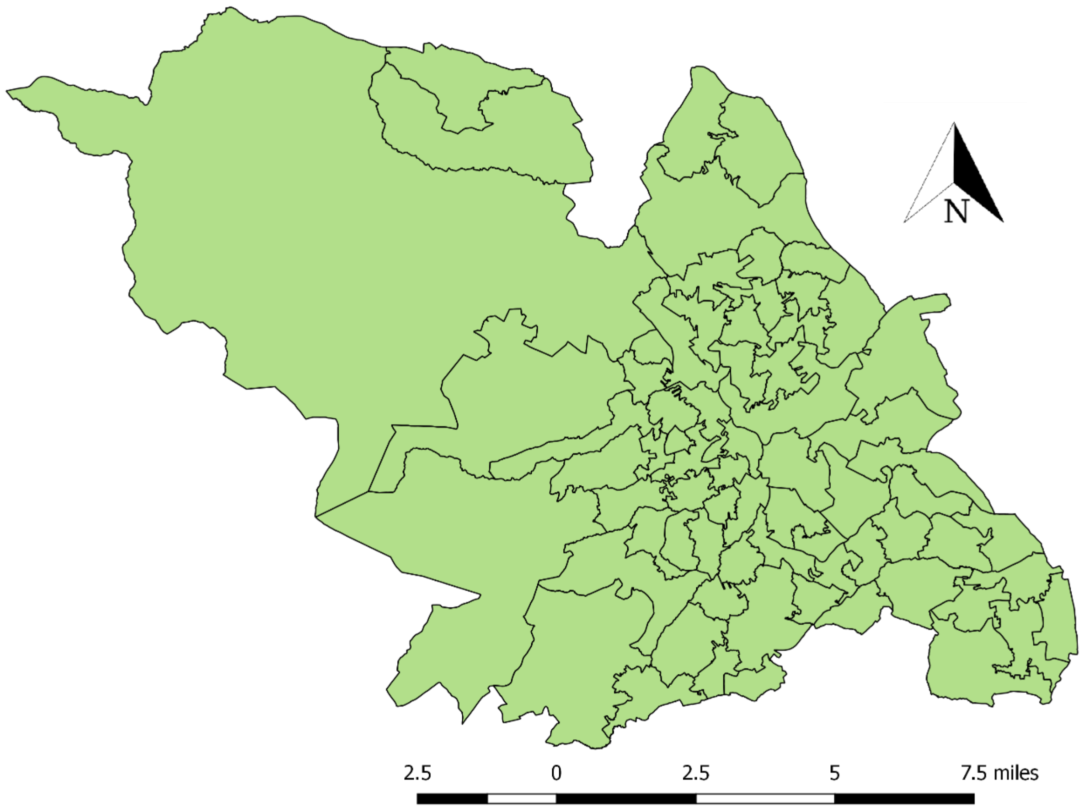


Figure 10: Lower Layer Super Output Areas (LSOAs) within Sheffield

Figure 11: Middle Layer Super Output Areas (MSOAs) within Sheffield



4.4.1. Choice of constraint variables

A key requirement when choosing the constraints is that the variables of choice are present in both the Census and survey datasets, and are comparable in some way. It may be that the data from both sources are directly comparable due to the way they have been collected, or that data manipulation may be required to re-categorise the two datasets so that they match. With regard to the ADHS, the following variables fulfilled the above criteria:

- Household size ('DVHsize')
- Sex ('Sex')
- Marital status ('xMarSta2').
- Age structure ('ageband5')
- General health ('QHealth1')
- Car ownership ('Car')
- Educational attainment (any educational qualification – 'EdAttn1')
- Educational attainment (degree level or below – 'EdAttn3')
- NS-Sec ('NS-SEC8')
- Ethnicity ('ethnicg')

Variables for tenure type ('Hhldr'), and long term illness ('LSIll') were present in the study, however the way they had been categorised and labelled meant that they could not be matched up to the corresponding Census data.

Household size refers to the number of household members present within the dwelling at the time of the survey. The sex variable was dichotomised as male or female, while the age data was split into 8 age bands, with 10 year intervals per band. Age data split into 3, 4, 5, 6 and 7 bands were also available, however the 8 band classification was chosen as this allowed for a more specific age grouping to be given to each member of the population. The general health data refers to self-reported general health, recorded on a five point Likert scale (ranging from 'very good' to 'very bad'). Car ownership was a dichotomous variable relating to whether a household owns a car/van or not. Unlike the Census data, the exact number of cars and/or vans that were owned was not included. The two educational attainment variables contained data on whether participants had educational qualifications or not (EdAttn1), and whether participants held a degree level qualification or not (EdAttn3). In the end it was decided 'EdAttn3' would be the more appropriate variable to use. This assumption was made on the basis that degree vs non-degree would offer a better differentiation within the population than looking at those with and without any qualifications. Therefore, 'EdAttn1' was dropped from the analysis. The NS-SEC data was split into 9 social classifications, having replaced previous versions based on social class and socio-economic position. Finally, the ethnicity variable asked participants to identify their respective grouped ethnicity.

Within the spatial microsimulation literature there seems to be no consensus on how to decide on a final set of constraint variables, with some advocating a statistical approach (Anderson, 2007), while other posit that a theoretical approach is the most appropriate (Smith et al, 2007). In this research it was decided that a combination of both theoretical and statistical approaches would be used, by first testing the significance of all of the available variables, before assessing the theoretical links between those that remained significant and tooth decay. For the statistical analysis the r-squared statistics for each of the significant variables were used to list the variables in reverse order (i.e. smallest r-squared value appearing first, the largest appearing last) to reflect the order in which the constraint variables would be applied in the microsimulation. It is generally accepted that applying the most influential variable last is the best approach (Anderson, 2007). The results of the statistical analysis can be seen in Table 12 below.

Table 12 - Statistical testing of the potential constraint variables

Variable	p-value	R2 value
Household size	0.001965	0.003042
Sex	0.00000000553	0.006107
Age	0.00122	0.003123
Marital status	0.00000000284	0.008043
Health	<0.0000000000000002	0.01401
Car ownership	<0.0000000000000002	0.02202
Educational attainment	0.0000000000519	0.007788
NS-SEC	0.0000000000000768	0.01311
Ethnicity	0.0727	0.0004122

The statistical analysis showed that the 'ethnicity' variable was not significant, and was therefore excluded from the analysis. Despite their statistical significance, marital status and household size are not theoretically linked to tooth decay, and so were both excluded from the analysis as well. Regarding the latter, within the oral health literature a number of studies have indicated that household overcrowding may be associated with tooth decay (Antunes et al, 200; Antunes et al, 2002; Pereira et al, 2010), however the number of people in a household is not necessarily the same as overcrowding. This left the following set of six constraint variables, which are listed below in reverse order according to their R^2 values, the order in which they were applied in the model.

- Age structure ('ageband5')
- Sex ('Sex')
- Educational attainment (degree level or below – 'EdAttn3')
- NS-SEC ('NS-SEC8')
- General health ('QHealth1')
- Car ownership ('Car')

Age has been shown to be important for tooth decay up until the age of 50, when tooth loss is more of a problem (Steele et al, 2015) and perhaps has the weakest theoretical linked with tooth decay of the six constraints. However, as a strong demographic indicator this variable would likely help to constrain the model more accurately. The rest of the variables all have theoretical as well as statistical links with tooth decay. Differences in sex (or gender) have clearly been demonstrated between males and females, with females tending to have higher levels of decay (Lukacs and Largaespadad, 2006) despite having better oral health related behaviours overall (Bertea et al, 2007). Tooth decay has also been strongly linked to educational attainment (Brennan et al, 2007; Momai-Homata et al, 2012), as well as employment (Roberts-Thomson and Stewart, 2008; Costa et al, 2012) and socio-economic position (Hobdell et al, 2003; Schwendicke et al, 2015) which are perhaps the closest concepts to the NS-SEC classification. Oral health outcomes are also associated with more general health outcomes (Bailey et al, 2004; Chroinin et al, 2016), so using a (self-reported) measure of general health seems appropriate in this case. Finally, material circumstances have also been linked to tooth decay (Nicolau et al, 2005), as well as with oral cancers through the Townsend Index (O'Hanlon et al, 1997) which uses car ownership as a constituent variable. The use of car ownership as a constraint variable therefore fits within the remit of the dental literature.

4.4.2. Data transformation

Having decided on the constraint variables, data transformation was required for the Census data on car/van ownership and the NS-SEC classification, due to the mismatch between the data classification in the Census and the ADHS. The data for car (and van) ownership was presented as a binary ('yes', 'no') variable in the ADHS, whereas a count was recorded within the Census. Regarding the NS-SEC data, the counts associated with these did not always match up to the area's Census population totals, due to some individuals being included in a 'not classified' category. Constraint variables must match up to the area's population totals or the sums in the IPF process will not work. Thus manipulation of the NS-SEC data was needed to solve this.

Initial data for the car/van variable was created by combining the columns of data from the Census referring to households that owned 1, 2, 3 or 4(+) cars/vans into one column labelled 'has_car', while the column for those who did not own a car/van was labelled 'no_car'. This meant that both data sources were now in a binary format. Table 13 demonstrates a further problem (and solution) with the car/van data. This problem arose as the rest of the Census data counted individuals, whereas the data on car ownership was collected by household, which meant that the sum for the car variable (element 'a' in Table 13) was always lower than the sum of the other variables in a given area (b). Such an issue can affect the accuracy of a model. To counter this problem, the difference between the population totals (b) in the area and the car total (a) was calculated (c). This number was then split between the two categories ('has_car', 'no_car') proportionally to match the binary nature of the ADHS question. In the worked example below, for LSOA 'E01007825' (an area along Abbeydale Road, south of the city centre) the car total was 53 less than the population total. The proportion of people who owned a car in this area was 92.23% (d) with 7.77% (e) not owning a car/van. This equates to 1069 and 90 people respectively. The difference of 53 was split proportionally between these two categories, leading to 49 people being added to the 'has_car' column, and 4 to the 'no_car' column. This resulted in new figures for the area of 1118 (f) and 94 respectively (g), which summed to the overall population total for that LSOA.

Table 13 – Data transformation of car data from the Census

Variable	Value
LSOA	E01007825
Number of people with cars	1069
Number w/out cars	90
Total (a)	1159
Population total for area (b)	1212
Difference between totals (c)	53
Proportion with car (d)	0.922346851
Proportion w/out car (e)	0.077653149
new_hascar (f)	1069 + 49 = 1118
new_nocar (g)	90 + 4 = 94
New car total	1212

A similar issue was identified with the NS-SEC data. In order for the different NS-SEC categories in the Census to sum to the area totals the ‘not classified’ column needed to be included. In the ADHS those not classified into an NS-SEC category were coded as missing, and were thus not included in the data from the survey. This left the question as to what to do with those from the Census who were classified as ‘not categorised’. In a similar manner to the car data, these individuals were re-distributed among the remaining nine NS-SEC categories based on the proportion of people in each category in a given area. Often this led to decimal numbers, which are not appropriate when attempting to represent numbers of people in area, so the numbers were rounded to give integers. Due to this rounding, the population totals of some areas were out by +/- 1 or 2 people, so for each LSOA these ‘missing people’ were added to the NS-SEC category with the largest count. The rationale behind this was that adding these individuals to a column with a large number would have less of an effect, proportionally, than adding them to a column with smaller numbers.

Once this data transformation had taken place the constraint variables could be matched to the survey data from the ADHS. Table 14 demonstrates the final categorisation of the constraint data used in the spatial microsimulation modelling.

Table 14 – Final categorisation of constraint data

Variable	Groupings
Age	16-24, 25-34, 35-44, 45-44, 55-64, 65-74, 75-84, 85+
Gender	Male, Female
General health	'Very good', 'good', 'fair', 'bad', 'very bad'
Car ownership	'Yes', 'No'
Educational attainment	'Degree level or above', 'below degree level'
NS-SEC	1.1, 1.2, 2, 3, 4, 5, 6, 7, 8

4.5. Spatial microsimulation using R

The spatial microsimulation modelling was conducted using R, an open source programming language (R Core Team, 2015). R is an object oriented language, with a particular focus on statistics and graphical output. The object oriented nature of the language makes it highly suited for data analysis and manipulation, which are particularly useful for population analysis. The analysis conducted in this research was based on methods used in Lovelace and Dumont's (2016) book 'Spatial Microsimulation in R'. The source code used for the spatial microsimulation in this study is available in full online (https://github.com/tombro1987/SMStoothdecay) with notations.

The main function used for the analysis came from the 'ipfp' package (Blocker, 2015), which is a fast way of implementing the IPF procedure using the C programming language. In order to work, the function requires three types of input data – a data frame containing aggregated Census data for the geographical area of interest (the constraint variables – or ‘wide’ data – Figure 12); a data frame containing individuals from the ADHS, with associated constraint data, and variables that the user wishes to ‘create’ in their new dataset (or ‘long’ data – Figure 13); and a data frame of individuals from the survey containing constraint data in a Boolean (or model matrix) format (Figure 14). This means that a ‘1’ is present in the column which is relevant to that individual, while a ‘0’ is placed in the others. For instance, if an individual is 25 years old a ‘1’ will be placed in the ‘25-34’ age grouping, with zeros in the other seven age columns. The row sums of each individual were taken to make sure that these summed to 6, as there should be a ‘1’ present for each constraint variable. The column order of the aggregated Census data and the Boolean format survey data should be the same, and should preferably be in order of constraint application (i.e. least influential to most).

The reason for this extra Boolean formatted data frame is that the ADHS and Census data are not directly comparable in their original formats, and therefore this third data frame ‘flattens’ the individual level data in the ADHS, so that it matches the format of the Census (Lovelace and Dumont, 2016). It then becomes clear which categories an individual belongs to for each constraint variable, and the two data sets can then be compared.

Figure 12 – Constraint variables in R (each row represents the population total of an LSOA)

additional_variables_test.R × census.data ×

Filter

	X16to24	X25to34	X35to44	X45to54	X55to64	X65to74	X75to84	X85plus	Males	Females	High_qual	other_qual	nssec1_1	nssec1_2	nssec2	nssec3	nssec4	nssec5	nssec6	nssec7
1	156	164	211	213	237	210	148	67	656	750	343	1063	30	104	300	234	128	119	261	169
2	120	125	203	214	192	137	99	39	540	589	453	676	41	131	373	201	132	60	119	50
3	97	100	198	224	220	185	118	70	559	653	512	700	47	154	407	194	133	58	133	52
4	139	80	257	283	184	139	82	32	579	617	481	715	46	144	404	211	123	58	127	60
5	124	177	213	258	223	178	83	36	641	651	423	869	39	137	351	198	124	103	205	111
6	176	215	251	279	161	126	106	45	609	750	377	982	21	104	325	197	98	120	245	185
7	151	338	244	220	151	112	58	27	649	652	462	839	34	127	368	197	115	88	198	127
8	147	99	207	226	188	200	154	106	627	700	286	1041	21	85	253	241	139	92	247	148
9	153	220	250	212	188	116	80	41	592	668	502	758	33	175	347	189	114	81	168	92
10	151	221	263	225	174	118	64	33	588	661	447	802	25	119	347	226	113	99	166	115
11	125	103	198	249	236	248	193	78	671	759	407	1023	32	126	376	249	154	96	214	135
12	112	139	226	252	217	187	159	61	610	743	487	866	40	143	374	213	128	73	193	145
13	115	141	180	188	195	176	175	57	586	641	207	1020	20	73	225	231	116	104	235	167
14	145	136	218	238	174	213	195	36	660	695	242	1113	33	87	258	253	121	131	227	197
15	159	196	217	231	179	148	74	33	583	654	174	1063	18	40	209	188	117	142	258	215
16	152	209	304	276	155	120	68	13	645	652	224	1073	21	45	272	258	115	129	251	176
17	172	166	295	231	173	157	91	45	630	700	216	1114	23	63	223	235	142	141	238	195
18	180	155	223	234	185	155	79	75	615	671	178	1108	23	38	211	189	110	123	286	247
19	152	170	223	207	186	154	110	96	627	671	130	1168	19	35	176	173	120	131	292	251
20	171	155	201	221	186	149	110	45	583	655	223	1015	29	82	236	243	113	114	225	157
21	136	161	184	191	154	146	78	29	536	543	126	953	10	30	163	138	73	109	241	210
22	215	206	245	196	178	140	74	60	596	718	145	1169	3	24	156	146	64	104	273	307
23	306	294	314	265	142	150	80	24	733	842	250	1325	16	58	239	219	99	140	365	256
24	196	153	211	237	201	213	62	20	626	667	171	1122	26	41	190	216	105	133	292	226
25	221	173	192	196	134	89	53	15	500	573	152	921	8	33	137	130	90	80	225	178
26	160	235	237	240	176	127	35	6	612	604	170	1046	25	52	210	171	120	143	242	196
27	167	201	253	207	166	191	82	39	643	663	184	1122	6	38	219	215	80	150	289	243
28	196	290	292	190	117	71	32	10	549	649	184	1014	16	39	163	156	103	82	262	211
29	266	393	278	234	190	110	40	10	740	781	264	1257	20	54	263	196	107	139	315	244

Showing 1 to 31 of 345 entries

Figure 13 – Participant data from the ADHS (each row represents an individual)

	Sex	ageband5	xMarSta	QHealth1	Car	EdAttn3	NSSEC	ethnicity	numdu98	CostDly	TPaste	CostTyp	PsycDisc	EvrAdSm	highsug5	Regular	fluoride	Sw
1	Female	65to74	Widowed	Very good	Yes	Degree OH	Neverworked	White Br/O	1	No	No	No	Never	Yes	Not high	Regular check-up	1000-1300ppm	
2	Male	25to34	Single	fair	No	Other qual	Routine	White Br/O	9	Yes	Yes	No	Occasionally	Never smoked	Not high	Only when having trouble	1350-1500ppm	
3	Female	35to44	Divorced	Good	No	Other qual	semi-routine	White Br/O	0	No	No	No	Never	Never smoked	Not high	Regular check-up	1350-1500ppm	
4	Male	75to84	Married	fair	Yes	Other qual	LowManProf	White Br/O	0	No	No	No	Never	Never smoked	Not high	Regular check-up	1350-1500ppm	
5	Male	55to64	Married	Good	Yes	Degree OH	LargeEmpHighMan	White Br/O	0	No	No	No	Never	No	High sugar intake	Regular check-up	1000-1300ppm	
6	Male	35to44	Married	Very good	Yes	Other qual	LowManProf	White Br/O	0	No	No	No	Fairly often	Never smoked	High sugar intake	Regular check-up	1350-1500ppm	
7	Male	45to54	Divorced	fair	Yes	Other qual	Intermediate	White Br/O	2	No	Yes	Yes	Never	Yes	High sugar intake	Regular check-up	1350-1500ppm	
8	Female	25to34	Single	Good	No	Other qual	LowerSupTechOC	White Br/O	0	No	No	Yes	Never	No	High sugar intake	Regular check-up	1350-1500ppm	
9	Male	65to74	Married	bad	Yes	Other qual	Routine	White Br/O	0	No	Yes	Yes	Never	No	Not high	Regular check-up	1350-1500ppm	
10	Male	55to64	Married	Good	Yes	Other qual	SmallEmpOAW	White Br/O	0	No	Yes	No	Never	Never smoked	High sugar intake	Regular check-up	1350-1500ppm	
11	Female	45to54	Married	Good	Yes	Degree OH	LowManProf	White Br/O	0	Yes	Yes	No	Never	Never smoked	High sugar intake	Occassional check-up	1350-1500ppm	
12	Male	45to54	Divorced	bad	No	Other qual	Routine	White Br/O	1	Yes	No	Yes	Never	No	High sugar intake	Regular check-up	1350-1500ppm	
13	Female	35to44	Married	Very good	Yes	Degree OH	LowManProf	White Br/O	0	Yes	Yes	Yes	Fairly often	No	High sugar intake	Occassional check-up	1350-1500ppm	
14	Female	25to34	Married	fair	Yes	Degree OH	LowManProf	Asian-Ind	0	No	No	Yes	Never	Never smoked	High sugar intake	Only when having trouble	1000-1300ppm	
15	Male	35to44	Married	Very good	Yes	Other qual	Intermediate	White Br/O	0	Yes	No	Yes	Hardly ever	No	High sugar intake	Regular check-up	1000-1300ppm	
16	Female	35to44	Seperated	Good	Yes	Degree OH	LowManProf	White Br/O	1	Yes	Yes	Yes	Never	No	Not high	Regular check-up	1350-1500ppm	
17	Male	16to24	Single	Good	Yes	Other qual	Neverworked	White Br/O	2	No	No	No	Fairly often	No	Not high	Occassional check-up	1350-1500ppm	
18	Male	45to54	Divorced	fair	Yes	Other qual	SmallEmpOAW	White Br/O	4	No	Yes	No	Never	No	High sugar intake	Only when having trouble	1350-1500ppm	
19	Female	25to34	Single	Good	Yes	Degree OH	semi-routine	Asian-Ind	0	No	Yes	No	Never	Never smoked	Not high	Only when having trouble	1350-1500ppm	
20	Male	16to24	Single	fair	Yes	Other qual	LowerSupTechOC	White Br/O	0	No	No	No	Never	No	Not high	Regular check-up	1000-1300ppm	
21	Female	25to34	Seperated	Very good	Yes	Degree OH	LowManProf	White Br/O	0	Yes	Yes	Yes	Never	No	High sugar intake	Regular check-up	1000-1300ppm	
22	Male	25to34	Married	Good	Yes	Degree OH	LargeEmpHighMan	White Br/O	1	No	Yes	No	Occasionally	Never smoked	Not high	Regular check-up	1350-1500ppm	
23	Female	45to54	Married	Very good	Yes	Other qual	LowManProf	White Br/O	0	No	Yes	No	Never	No	High sugar intake	Regular check-up	1350-1500ppm	
24	Male	65to74	Married	Good	Yes	Other qual	LowerSupTechOC	White Br/O	0	No	Yes	No	Never	No	High sugar intake	Regular check-up	1350-1500ppm	
25	Female	45to54	Single	Good	No	Degree OH	LowManProf	White Br/O	1	No	Yes	No	Hardly ever	No	Not high	Regular check-up	No fluoride	
26	Male	35to44	Married	Good	Yes	Degree OH	LowManProf	White Br/O	2	No	Yes	No	Never	Never smoked	High sugar intake	Regular check-up	1350-1500ppm	
27	Male	45to54	Single	Very good	Yes	Other qual	SmallEmpOAW	White Br/O	0	No	No	No	Never	No	High sugar intake	Regular check-up	1350-1500ppm	
28	Female	55to64	Divorced	Good	Yes	Other qual	semi-routine	White Br/O	1	No	Yes	No	Hardly ever	Never smoked	High sugar intake	Regular check-up	1000-1300ppm	
29	Female	55to64	Widowed	Good	Yes	Other qual	LowManProf	White Br/O	1	No	No	Yes	Never	Never smoked	High sugar intake	Only when having trouble	1350-1500ppm	

Showing 1 to 31 of 4,441 entries

Figure 14 – Model matrix of survey constraint data in dummy coded format

	a16to24_s	a25to34_s	a35to44_s	a45to54_s	a55to64_s	a65to74_s	a75to84_s	a85plus_s	male	female	Highqual_s	Otherqual_s	N_1.1	N_1.2	N_2	N_3	N_4	N_5	N_6	N_7	N_8	Hvgood
1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	
2	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	
3	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	
4	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	
5	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	
6	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	
7	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	
8	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	
9	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	
10	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	
11	0	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
12	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	
13	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
14	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
15	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	
16	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
17	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	
18	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	
19	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	
20	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	
21	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
22	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	
23	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	
24	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	
25	0	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
26	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	
27	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	
28	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	
29	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	

Showing 1 to 31 of 4,441 entries

The reweighting process used in the 'ipfp' package is similar in nature to the method undertaken by Anderson (2007). All individuals in the survey are given an initial weight, sometimes this is automatically set to 1, but it can also be calculated by dividing the number of individuals in an area (from the Census data) by the number of households from the survey data. Anderson (2007) suggests using a regional weighting technique which involves excluding (or giving a weight of 0 to) individuals from outside the region being simulated, the idea being that this 'avoids filling, for example, Sheffield with Londoners' (p.12). In practice this sounds like a reasonable approach, however the ADHS data is only available at the Strategic Health Authority (SHA) level, meaning that the lowest spatial scale relevant to Sheffield was the Yorkshire and Humber region. The Yorkshire and Humber region is far from homogenous, and Sheffield is just as likely to have characteristics in common with towns and regions from other parts of the UK as with the rest of the Yorkshire and the Humber region.

This assumption is in line with work conducted by the ONS on ‘statistical neighbours’ (Office for National Statistics, 2011c), which according to the classification would have Leeds as the most similar local authority to Sheffield, followed by Newcastle-upon-Tyne, Cardiff, Preston and Derby. Based on this classification it can be seen that Sheffield is more similar to a number of towns and cities from outside the Yorkshire and Humber region than some of those in it. Therefore, it did not necessarily make sense to exclude other regions from the analysis, which would also reduce the sample size significantly. The accuracy of spatial microsimulation models can suffer through the reduction of sample sizes, as with a smaller pool of individuals and a potential reduction in the variety of characteristics amongst the sample, it may be harder for the method to create the target variables as accurately as would be desired. Target variables are those that are simulated from the survey data, that do not currently exist in any data sources produced for small area geographies (i.e. tooth decay). This theory is supported by Ryan et al. (2009), who found that ‘as input sample size increases, resulting populations experience gains in accuracy’ (p.201).

The reweighting method used in the ‘ipfp’ package is the same as that displayed in the worked example of the IPF procedure in Section 4.3, where the formula below is applied to the six constraint variables iteratively until convergence of the datasets is achieved.

$$ni = wi \times sij/mij \qquad \text{(Equation 1)}$$

Anderson (2007) states that 20 iterations were enough to achieve convergence, whereas other authors have suggested only 10 are required (Ballas et al, 2005a). As a compromise 15 iterations were used in this research. For this research more than one target variable needed to be simulated, due to the need for individual level data to fill out the non-neighbourhood based variables in the pathways. This was a fairly simple process, and involved adding additional target variables to the dataset of individuals from the ADHS before data manipulation took place, checking to make sure that the number of people matched up between the two input data sources taken from the survey data (the participant data, and the model matrix data). The original sample size for a model simulating only tooth decay was 5388 individuals, however once the additional target variables had been added this was reduced (through missing data) to a final sample size of 4840 individuals.

4.6. Integerisation of the microdata

One of the disadvantages of the IPF method is that output datasets contain decimal numbers, which is not a practical way of studying individuals in their local environments (it is not possible to have 0.34 of a person for example). Therefore, integerisation methods were employed to turn these decimal numbers into whole numbers, which is a format much more suited to further research using agent-based modelling. Lovelace and Ballas (2013) have highlighted that integerisation can reduce model fit, however for the needs of this research it was a necessary step. For this research the ‘truncate, replicate, sample’ (TRS) integerisation method was used, developed by Lovelace and Ballas (2013). This method involves converting the weights created in the spatial microsimulation modelling into integers, ‘that represent whether the associated individuals are present (and how many times they are replicated) or absent’ (p.2). A key feature of the TRS method is that it should not lead to any increases in differences between the simulated data and real data, as has been found in previous attempts (Ballas et al, 2005c).

The integerisation process was key, both in terms of creating the synthetic microdata, and in terms of being able to externally validate the model. Integerisation was therefore undertaken after the internal validation, in order to change the data into a format on which external validation could be conducted (i.e. with integers rather than decimals). Before integerisation it was possible to conduct internal validation as this tests the fit of the Census data against the aggregated survey data. However, to compare the model’s findings with alternative datasets from the real world, the final dataset needs to be produced. Such data is not available, or in the correct format, during the internal validation stage as this is only measuring the data used as constraints. Thus the key variables in this study such as tooth decay scores, oral health habits, and attitudes were also not obtainable until after integerisation.

The TRS method has been tested alongside other integerisation techniques, and proved to be the most accurate, introducing less error than its counterparts (Lovelace and Ballas, 2013). It also considers the weights created by the spatial microsimulation to have two types of data, both the probability of an individual being selected, and also the number of times that individual should be replicated. The method works by first removing all information past the decimal point for all the weights (the 'truncation'), leaving integer values (with the decimal remainders saved into a separate data frame). The next stage involves the 'replication' process, where rows of data (or individuals) are replicated based on their integer weight. A weight of 7 for example would be replicated seven times, while a weight of 1.73 would be replicated once (due to the truncation). The final stage involves weighted random sampling of the decimal weights left over from the truncation process, which determine the probability of an individual being selected. Thus individuals are replicated based on their truncated weight, while those with low weights (less than one) still make up a large proportion of those selected, but are replicated fewer times due to their weights being lower. The whole process is conducted deterministically, leading to identical results each time the model is run.

4.7. Validation

One of the most important parts of the IPF process is ensuring the validity of the output, particularly given that the models are estimating variables which previously did not exist in a geographical format. Models can be validated both internally and externally. Internal validation involves checking the fit between the Census constraints and the aggregated survey data, and Lovelace and Dumont (2016) list several ways to test this. Firstly, a simple check of the data can be performed using the correlation function in R ('cor'), which calculates a Pearson's *r* coefficient (Pearson, 1896). This value should be as close to 1 as possible, and preferably over 0.99 to show the two datasets have converged. This can also be assessed visually by plotting the two datasets against each other. Two further tests of the fit of the data include the total absolute error (TAE) and the standardised absolute error (SAE). The TAE calculates the differences between the observed and simulated values for each constraint in each zone, and sums these. The SAE is perhaps a more useful measure of error, as this takes the TAE for each zone and divides it by the population of each area, thus making comparisons between different areas more relevant. Having calculated the TAE and SAE it is then possible to investigate which areas have the highest error levels, how much these areas contribute to the overall level of error, as well as which of the variables have not converged and are causing the error to be higher in these areas. Statistical tests such as regression and t-tests can also be used to test the convergence of the two datasets (Edwards and Clarke, 2007)

External validation of spatial microsimulation models, while the ‘gold standard’, can be more difficult to conduct. Some models can be caught in a ‘catch 22’ situation, as the outputs are often new data with nothing else to compare them against, hence the need to use the method in the first place (Lovelace and Dumont, 2016). Where external validation is possible, there are a number of ways of achieving this, with Tanton and Edwards (2013) proposing four potential approaches. The first of these involves aggregating the results of the model to a spatial scale where relevant data already exist that can be compared to the output. The authors warn of the importance of the ecological fallacy during this approach, and the importance of not making assumptions about populations from aggregate data. The ecological fallacy concerns the ‘aggregational variability inherent in areal data’ (Openshaw, 1984 – p.18), whereby aggregate levels group statistics from Census data can become unrepresentative of the individuals that comprise that group. The danger then comes from making assumptions about the nature of individuals based on these aggregated statistics. This also brings into the question the areal unit used for the analysis, which in turn will influence the resultant statistics. This is another reason why the neighbourhood based variables in this study were, where possible, not based on aggregated individual data, so as not to create difficulties in interpreting group level data.

The second approach, which is far more time and resource consuming, is to collect primary data on the outcome variable of choice in either one, or a sample of small areas, assuming no other relevant data exists. The third approach suggested by the authors is to compare the simulated data to different variables that happen to be correlated with the simulated output, using data that already exists at the micro level. Using small area geographical data from the Census is perhaps the most obvious option in this regard, while Anderson (2013) used the Welsh Indices of Multiple Deprivation (StatsWales, 2011) as an alternative source of data. The fourth and final suggestion is to run the model at a larger geographical scale, before measuring the results against (reliable) estimates for larger scales from another dataset. This approach has similarities with the first method, and in the case of this study data from the Yorkshire and Humber region could be used. A fifth technique for external validation is to use unconstrained variables, i.e. one that is present in both datasets but was not used as a constraint (Campbell, 2011). By including these variables as additional target variables, it is possible to compare the model estimates against the Census version of the variable. This gives an idea as to whether the model is accurately producing data for the target variables. This approach was used in this research, through the inclusion of ethnicity and marital status as additional target variables.

Internal validation should be conducted as the bare minimum in the validation process. It may not be possible to conduct external validation on a model however, and this will very much depend on the target variable of choice in the research, as well as the spatial scale at which the data is produced. Once the data has been validated it is possible to export data frames from R as .csv files using the ‘xlsx’ package and ‘write.csv’ command. Once in this format it is also possible to read such data into agent-based modelling software, and this will be explained in more detail in Chapter 5. The next section will focus on the results of the models and their validation, before moving on to geographical and statistical analysis of the output.

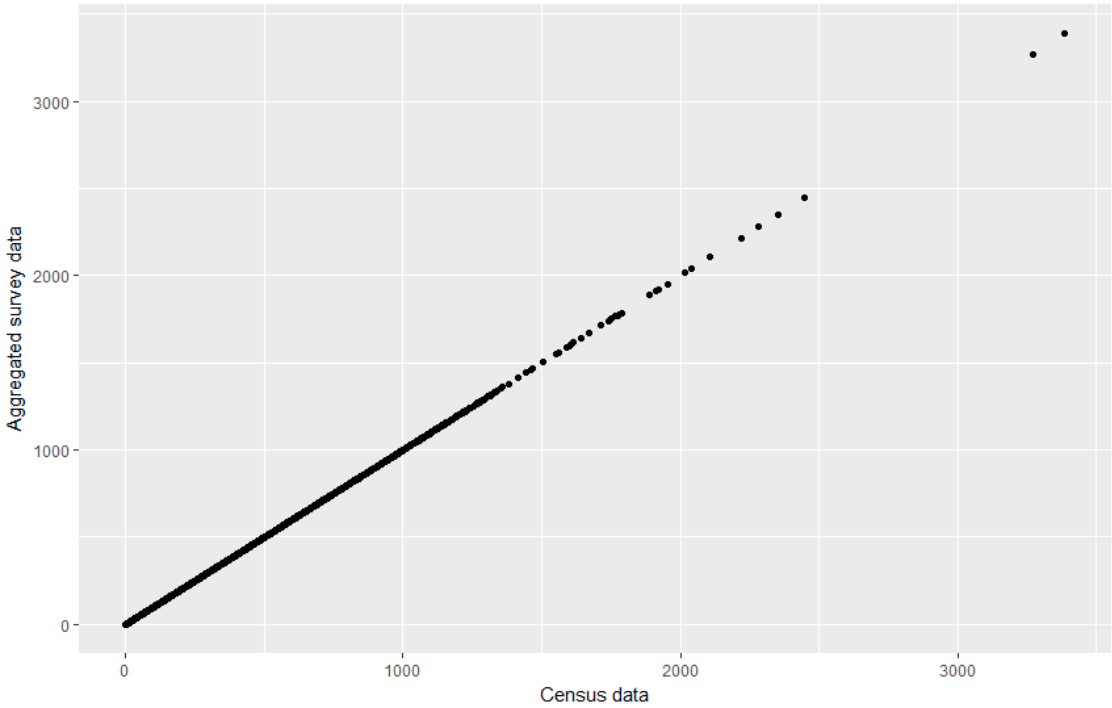
4.8. Results of the spatial microsimulation model

This section is split into two parts. The first (Section 4.8.1) will deal with the internal validation of the spatial microsimulation model, measuring how well the aggregated survey data constrained to the Census population totals, and a discussion surrounding this. The second part (Section 4.8.2) will look at the external validation of the model through the use of the ‘ethnicity’ and ‘marital status’ variables, to judge how well the model was able to predict these unconstrained Census variables, accompanied by a second discussion on the success of this approach.

4.8.1. Internal validation of the model

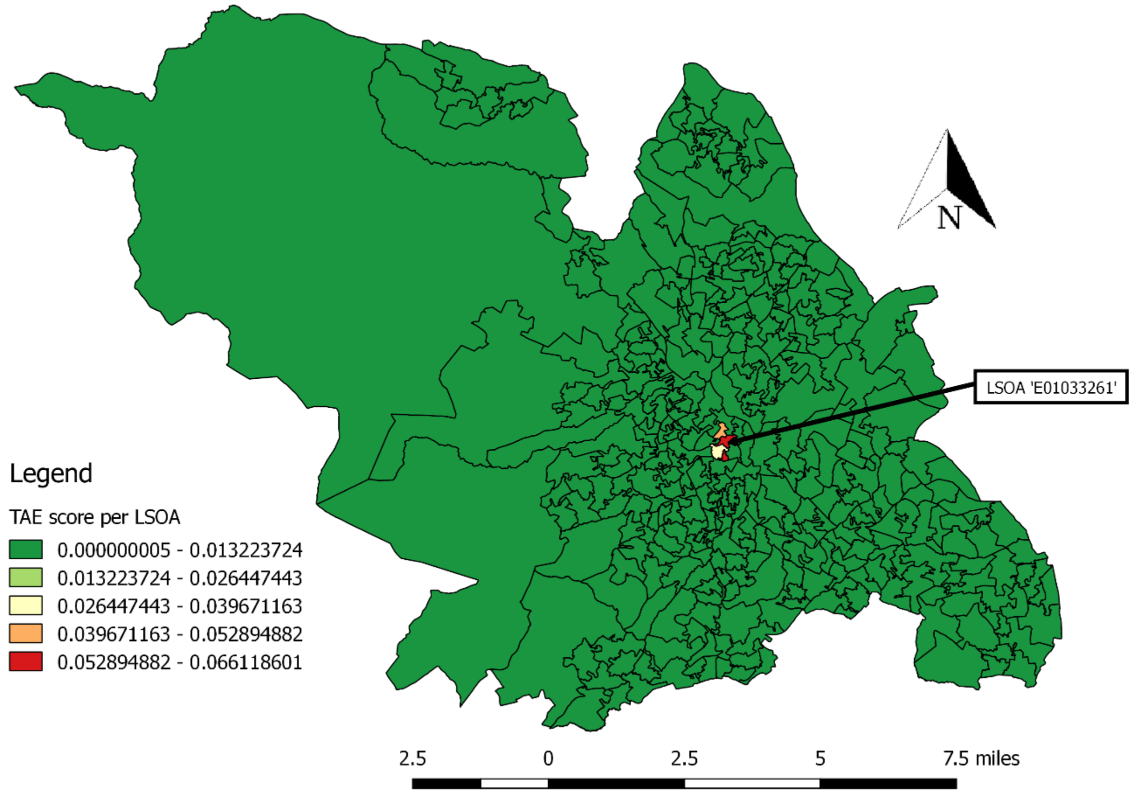
The overall model constrained extremely accurately. Using Pearson’s *r* correlation coefficient, the Census data constraints and aggregated survey data were plotted and compared, obtaining a perfect score of 1. The correlation between the aggregated survey data and the Census data can be seen in Figure 15. The correlations for the individual constraint variables were also very accurate, with each constraint also attaining a correlation coefficient of 1.

Figure 15 – Census data plotted against the aggregated survey data



Beyond the use of the Pearson’s *r* statistic already presented, the model was internally validated in a number of other ways. As mentioned in Section 4.7, the total absolute error (TAE) is a measure of the absolute difference between the Census data and simulated data for each zone. Having run the model for 15 iterations the TAE score for the overall model was again extremely encouraging at just 0.2985, out of a total adult population of 452,014. This ranged from 0.000000005 to a maximum of 0.066118601 across the 345 LSOAs. The population total created from the simulation was also very accurate, as the simulated population figure of 452,014 compared exactly to the adult population supplied by NOMIS (NOMIS, 2011) at the time of the 2011 Census (n=452,014). The standardised absolute error (SAE), where the TAE is divided by the population of the LSOA, also performed extremely well, with a score of just 0.0000001101. This ranged from 0.0000000000007 to a high of 0.00000776133.

Figure 16 – Total absolute error (TAE) scores per LSOA



Figures 16 and 18 show the TAE and SAE scores mapped onto the Sheffield LSOA boundaries. Figure 16 shows that a cluster of zones to the west of the city centre featured the highest TAE scores. The zone with the highest error score was LSOA E01033261, an area near Broad Lane and the University of Sheffield's city campus (Figure 17). Further analysis revealed this zone accounted for 27.47 % of the TAE in the model.

Figure 17 – Location of LSOA E01033261 near Broad Lane and the university

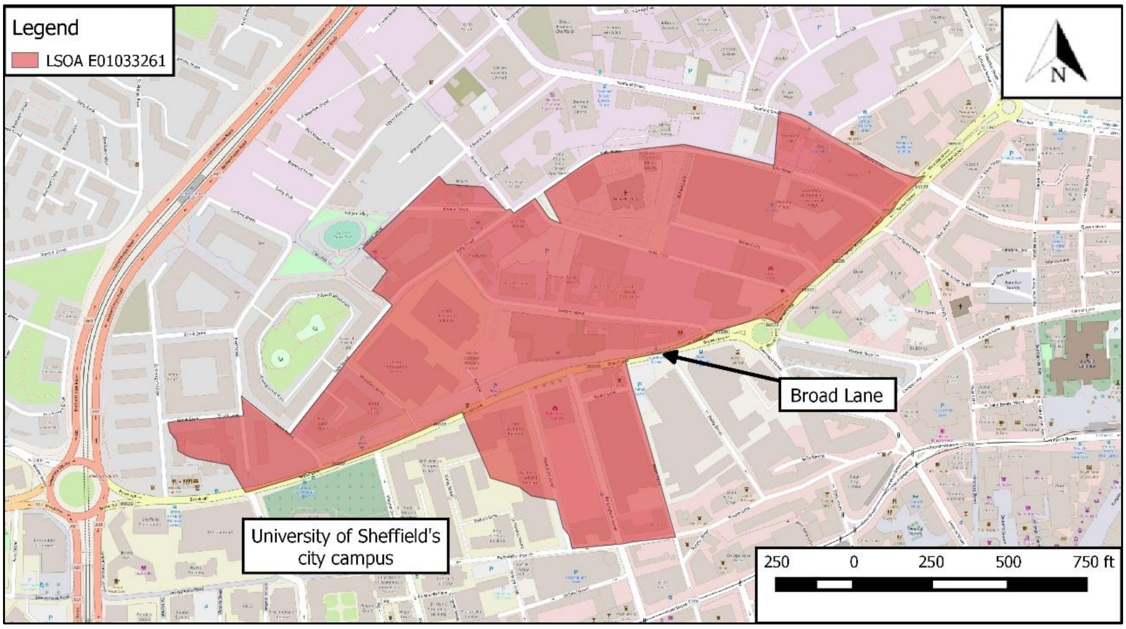
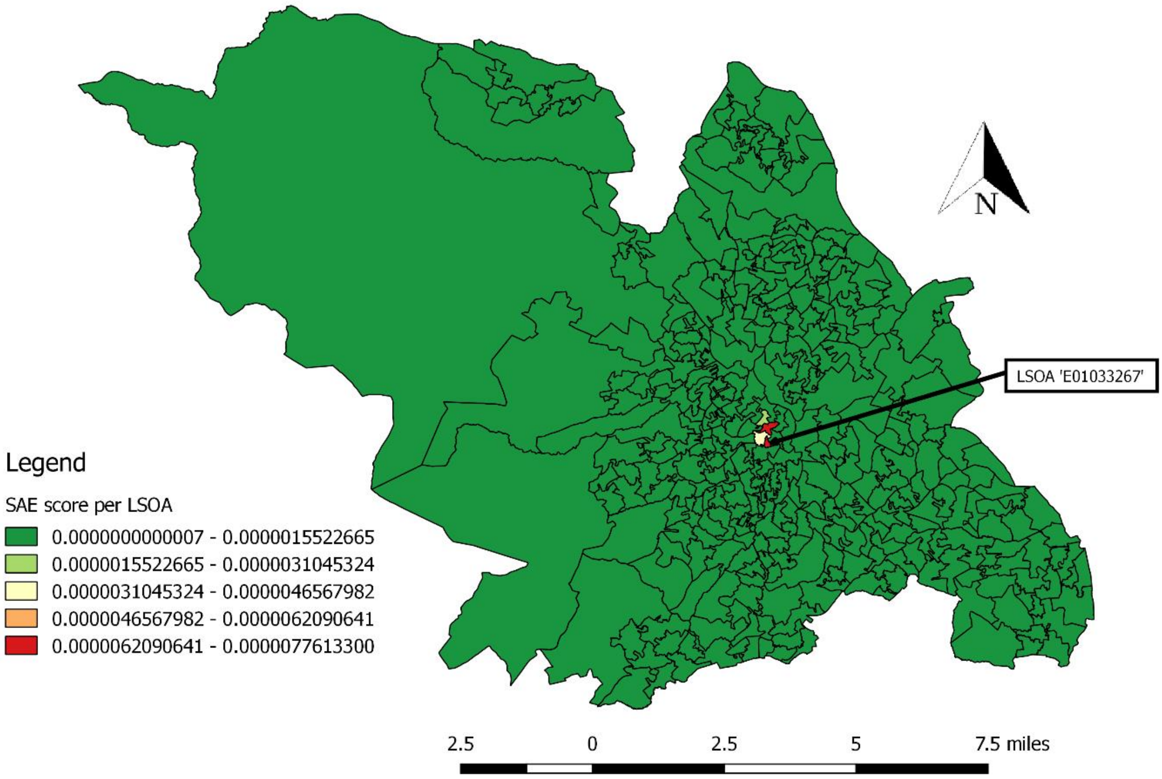


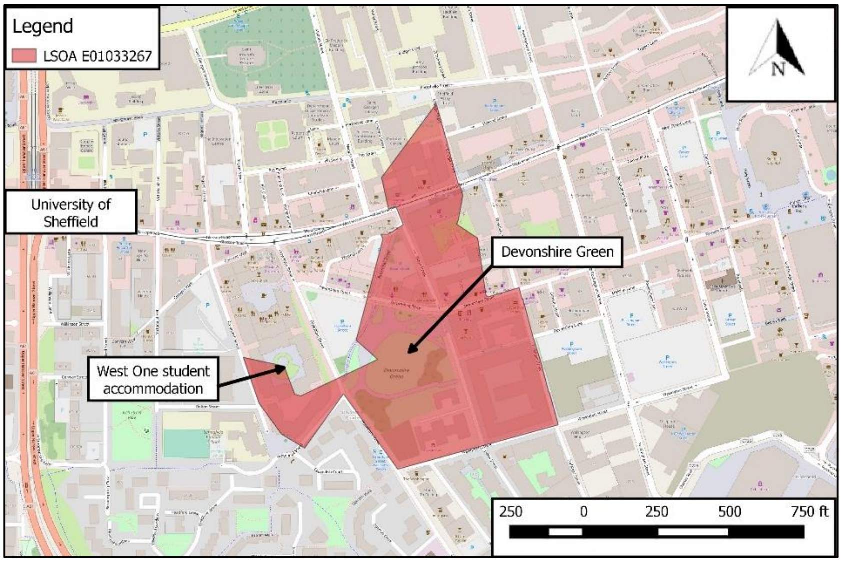
Figure 18 demonstrates a similar pattern for the standardised absolute error, with a cluster of LSOAs to the west of the city centre again providing the highest error scores.

Figure 18 – Standardised absolute error (SAE) scores per LSOA



This time LSOA E01033267 had the higher error score. This area was situated around Devonshire Green, and was also located near to a number of buildings affiliated with the University of Sheffield, as well as student accommodation at West One (Figure 19). This LSOA accounted for 33.25% of the standardised error in the model.

Figure 19 – Location of LSOA E01033267 near Devonshire Green and West One



It would seem that one zone did not account for the majority of the absolute or standardised error within the model. If one zone had been principally responsible for the error, then the rest of the model could be assumed to be accurate. However, this is not necessarily a problem when the error scores are a small as they were in this research. Interestingly, a similar cluster of LSOAs to the west of the city centre appeared to account for the majority of the error (both absolute and standardised) in the model, with these areas being located near the main campus of the University of Sheffield. As these LSOAs cover areas occupied by university buildings as well as student accommodation, it may be that the characteristics of the student population in the area, which will differ to that of the usual resident population near the university, caused the higher error counts in these areas. Ballas et al. (2005a) have stated that spatial microsimulation is not suitable for estimating populations when 'affected considerably by external and localised factors, such as transport networks and public transport services, or the presence of a disproportionally large university or a single major employer in the region' (p.14). It is possible that the presence of the university and its students may have affected the output of the model, although it is curious that higher error counts were not seen in areas occupied by buildings associated with Sheffield Hallam University and its students.

While the TAE and SAE are helpful in assessing error, they do not give any indications of statistical difference, and tests of this type provide an alternative way to judge the fit of the data internally. Edwards and Clarke (2007) used two-tailed equal variance t-tests to check whether there were statistically significant differences between the Census data and aggregated survey data, with the results as applied to this research presented in Table 15. The data in this table demonstrates the excellent fit of the model internally, with no statistically significant differences between the Census and survey data for the overall model, or any of the six constraint variables.

Table 15 – Results of the two-tailed equal variance t-tests

Constraint	p-value
Overall model	0.995
Age	0.9996
Gender	0.9993
Qualification	0.9998
NS-SEC	0.9995
General health	0.9998
Car ownership	1

Edward and Clarke (2007) also suggest linear regression as a further internal validation method, as the p-values and R^2 statistics give a good indication of internal fit. Again the simulation performed well in this regard, as the overall model and all of the constraints attained statistically significant p-values ($p < 0.01$), and R^2 values of 1. Finally, Smith et al. (2007) contend that for a rare disease such as diabetes, there should be less than 10% error in 90% of the areas (output areas in their research). In this research all of the SAE scores for each LSOA were below 10%, with the highest standardised error score being 0.00000776133, indicating excellent internal model fit. Based on the findings of the internal validation, it was assumed that the model was internally valid, and had constrained accurately to the Census data.

4.8.2. External validation of the model

While the results of the internal validation were a positive, the point of spatial microsimulation modelling is to constrain and converge the two datasets, so the fact this has occurred should not be too much of a surprise. Having internally validated the model, the data was integerised in order to create ‘whole’ individuals within the data. This in turn allowed for the external validation to take place.

Following the integerisation of the data, a key error in the modelling was discovered during the external validation process. As mentioned in Section 4.4.2, data transformation was undertaken on the NS-SEC data to account for the ‘not-classified’ data. Also mentioned in this section was that this process left some LSOAs (n=175) one or two (both positive and negative) out from the actual Census numbers. An additional section of code was created that split these 175 zones into four data frames: zones out by negative one; zone out by positive one; zones out by negative two; zones out by positive two. Once these zones had been corrected they were combined into a further data frame, then joined with the LSOAs without error to form one data frame with all 345 LSOAs, complete with correct NS-SEC data. However, due to the way the zones were re-combined, the data no longer matched the order of the LSOAs in the Census data. This led to spurious results during the external validation. While it would be possible to re-order the Census variables, a more practical approach was to fix the code to leave the LSOAs in the desired order. This was done by adding a key to the original Census data, which acted as a reference point for re-ordering the LSOAs later in the code. This allowed for more reliable external validation to take place, but also reinforces the importance of model checking throughout. The internal validation would not have picked this up as this stage is concerned with checking whether the survey data had aggregated correctly to the Census constraints. Thus the order of the LSOAs was not as important as the data associated with each zone at this step.

As alluded to in Section 4.7, external validation of spatial microsimulation models is harder than internal validation. This is because there is rarely appropriate data available to test the model (hence the need for the modelling in the first place), and often the model outputs will need to be adjusted to match data at other spatial scales. With this in mind, the UK Data Archive, the Office for National Statistics (ONS) and the Health and Social Care Information Centre (HSCIC) were contacted regarding possible data sources for external validation. The UK Data Archive was contacted specifically regarding differing spatial scales for the ADHS (which is only available publically at the Strategic Health Authority level), however such data were not available. Following a recommendation from data analysts at the UK Data Archive to check with the ONS, it was found that data were available under special agreements at the old Primary Care Trust level. However, this equated to only 77 cases for Sheffield (out of 1,021 for Yorkshire, and 11,380 Overall), of which only 42 had undertaken a dental examination and included the 'numdu98' variable. Given that the sample size for the original spatial microsimulation model contained 4840 individuals, 42 was not seen as an adequate number for validating the model. Finally, the HSCIC was contacted with regard to whether they had any alternative datasets on tooth decay (not related to the ADHS) at a spatial scale lower than SHA (or similarly sized regions). No such data were available however.

Thus it became impossible to test how well the spatial microsimulation model had performed regarding the tooth decay variable. While this was unfortunate, it is a common problem within the field. The performance of the model could still be assessed by evaluating it against 'non-constrained' variables though (Campbell, 2011). This involved taking data from variables that were present in both the Census and ADHS, and that were not included as constraint variables in the model. The idea behind this was that if the model was able to predict these target variables accurately, it indicates reliability on the models part when creating new data, and gives confidence that the other target variables have been predicted accurately as well. Variables for household tenure and limiting long-term illness were excluded based on their incompatibility with the Census data, while data on household size were also excluded as it counted the whole population, rather than those aged 16 and over.

This left ethnicity and marital status as potential variables to use. The data for marital status was collected in a very similar manner in the ADHS and the Census data, making it comparable without any data manipulation. The data for ethnicity was found to be less comparable however. When using data for the usual resident populations of each LSOA the ethnicity data from the Census can be manipulated to match the categories from the ADHS exactly. However, this usual resident population included residents aged under the age of 16, therefore making it incomparable with the microdata, which contained only those who were aged 16 or over. Ethnicity data on individuals aged 16 and over was available from NOMIS (NOMIS, 2011), however this data was not available in the same format, meaning that when cross-tabulated with age (so as to exclude those aged under 16) the only comparable categories were 'white', 'mixed race' and 'other ethnic group'. It is unusual and unfortunate that the data was not available by age in a format that allowed for a greater disaggregation of different ethnicities, however as the other more comparable datasets were not measuring the same population that was created in the spatial microsimulation modelling, there was little choice but to use these data. The data would still allow for external validation to be conducted, however it is a shame that a more extensive validation using ethnicity was not possible.

Two statistical measures were used to judge the success of the external validation. The first of these was the R^2 value, as a general indicator of the fit of the data. This measures the fit of the data, but not necessarily around a given point. Therefore, the second measure used was the standard error around the identity (SEI), which is a measure of how well the data from the model falls around the 45-degree line (Tanton et al, 2011). Ideally data that fits perfectly should fall along this line, in the same way that it did during the internal validation (see Figure 15). The SEI score is calculated using the formula demonstrated in Equation 2.

Equation 2 – SEI calculation formula (Tanton et al, 2011)

$$\mathrm{SEI} = 1 - \frac{\sum (y_{\mathrm{est}} - y_{\mathrm{ABS}})^2}{\sum (y_{\mathrm{ABS}} - \bar{y}_{\mathrm{ABS}})^2},$$

In Equation 2 *Yest* are the spatial microsimulation estimates (of marital status or ethnicity in this case) and *YABS* are the equivalent data from the Census. Tanton et al, (2011) state that the SEI is interpreted in a similar way to the R^2 value, in that a higher figure indicates a better fit, only this time it refers to the 45-degree line. Presented below are graphs and statistics for each of the variables used for the external validation (Figures 20-28). The plain black lines represent the 45-degree line around which the SEI is measured, while the blue lines represent the fit of the data through the R^2 statistic. The blue shading around this line represents 95% confidence intervals.

Figure 20 – External validation using the marital status 'single' variable. $R^2 = 0.973$, *SEI = 0.9481061*

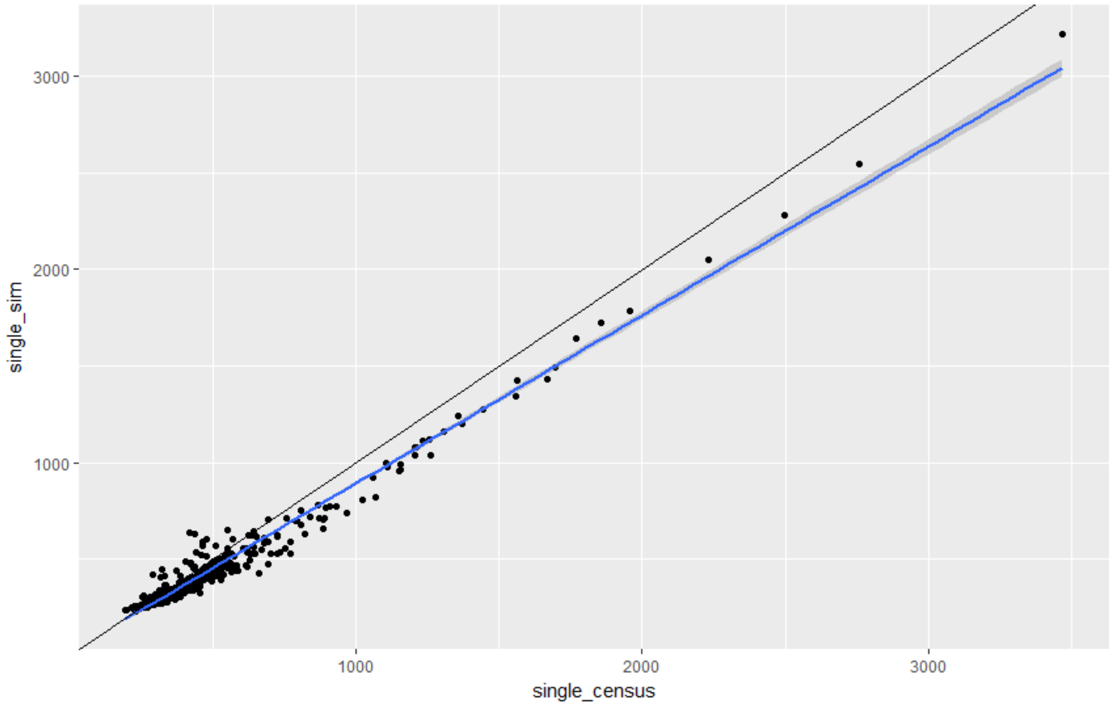


Figure 21 – External validation using the marital status 'married' variable. $R^2 = 0.813$, *SEI = 0.7381038*

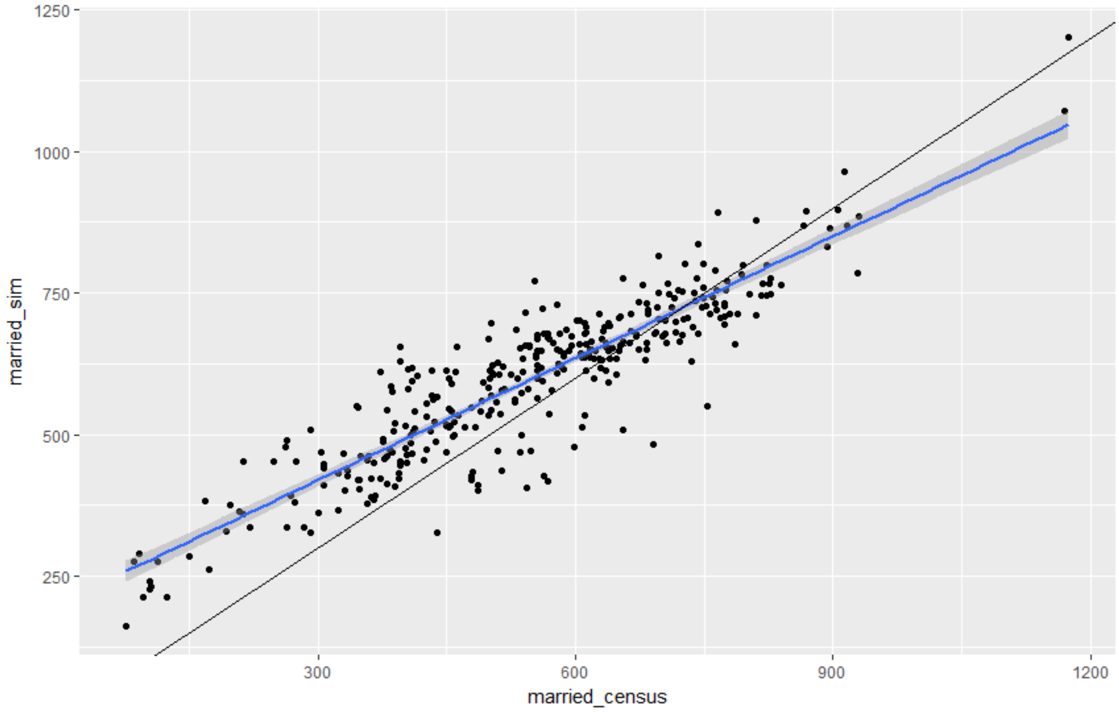


Figure 22 – External validation using the marital status 'civil partnership' variable. R^2 = -0.0000744, SEI = -0.303626

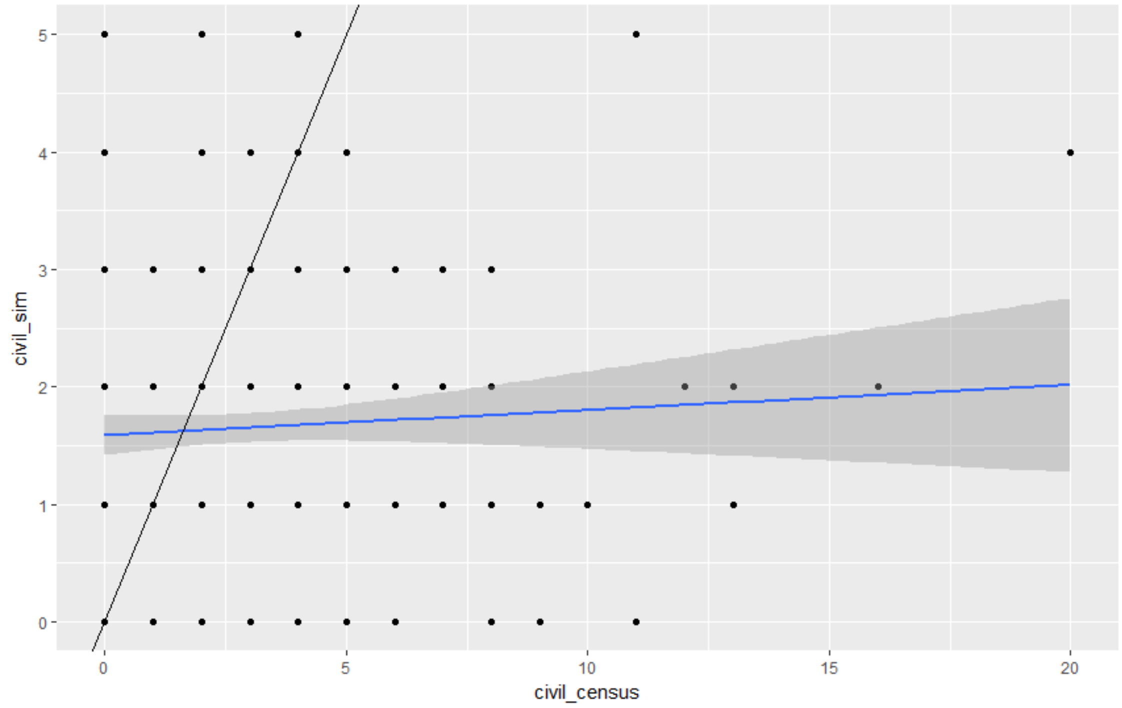


Figure 23 – External validation using the marital status 'separated' variable. R^2 = 0.293, SEI = -0.2427851

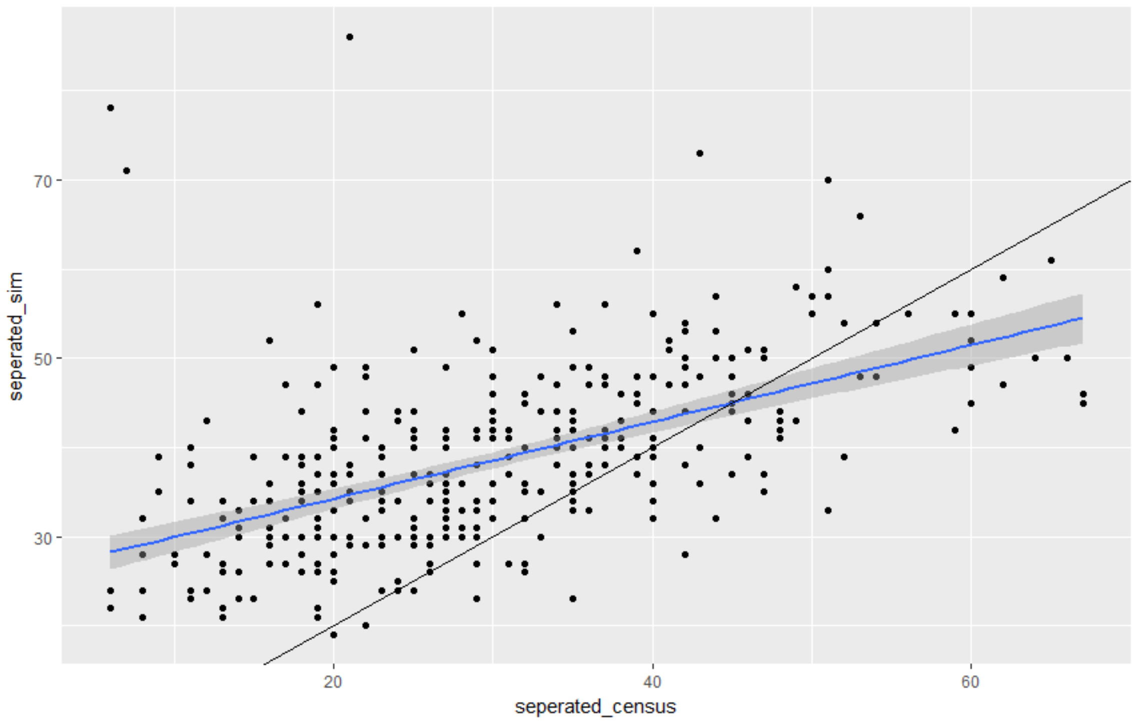


Figure 24 – External validation using the marital status 'divorced' variable. R^2 = 0.605, SEI = 0.5857174

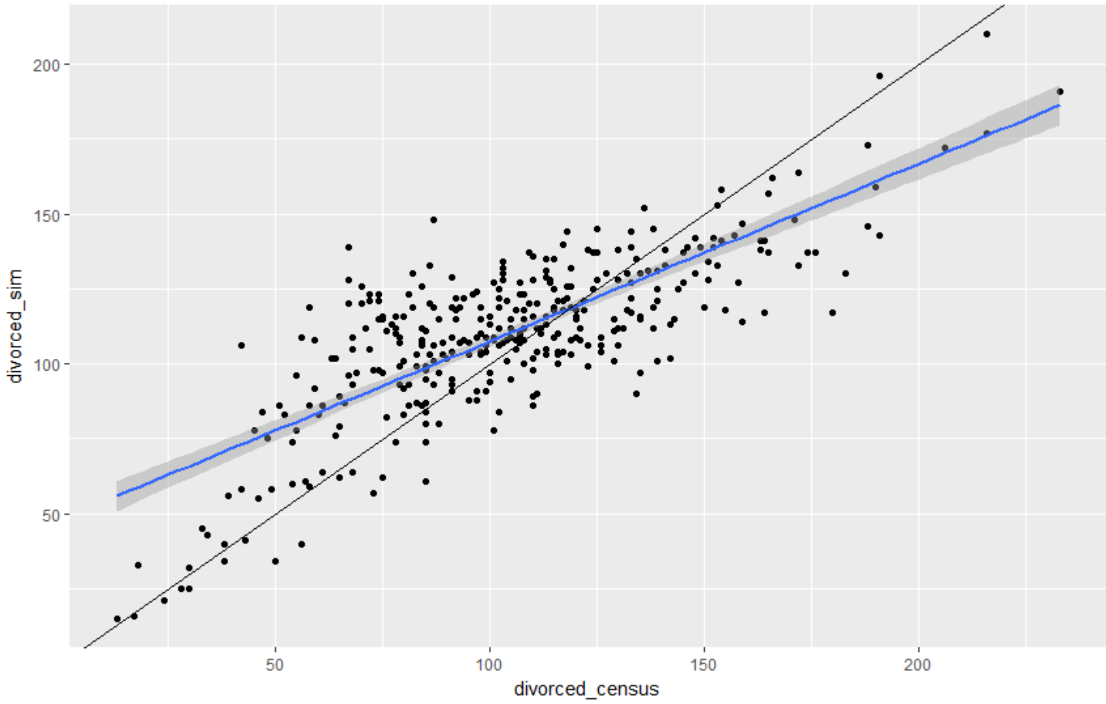


Figure 25 – External validation using the marital status 'widowed' variable. R^2 = 0.874, SEI = 0.7056197

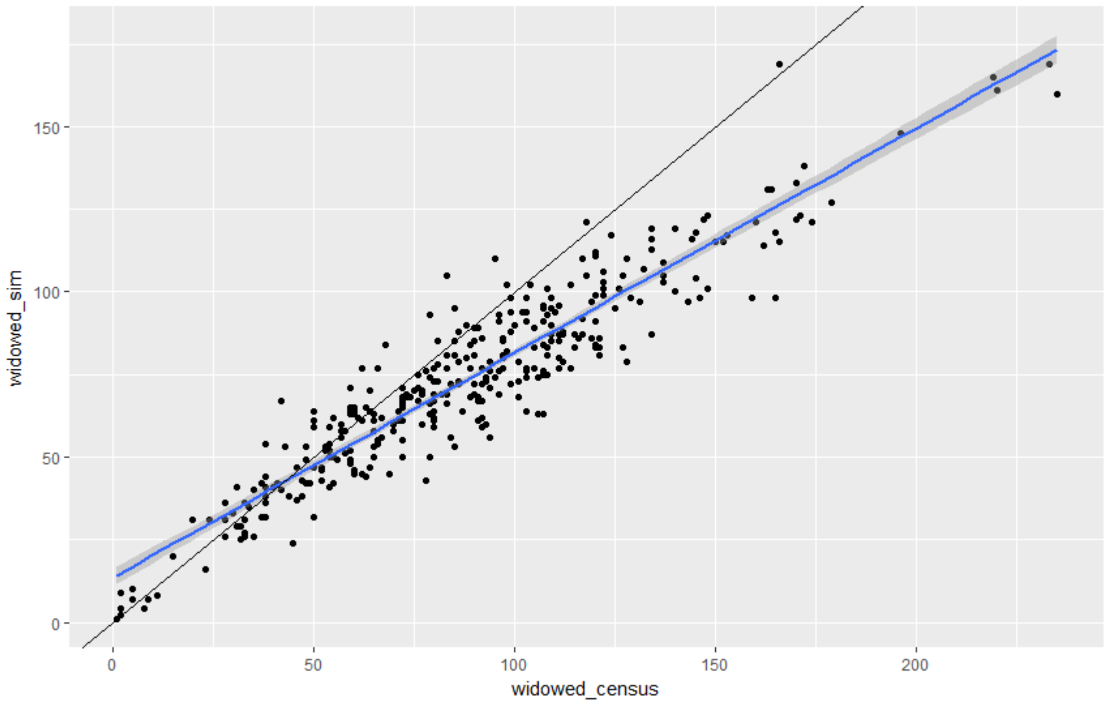


Figure 26 – External validation using the ethnicity 'White British/other white' variable. R^2 = 0.5314, SEI = 0.4424715

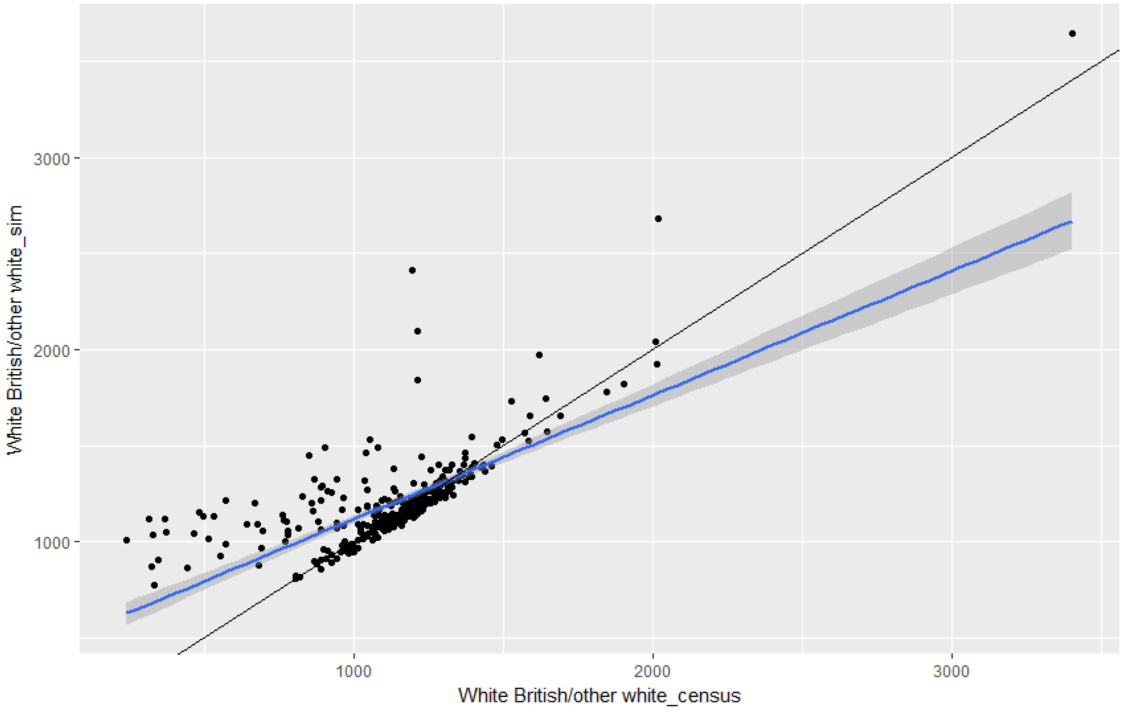


Figure 27 – External validation using the ethnicity 'Mixed/multiple ethnic group' variable. R^2 = 0.3522, SEI = -0.4144036

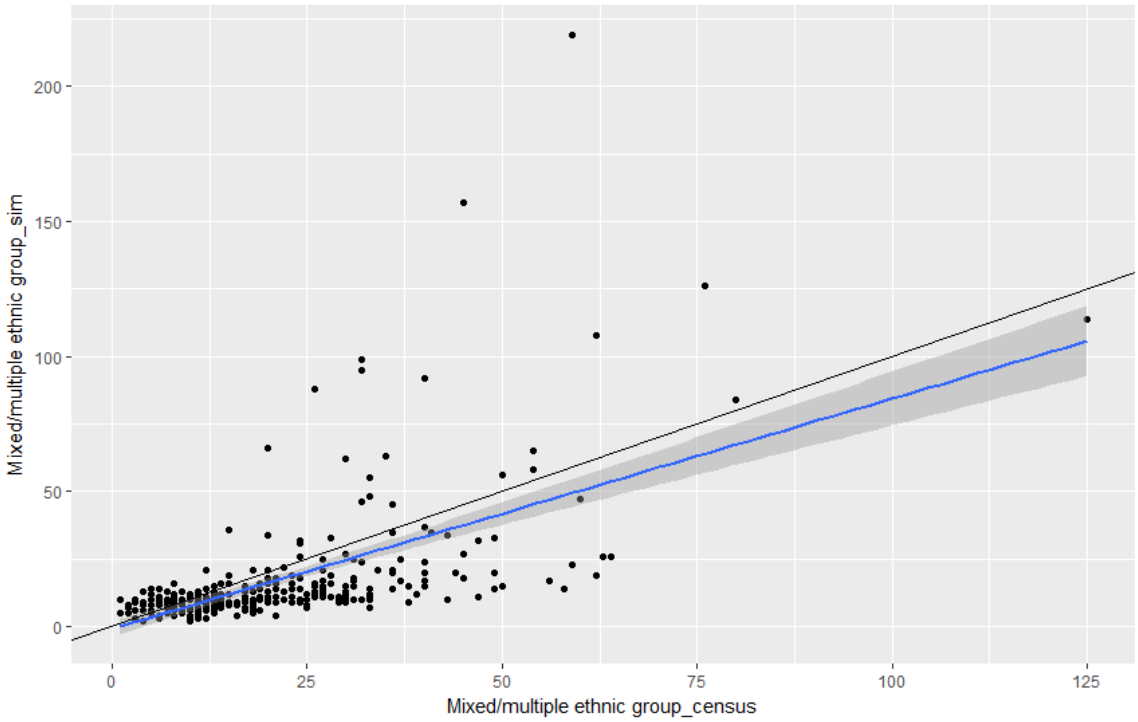
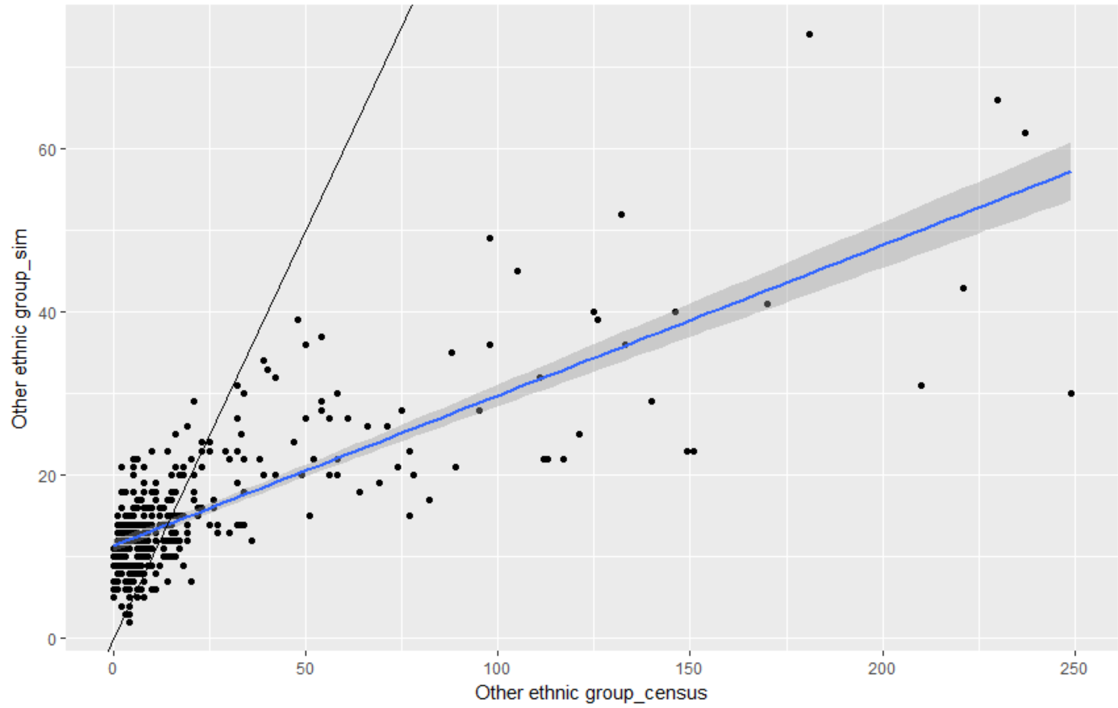


Figure 28 – External validation using the ethnicity 'Other ethnic group' variable. R^2 = 0.6123, SEI = 0.2682731



As can be seen from Figures 20 to 28, the results of the external validation were mixed, and varied far more than those of the internal validation. Many of the tests obtained R^2 values over 0.5, and SEI scores near to or above this figure. The scores for the 'civil partnership' variable (Figure 22) in particular were very low, with the 'separated' and 'mixed/multiple ethnic group' variables also having very low SEI scores (Figures 23 and 27). The latter two variables also showed the most discordance between the R^2 value and the SEI value, an important reminder that they are not measuring the data in the same way.

The general trend in the data seemed to be that that the larger the counts associated with a different level of a variable the better the validation scores, and that the variables that did not perform as well in the external validation had the smallest counts. This pattern was seen consistently throughout the external validation. Therefore, any errors in the simulated variables with lower counts would make more of a difference proportionally than it would with a variable with larger numbers. For example, an error of 6 is not that large out of a population of 1000, whereas it would be out of a population of 15-20. Burden and Smith (2015) have commented on a similar issue regarding validation, stating that 'the categories with small counts and those with high within-area homogeneity were the most highly variable' (p.575). This is not an attempt to 'blame the data' or to excuse the performance of the model however. It is an interesting discussion point though that the presence of external validation variables with smaller numbers may make a difference to the results, and perhaps are not best suited to such tasks. The analysis also showed that the model over and underestimated the populations for all of the target variables, indicating that the counts of the validation variables are not the only potential issue. It is of course highly subjective what a 'big' or 'small' number is, and may depend on the counts of the other variables in the model. Thus finding a suitable cut-off point for externally validating a model could be difficult if this approach was taken.

Part of the issue with the external validation may be due to the sample size. As already mentioned, Ryan et al. (2009) have commented that increased sample sizes lead to more accurate models. The sample size in this research was 5,388 when an initial model containing just tooth decay was completed, however this number fell to 4,840 upon the addition of the extra variables associated with the theoretical pathways. This was due to presence of missing data, as any individual with missing data points had to be removed from the analysis, as spatial microsimulation cannot work with missing data values. It may be that a larger sample would allow for a more accurate external validation.

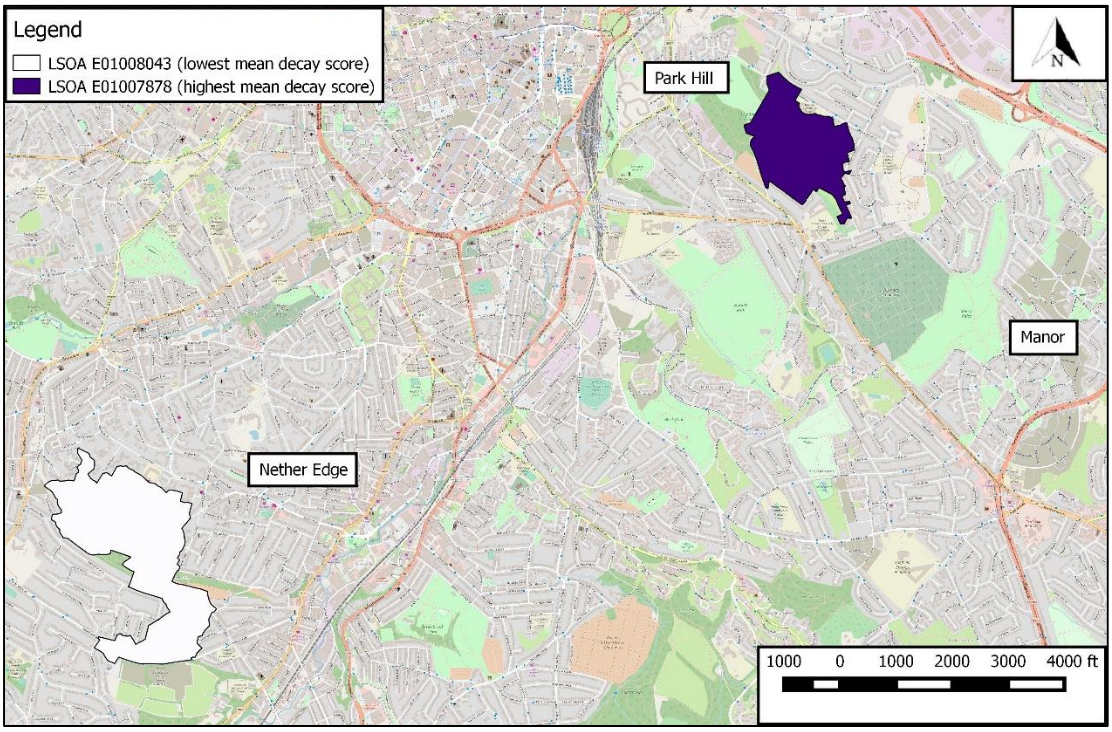
Given all of these concerns, most of the variables that had higher numbers (i.e. approaching 100 or above) still scored over 0.6 for both the R2 and SEI, indicating a good, if not great fit. Given these scores and those of the internal validation, the model was considered accurate enough to proceed with. Ideally more accurate values would have been obtained, and future work should investigate how the fit of such models could be improved, however such analysis was beyond the scope of this research.

4.10. Using microdata to analyse spatial trends in tooth decay

Having integerised the weightings from the spatial microsimulation model and validated the output, it was then possible to calculate other aggregate statistics for the city of Sheffield, and study a number of trends in the data. The results showed that the city had a mean tooth decay score of 1.047, with 40.1% of the city's residents experiencing some sort of decay. Within the 345 LSOAs of Sheffield, the mean tooth decay score ranged from 0.77 (LSOA E01008043 – an area near Nether Edge to the south-west of the city centre) to 1.69 (LSOA E01007878 – an area between Park Hill and Manor to the east of the city centre). The east-west divide between these two LSOAs is mirrored by the geographical divide between the highest and lowest scoring areas for mean tooth decay scores at the city level (Figure 29). The locations of the highest and lowest scoring LSOAs is presented in Figure 30. These were mapped using Quantum Geographic Information Systems (QGIS - version 2.8 – Wien, Development Team, 2015) and the British National Grid coordinate system. Sadly, the pattern shown in Figure 29 is not surprising when the spatial patterning of the city is considered in a wider social and historical context (Thomas et al, 2009).

Figure 29 – Map of simulated mean tooth decay scores per LSOA in Sheffield

Figure 30 – Location of the highest (purple) and lowest (white) scoring LSOAs

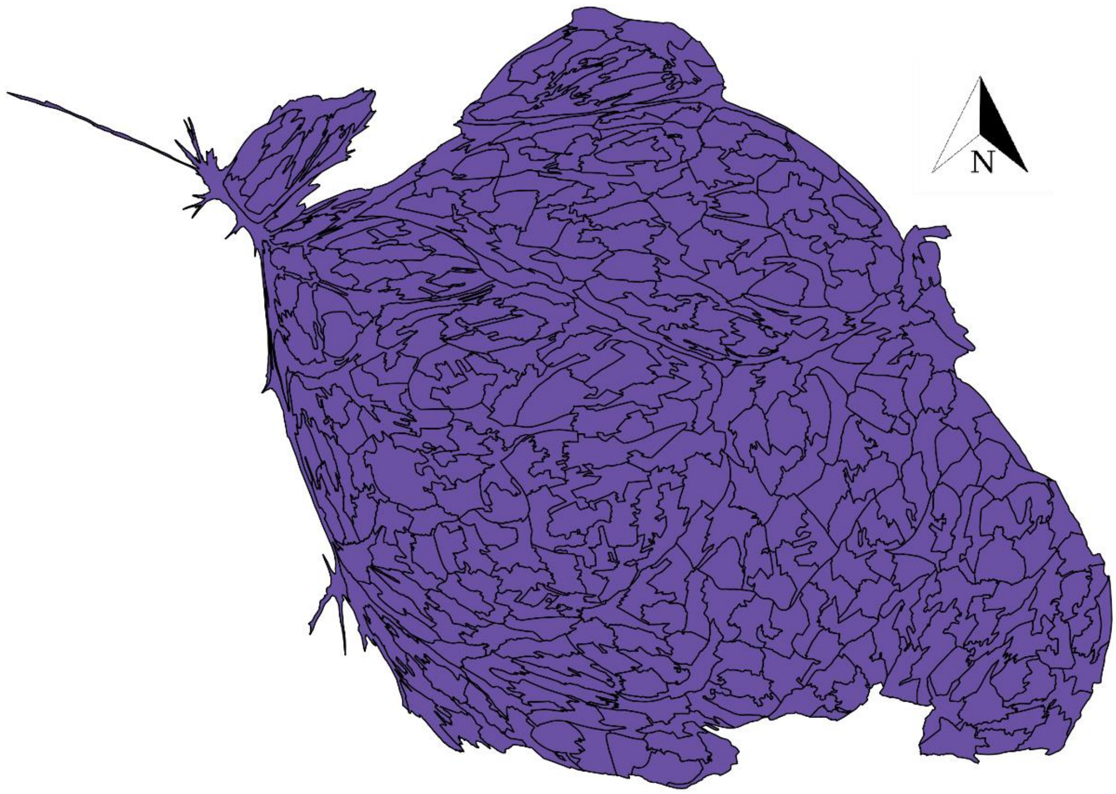


As can be seen in Figure 29, LSOAs in central, northern and eastern areas of Sheffield tended to have higher levels of decay than more westerly LSOAs. In a general sense this divide within the city is historical, and can be traced back to the positioning of heavy industry in Sheffield during its heyday. This was mainly located in the east of Sheffield, particularly along the Don Valley, while more affluent residents were located in the west of the city, away from the smoke and nearer to the Peak District. This divide is well recognised, and has manifest itself across a wide variety of other indicators including health, poverty and wealth, education, housing and crime (Thomas et al, 2009).

Positive associations were also found between deprivation scores per LSOA (IMD, 2015) and mean tooth decay scores from the spatial microdata through the use of linear regression, with an R^2 statistic of 0.8467. While use of the 2010 IMD would have been conceptually more favourable (being closer in date to the collection of the Census data, as well as the ADHS), this was not possible due to boundary changes at the LSOA level in Sheffield between the 2001 and 2011 Census'. The 2010 IMD was based on the 2001 LSOA boundaries, thus the later IMD data from 2015 was the most recent compatible source. While not as convenient chronologically, it still demonstrates an alarming pattern of inequality within the city.

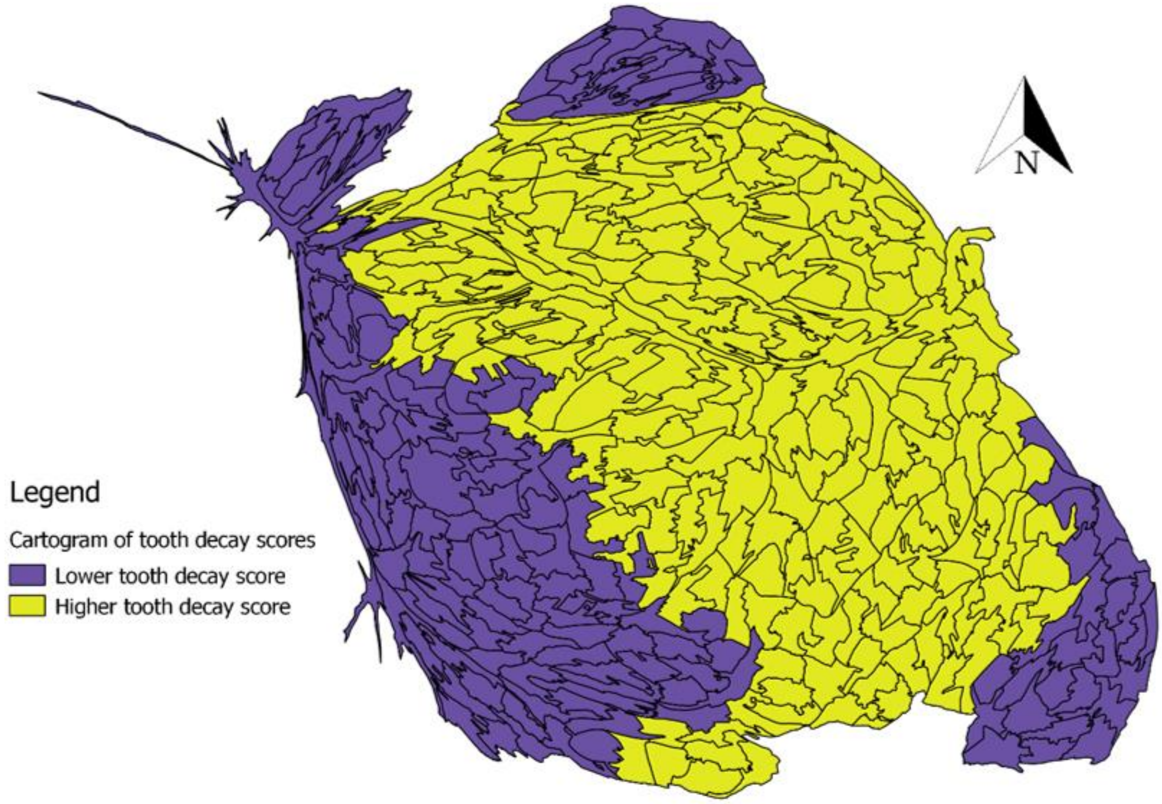
The unequal distribution of tooth decay in the city is further demonstrated by the human cartogram displayed in Figure 31. Human cartograms can only be used to display counts, rather than proportions, percentages or rates, so the total number of decayed teeth in each LSOA was used to create this figure. Human cartograms aim to represent geographical regions in proportion to their population, or some other property of these areas (Gastner and Newman, 2004). In this case the maps are representing the counts for tooth decay in each LSOA, and emphasising, proportionally, those LSOAs with the highest counts. This is interesting when comparing the cartogram with the standard geographical map of Sheffield, as they differ greatly. The west of the city has decreased in size proportionally, with the LSOAs in the east of the city being far more prominent.

Figure 31 – Human cartogram of tooth decay counts by LSOA in Sheffield



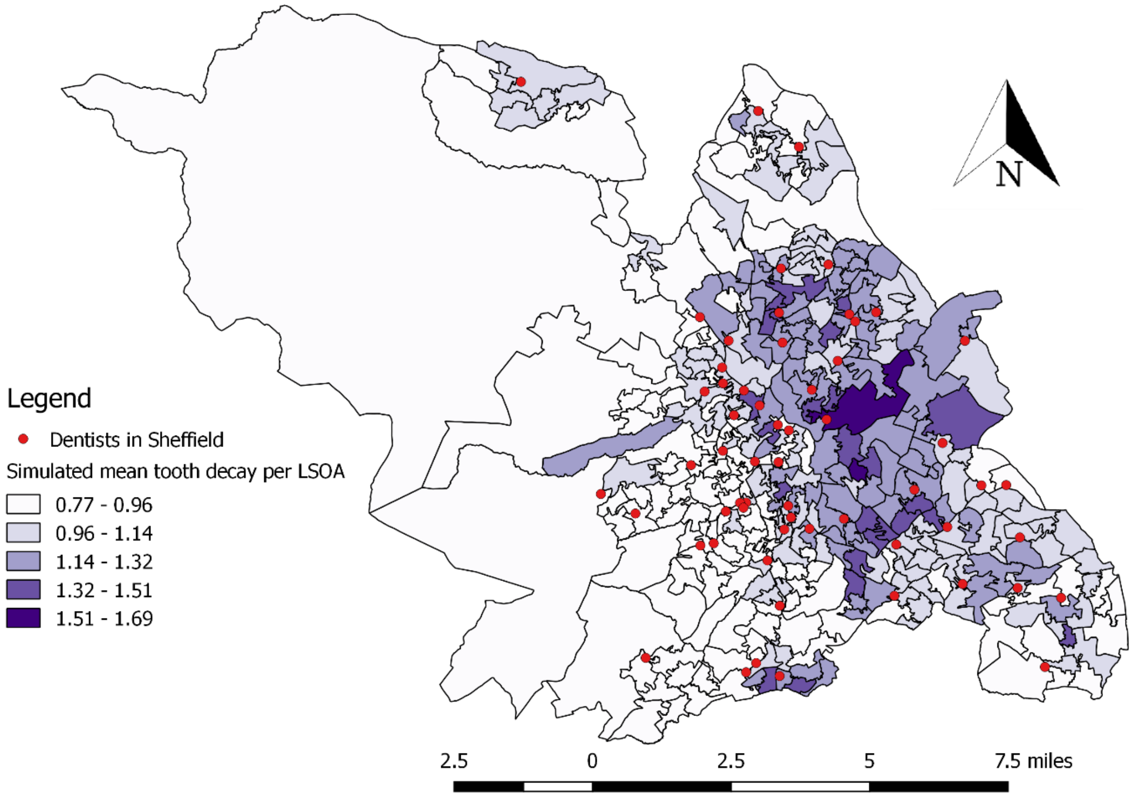
This can be seen in Figure 32, where the LSOAs in the east with higher tooth decay scores have been highlighted in yellow in an exploratory example. It should be stated that values of ‘higher’ and ‘lower’ are somewhat subjective and purely for demonstration purposes, rather than being based on exact calculations. From this it can be seen that these LSOAs now take up a far larger proportion of the total area in Sheffield compared to the standard geographical map presented in Figure 29. This distorted image demonstrates the nature of the divide in the city with regard to the number of decayed teeth, emphasising the spatial clustering of higher levels of tooth decay in areas of east Sheffield. This is another example of the possible use of data from a spatial microsimulation model, as well as an example of a novel spatial technique to highlight inequalities and the uneven distribution of variables. These cartograms were created using the ScapeToad open source software tool (ScapeToad, 2008).

Figure 32 – Human cartogram of tooth decay counts by LSOA in Sheffield, with higher scores highlighted in yellow



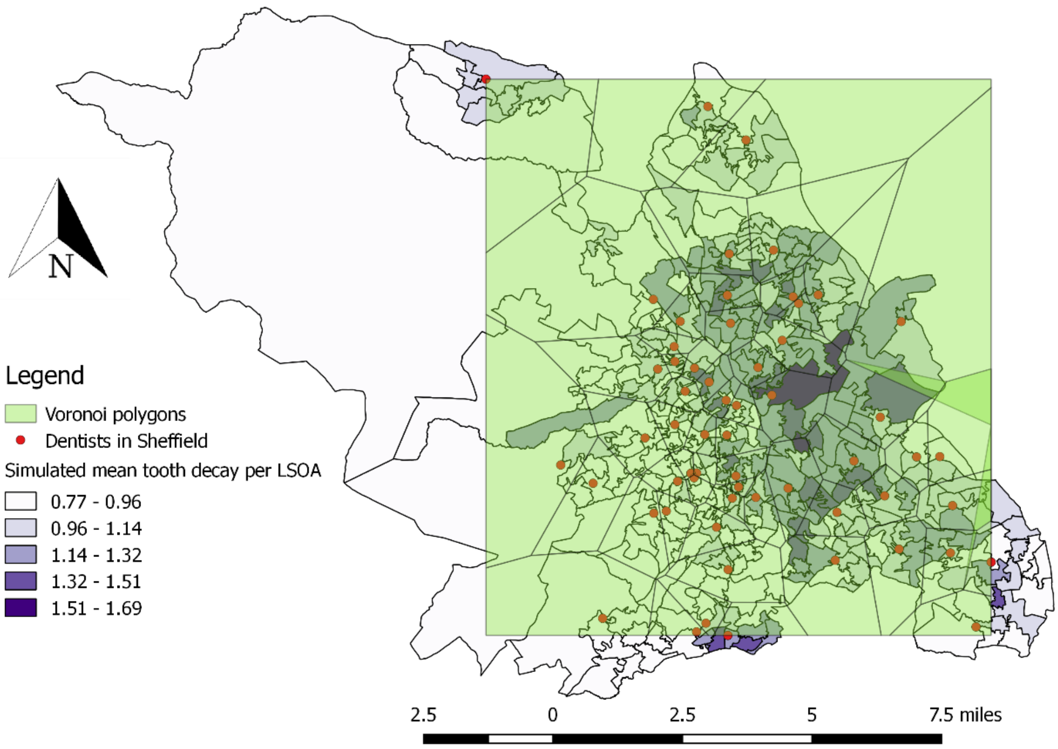
Patterns other than the spatial distribution of tooth decay in the city can also be studied. Thanks to data kindly provided by Kate Jones (National Consultant in Dental Public Health at Public Health England) on the location of all dental services in Sheffield, it was possible to map this along with the tooth decay scores (Figure 33).

Figure 33 – Tooth decay in Sheffield with dentist locations



From a visual analysis of the map it can be seen that dental surgeries tended to be located in areas not experiencing the highest levels of tooth decay. The areas characterised by the darkest purple shading (i.e. the highest levels of tooth decay) would seem to be poorly served by these services. This would be in line with the inverse care law theory (Hart, 1971), although of course a visual inspection is never going to be a thorough enough way to evaluate this in a policy context. There are more technically accurate ways to judge the spatial distribution of such locations, and this section is merely demonstrating the types of analysis that could be undertaken in a dental context once spatial microdata on a given subject is available. One of the more popular techniques in the early days of spatial analysis in Dental Public Health was the use of Voronoi polygons (or Thiessen polygons – Bradley et al, 1978) to delineate areas that were closest to each point, relative to all other points. This has been applied to the map in Figure 33, and is seen below in Figure 34.

Figure 34 – Voronoi polygons applied to dentist locations in Sheffield



A visual inspection of the Voronoi polygons created in Figure 34 shows that the areas around the LSOAs with higher levels of tooth decay tend to be larger, indicating fewer dentists in these areas, with those that are present having to 'cover' a larger area. Again this is more of a visual inspection, and there are specific allocation tools that may help in exact calculations (Horner et al, 2007), but again this demonstrates the type of analysis that is now possible with this data.

4.11. Choosing the study areas for the agent-based models

Having run the spatial microsimulation model for Sheffield, the decision was then taken to study smaller areas within the city. This decision was made for both conceptual and practical reasons. From a conceptual perspective, running a model for the whole of Sheffield may lead to a lack of focus on neighbourhoods and local environments, which was the main idea behind this research. By conducting the research at the city level it would also be difficult to monitor the whole city, or zoom in and out of specific areas without missing processes that were occurring in other areas. From a practical perspective this decision made sense too. Trying to run a model with 452,014 individual agents interacting over a whole city with hundreds of dentist, shops and further educational locations would be hugely computationally intensive, and challenging to extract meaningful data from.

Given these concerns, it was decided that the spatial microsimulation data would be used to identify suitable study areas for the research, which would then be used in the agent-based models. The idea was not to create a ‘perfect’ cluster containing the areas with the 1st, 2nd, 3rd and 4th (etc.) highest (or lowest) levels of tooth decay, but rather to find regions of Sheffield that contained areas which generally had some of the highest scores in the city, and vice versa. This would mean that the theoretical pathways could be applied in two very different areas, to see if different effects were important in different neighbourhoods.

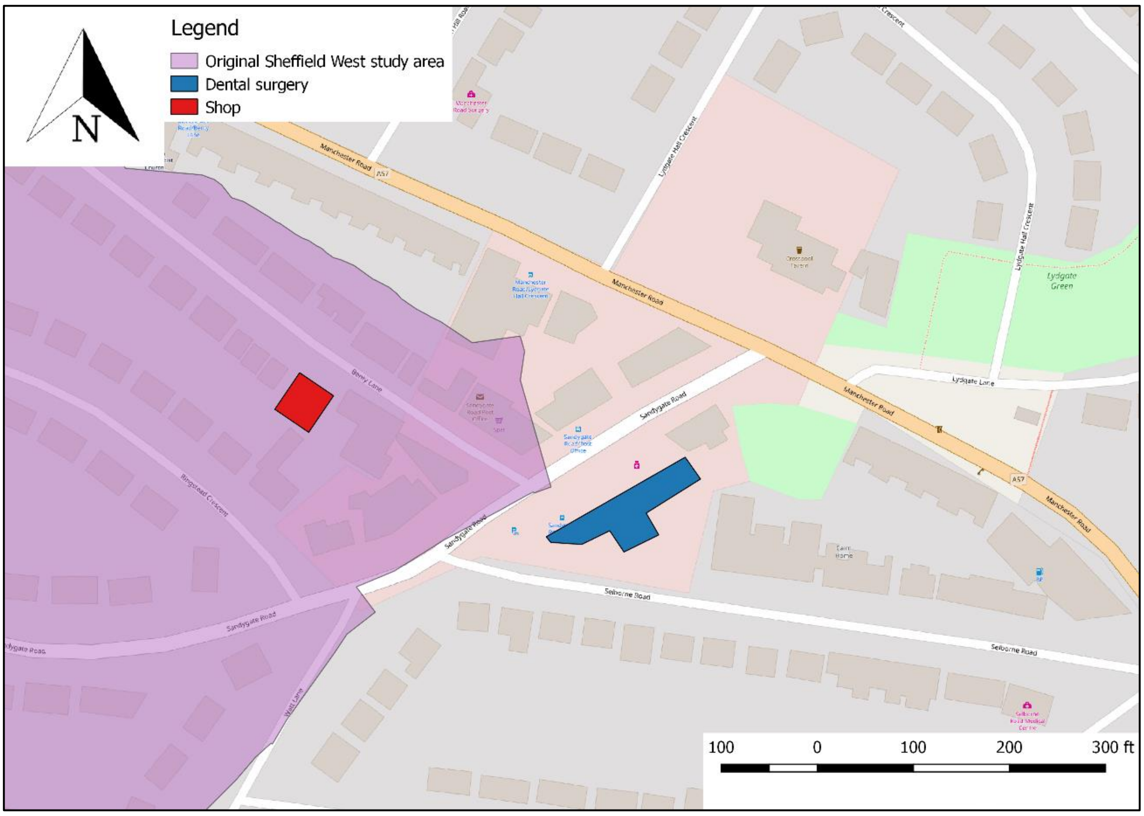
By mapping the mean tooth decay scores using QGIS, through the use of shapefiles from the UK Data Service’s UK Borders Boundary Data Selector (UK Data Service, 2011b), it was possible to highlight the areas with the 30 highest and lowest scores. To form a cluster these LSOAs needed to be adjoining. This analysis revealed two distinct clusters in the east and west of the city. In the east, including the areas of Wybourn, Attercliffe and Burngreave were LSOAs with the 1st, 2nd, 5th, 7th, 10th, 14th and 21st highest levels of tooth decay. In the west of the city, in the areas of Crosspool, Nether Green, Fulwood and Lodge Moor were LSOAs with the 1st, 4th, 11th, 12th, 14th, 20th and 26th lowest tooth decay scores. These areas are shown in Figures 36-39. The use of these two study areas seemed appropriate when conducting research into inequalities in Sheffield, given the city’s stark divide between the east and west on so many socio-demographic, economic and health related variables (Thomas et al, 2009). Both areas were made up of seven LSOAs, with the Sheffield East cluster having a total population of 8524 individuals, compared to a population of 8644 in the Sheffield West cluster. This represents another advantage of using spatial microsimulation modelling to investigate inequalities in oral health – while it would be possible to conduct a city wide analysis, due to the way the method has been applied smaller areas, and clusters of these, can also be selected for comparative analysis.

Given the shape and positioning of some of the LSOAs, buffers were added to the clusters in an attempt to counter border effects. For example, in the Sheffield West cluster, a dentist was located just outside the boundary of one LSOA, but it seems highly unlikely that people who live just inside that LSOA would discount it as a local service due to a somewhat artificial administrative boundary. This is demonstrated in

Figure 35, where the dark purple shading indicates the extent of the LSOA in question, with the blue building representing the building within which the dentist is located. The red building represents a shop that narrowly falls inside these same boundaries. The buffers that were applied subsequently included any facilities just outside the original clusters.

This is a difficult matter to deal with conceptually, as it is impossible to say that people would not travel to or use facilities outside of the clusters, but the buffers aimed to minimise this effect by including locations that would obviously be considered as 'local'. This issue is a more general one within geographical research, particularly where map boundaries are concerned. For example, even if the model had been conducted at the city level, it is still entirely possible that people may travel to Chesterfield, Rotherham or other nearby towns to use services that do not fall within the Sheffield city limits. The buffers also avoided roads at the edge of the clusters being cut off, which would have limited the movement of the agents in an unnatural way.

Figure 35 - Location of a dentist in Crosspool relative to the original Sheffield West boundaries



These newly created buffers were used to define the rest of the spatial data needed for the agent-based models. Data for supermarket locations for example were taken from the Geolytix website (Geolytix, 2016), which contained the postcode of each of these locations, as well as the easting and northing. The easting is the equivalent of the x coordinate (longitude), while the northing is the equivalent of the y coordinate (latitude). This made it easy to convert these locations into point data, which could then be mapped. The buffers for each area were then used to clip these shops locations. Clipping involves extracting data from within a defined spatial unit, in this case the shapefiles that formed the buffers for each of the study areas. This left two shapefiles containing data on the shop locations for the two study areas. A manual search of Google Maps (2016a) was conducted to identify any other shops that could be included. This was limited to those that were judged to sell a large enough range of products to support a weekly shop, rather than just confectionary, newspapers and tobacco. These locations were manually added where necessary to the existing shapefiles.

A similar procedure was undertaken for the dentist location data, this time without the additional manual search due to the completeness of the data. Finally, shapefiles for further education facilities were created. This data was not clipped as no comprehensive location file exists for these locations, so instead a manual search using Google and Google Maps (2016b) was conducted to obtain these locations. These facilities were roughly defined as those which provided either apprenticeships or further training, mainly for young adults (typically 16-24 year olds). This search yielded two locations for the Sheffield East study area, and none in the Sheffield West study area. Road and building data (Digimap, 2016) were also clipped to allow the agents to move around the model, and to add realistic detail in the case of the buildings. The integration of buildings into the model also allowed for specific locations (dentists, shops, etc.) to be designated as locations for the agents to visit.

Figures 36 and 37 demonstrate the exact locations of the two study areas within Sheffield, superimposed over an OpenStreetMap (2016) background, using the OpenLayers plugin (Sourcepole, 2016) in QGIS. The LSOA boundaries of the clusters are represented in pink, while the buffers applied to each study area are shown in yellow. As can be seen, particularly in the case of the Sheffield East study area in Figure 36, these buffers proved necessary in order to avoid 'cutting off' areas, and features within these such as roads, that would likely play a role within the study area. Figures 38 and 39 also demonstrate the locations of each study area in relation to the rest of the city, and the other LSOAs within Sheffield.

Figure 36 – Location of the Sheffield East study area within Sheffield, with original boundaries (pink) and new buffer (yellow)

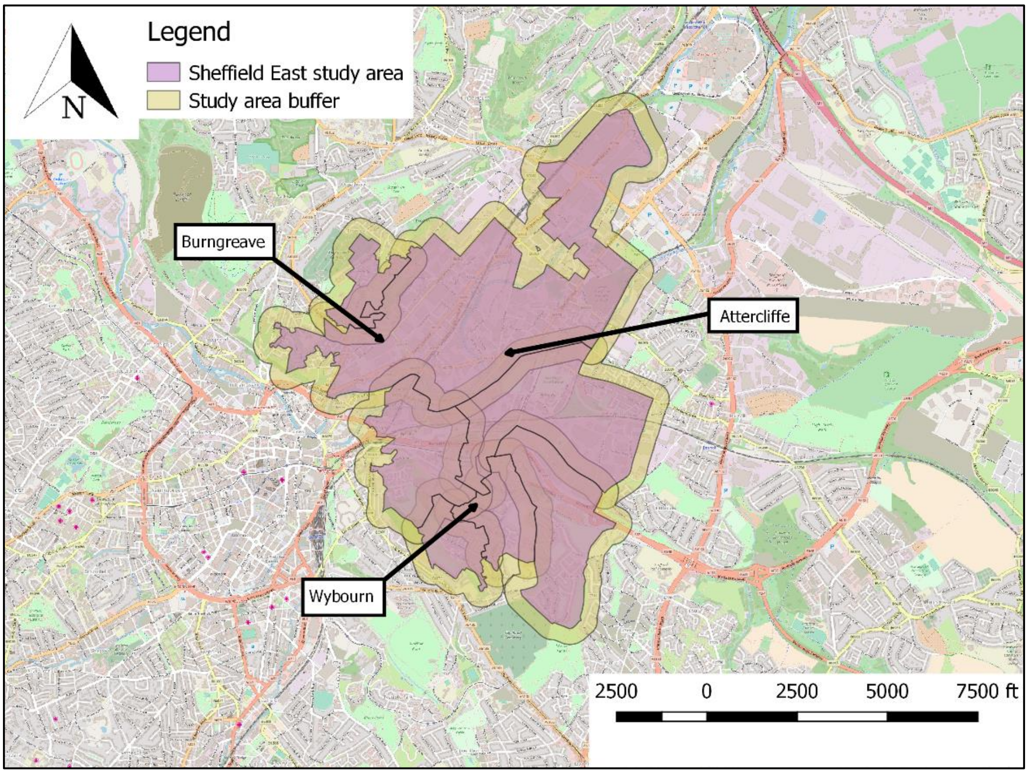


Figure 37 – Location of the Sheffield West study area within Sheffield, with original boundaries (pink) and new buffer (yellow)

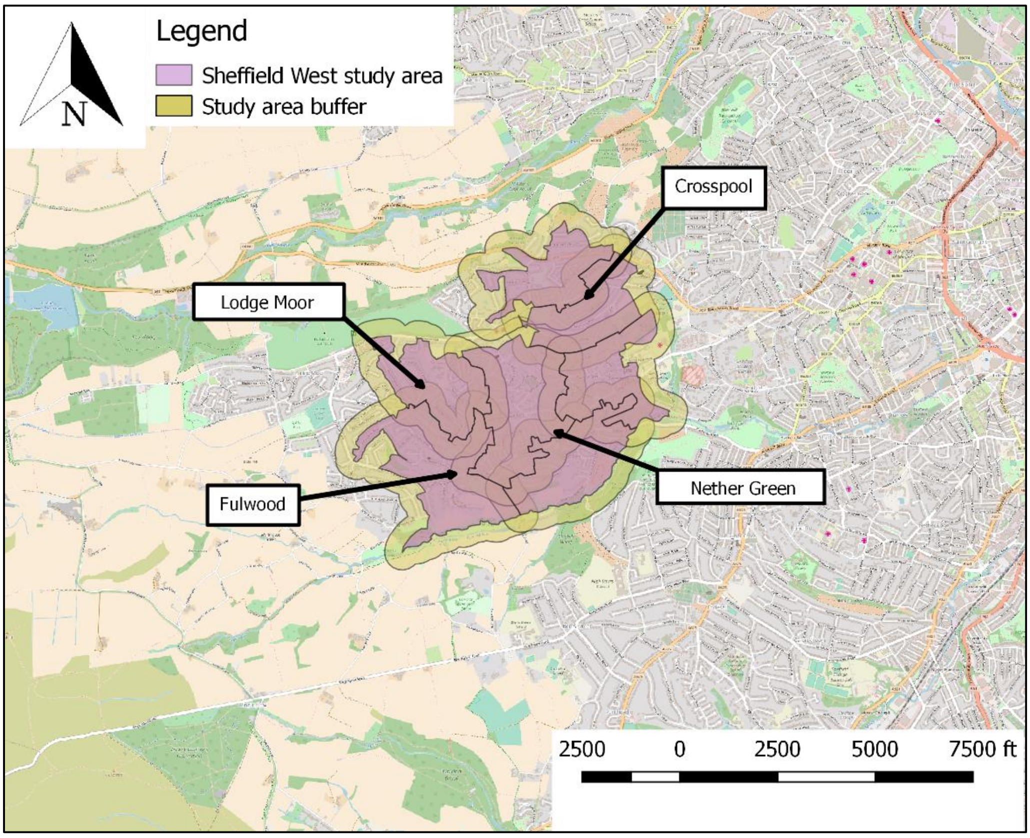


Figure 38 – Location of the Sheffield East (right) and Sheffield West (left) study areas in relation to each other within Sheffield

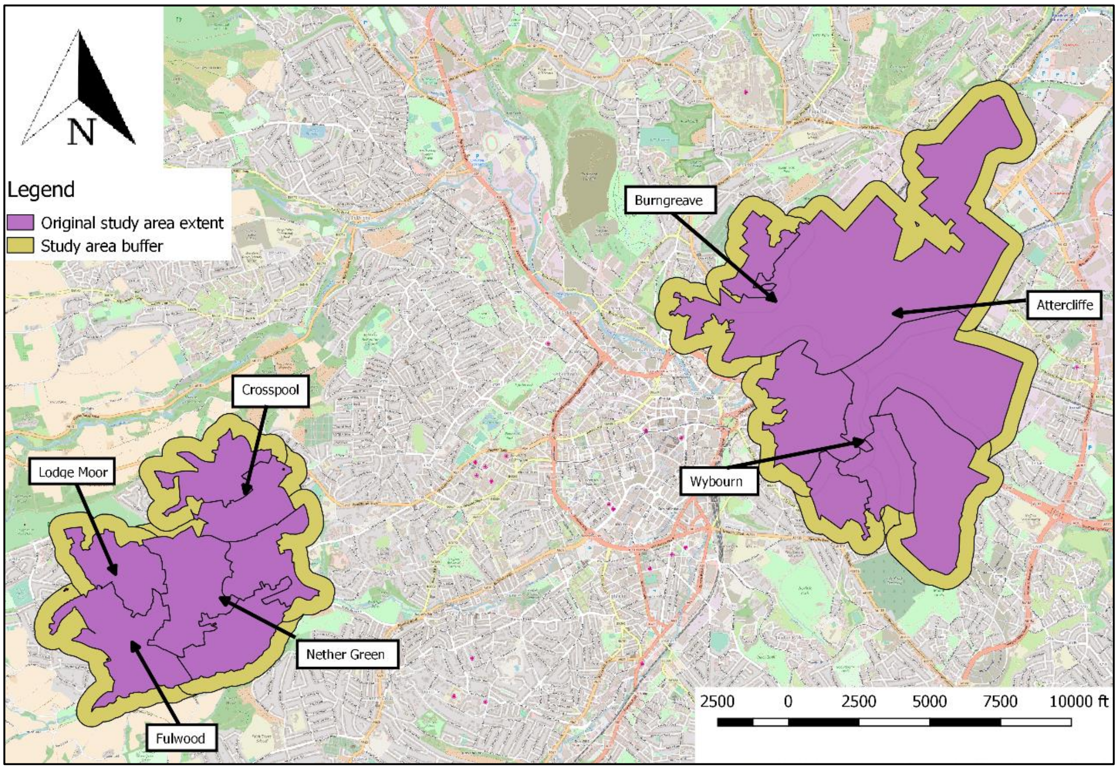
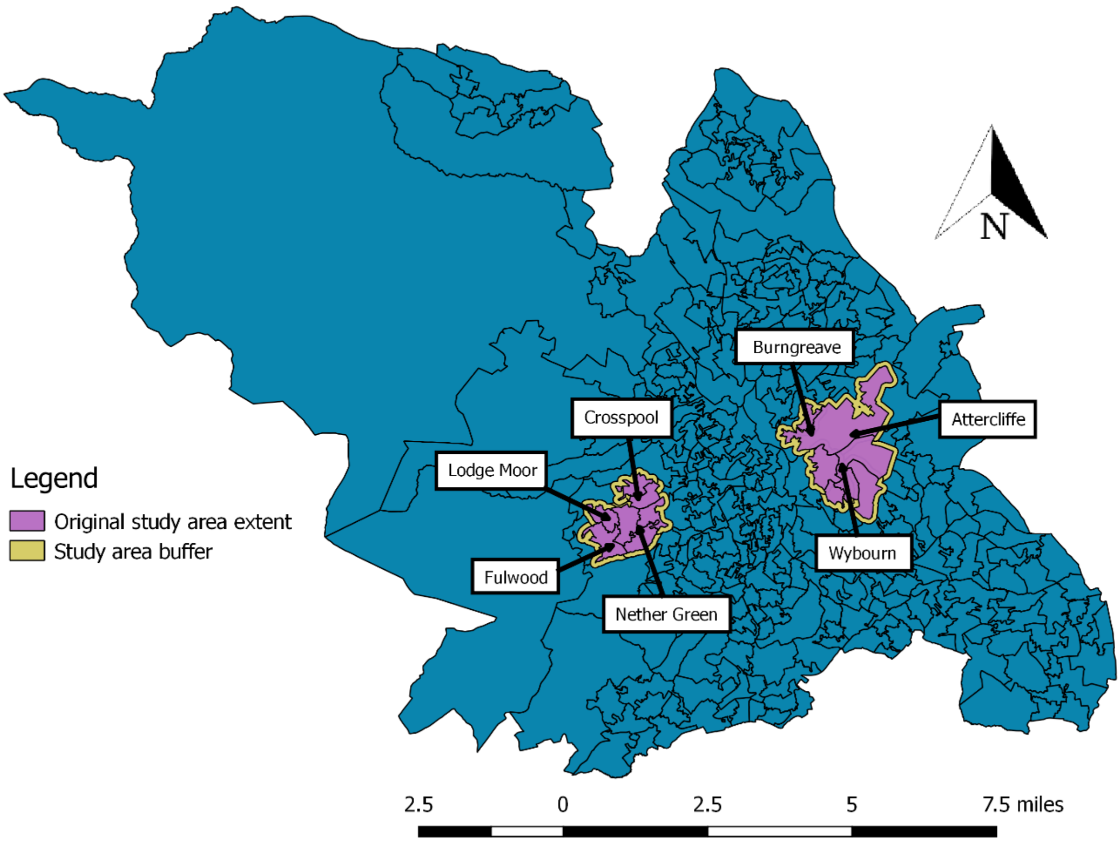


Figure 39 – Alternative view of the Sheffield East and Sheffield West study areas within the LSOA boundaries of Sheffield



4.12. ‘Painting a picture’ of the study areas

Ballas et al. (2005a) describe spatial microsimulation as being able to ‘paint a picture of the possible (or most probable) life of households of a city or region’ (p.14). Thanks to the selection of the study areas, and the aggregating of the individual data, it was now possible to get more insight into the characteristics of each of the areas. Table 16 shows the characteristics of the two study areas represented by the constraint variables and tooth decay scores.

Table 16 – Constraint characteristics and tooth decay scores for the two study areas

Variable	Sheffield East		Sheffield West	
	n	%	n	%
Age - 16-24	1383	16.22	822	9.51
Age - 25-34	2028	23.79	1003	11.6
Age - 35-44	1668	19.57	1563	18.08
Age - 45-54	1236	14.5	1693	19.59
Age - 55-64	952	11.17	1527	17.67
Age - 65-74	689	8.08	1034	11.96
Age - 75-84	422	4.95	707	8.18
Age - 85+	146	1.71	295	3.41
Females	4207	49.35	4544	52.57
Very good health	2760	32.38	4056	46.92
Good health	3100	36.37	3203	37.05
Fair health	1630	19.12	1059	12.25
Bad health	805	9.44	247	2.86
Very bad health	229	2.69	79	0.91
Car ownership	4269	50.08	7864	90.98
Has a degree	1084	12.72	4760	55.07
NS-SEC(1.1)	66		450	
NS-SEC(1.2)	210		2000	
NS-SEC(1)	276	3.24	2450	28.34

Variable	Sheffield East		Sheffield West	
	n	%	n	%
NS-SEC(2)	983	11.53	2772	32.07
NS-SEC(3)	806	9.46	1172	13.56
NS-SEC(4)	603	7.07	798	9.23
NS-SEC(5)	579	6.79	288	3.33
NS-SEC(6)	1854	21.75	598	6.92
NS-SEC(7)	1837	21.55	324	3.75
NS-SEC(8)	1586	18.61	242	2.8
Tooth decay (mean)	1.44		0.8	

As can be seen in Table 16, the two areas differ greatly on a number of variables. Sheffield East had a higher proportion of residents aged 16-34 for example (40.01%) compared to Sheffield West (21.11%), while Sheffield West has a higher proportion of residents aged 65 and over (23.55%) than Sheffield East (14.74%). Sheffield West also has a slightly higher proportion of female residents (52.57% versus 49.35%). Although the number of residents with perceived ‘good’ health is very similar between the two areas (37.05% for Sheffield West, 36.37% for Sheffield East), the proportion of individuals with perceived ‘very good’ health is much higher in Sheffield West than Sheffield East (46.92% versus 32.38%). Sheffield West also has a far higher proportion of car owners (90.98%) and degree holders (55.07%) compared to Sheffield East (50.08% and 12.72% respectively). Sheffield West also has a larger proportion of its residents in NS-SEC categories 1 and 2 (60.41%) than Sheffield East (14.77%), while the mean tooth decay scores also differ greatly (1.44 for Sheffield East versus 0.8 for Sheffield West). This paints a picture of Sheffield West as an area with a more elderly population, with most residents experiencing ‘good’ or ‘very good’ health, being likely to own a car, more likely to have a degree, and more likely to occupy positions towards the top end of the job market. It should again be noted that the two study areas were not selected based on any kind of socio-economic or demographic differences. Rather, they were selected based on clusters of tooth decay scores, with the stark socio-economic differences between the two areas only emerging following further analyses of these clusters. This adds weight to the theory that socio-economic circumstances and differences between them also reflect differences in health conditions (Wilkinson and Marmot, 2003), and that it would be hard to consider one without the other.

What is more interesting however is studying the trends associated with the additional target variables from the theoretical pathways that have been simulated. The socio-economic data above is interesting, however such a picture could be obtained without the use of spatial microsimulation, through taking data for each LSOA from the Census. The trends in the target variables in this research could not be assessed without this method however. This will be the next focus of this section, with the data for the target variables presented in Table 17.

Table 17 – Trends in the additional target variables for both study areas

Variable	Sheffield East		Sheffield West	
	n	%	n	%
Delayed treatment due to cost	1904	22.34	1347	15.58
Use toothbrush/paste	4702	55.16	5804	67.14
Cost affected treatment type	2527	29.65	2101	24.31
Psychological discomfort - very often	643	7.54	265	3.07
Psychological discomfort - often	533	6.25	361	4.18
Psychological discomfort - occasionally	1150	13.49	954	11.04
Psychological discomfort - hardly ever	570	6.69	617	7.14
Psychological discomfort – hardly	5628	66.03	6447	74.58
Had advice about smoking	1359	15.94	777	8.97
High sugar intake	4255	49.92	4295	49.69
Dental attendance - regular	4756	55.8	6403	74.07
Dental attendance – occasional	802	9.41	735	8.5

Variable	Sheffield East		Sheffield West	
	n	%	n	%
Dental attendance – trouble	2966	34.8	1506	17.42
Dental attendance – never	0	0	0	0
Fluoride intake - 1350-1500	6569	77.06	6564	75.94
Fluoride intake - 1000-1300	1522	17.86	1517	17.55
Fluoride intake - 550 or less	23	0.27	38	0.44
Fluoride intake – none	410	4.81	0	0
Sweets - 6+ per week	1526	17.9	1325	15.33
Sweets - 3-5 per week	1721	20.19	1858	21.49
Sweets - 1-2 per week	2345	27.51	2587	29.93
Sweets - less than once per week	1325	15.54	1539	17.8
Sweets - never	1607	18.85	1335	15.44
Severity of difficulty going out (average)	0.16		0.08	
Cakes - 6+ per week	2232	26.18	2483	28.73
Cakes - 3-5 per week	2214	25.97	2418	27.97
Cakes - 1-2 per week	2458	28.84	2310	26.72
Cakes - less than once per week	852	10	917	10.61
Cakes – never	768	9.01	516	5.97
Clean teeth - twice a day +	5969	70.03	6935	80.23
Dental visits - every 6 months	3999	46.91	4953	57.3
Dental visits - at least once a year	1546	18.14	1967	22.76
Dental visits - at least once every two years	409	4.8	457	5.29
Dental visits - less than every two years	919	10.78	563	6.51

Variable	Sheffield East		Sheffield West	
	n	%	n	%
Dental visits - Only when in trouble	1651	19.37	704	8.14
Had advice about visit frequency	6163	72.3	6460	74.73
Rate access to dentist - very good	2842	33.34	3003	34.74
Rate access to dentist - good	3925	46.05	3628	41.97
Rate access to dentist – fair	1182	13.87	1209	13.99
Rate access to dentist - poor	470	5.51	628	7.27
Rate access to dentist - very poor	105	1.23	176	2.04
OHIP (average)	5.28		3.17	

As can be seen from Table 17, there are some interesting differences in multiple variables between the two study areas. Despite Sheffield West seeming to have a higher proportion of what might be considered more favourable characteristics for good oral health, the picture is more mixed than might have been expected. Sheffield East for example had a lower proportion of resident regularly using other dental hygiene products including toothbrushes and paste compared to Sheffield West (55.16% versus 67.14%). Interestingly high sugar intake in both areas was remarkably similar with only 0.23% difference between the two areas. It is noticeable though that Sheffield West had a higher proportion of residents who attended dental services regularly (74.07% versus 55.80%), and a smaller proportion who only attended in emergencies (17.42% versus 34.8%), while neither area had any residents who had never visited a dentist. The fluoride intake scores were interesting in that Sheffield East actually had a slightly higher intake of fluoride at the top end of the scale, with the scores for the two areas remarkably similar across this variable. It was also noticeable that the two areas (similar to the high sugar variable) scored very closely with regard to cake and sweet consumption. Such trends point towards the conclusions of Sanders et al. (2006) who disputed the idea that poorer individuals care less about their oral health. Despite generally more favourable dental visiting patterns in Sheffield West, the opinions on access to dental services were again similar across the two areas.

A further benefit of the data produced by spatial microsimulation models is that they can be further broken down through cross-tabulation to provide even more insight into certain groups. Any of the behaviour variables could be broken down by gender, age or NS-SEC for example. The data can also be policy relevant. For example, the South Yorkshire and Bassetlaw Oral Health Needs Assessment (Public Health England, 2015) recommends that 'the longest interval between dental examinations for adults should be 24 months' (p.89). As can be seen from the data in Table 17, 10.78% of residents in Sheffield East, and 6.51% of residents in Sheffield West had not visited the dentist in the previous two years. This marks these areas out as two that could potentially benefit from oral health promotion, or other similar schemes, to improve this. Similarly, the Department of Health made a number of recommendations for oral health (Department of Health, 2015), including 'clean[ing] teeth effectively twice a day with a fluoride toothpaste' (p.32), as well as reducing the amount and frequency of sugar intake. As can be seen from Table 17, Sheffield West had 80.23% of its residents brushing twice a day or more compared to 70.03% in Sheffield East. With regard to sugar consumption, Sheffield East and Sheffield West had very similar percentages for high weekly sugar intake (49.92% versus 49.69%), while 17.9% of residents consumed sweets six or more time a week in Sheffield East, compared to 15.33% in Sheffield West.

These figures could be used in relation to areas of policy relevance, given that the aforementioned recommendations would aim to improve on all of these figures. Such data could help policymakers identify areas of need, and in some cases the data could be used to aid and inform policy goals. Obviously this is a simplified example which would require more accurate analysis of the data and guidelines, but it gives another example of the advantage of using spatial microsimulation. Once policies have been enacted, data from spatial microsimulation models could also be used to assess the impacts of these on individuals and households within these microunits (Ballas et al, 2005c). An example of this can be seen in the work of Ballas et al. (2006), who tested the effects of future policies on health in the city of York.

4.12. Conclusions

This chapter, and the methods described in it, helped contribute towards achieving the second research objective of this thesis:

- To build simulation models capable of representing these theoretical pathways

The data from the spatial microsimulation model in this chapter would be used to form the basis of the agent-based models. These in turn were used to test the theoretical pathways on this representative population dataset. This population dataset contained key demographic and socio-economic data through the inclusion of the constraint variables, as well as each individual being assigned an appropriate value for each of the variables used to populate the theoretical framework described in Chapter 3. This allowed for an easier translation between the framework and individual characteristics.

More broadly, this chapter has demonstrated the benefits of using spatial microsimulation for studying population level oral health. The flexibility to combine Census variables with those of a behavioural or attitudinal nature from survey data is a novel approach, and one that is still underused in health inequalities research, and particularly within Dental Public Health. The creation of a custom dataset has already been shown to be valuable in helping to populate the theoretical framework used in this research, and helps to address the lack of publicly available datasets for use on topics such as this. The creation of such data also allowed for a more detailed picture to be painted of the individuals and households involved in the analysis (Ballas et al, 2005a), both spatially and statistically. The deterministic nature of the IPF models used in this research also meant that the same output was produced when starting from a given position, due to the lack of randomness in the models (Ballas et al, 2005c).

As alluded to throughout the chapter, there are also a number of limitations associated with the spatial microsimulation method, and the approach taken in this thesis, which mainly revolve around data limitations and sample size. These issues will be discussed in greater depth in Chapter 6 (Sections 6.5.3 and 6.5.4).

Chapter 5 – Agent-based modelling of tooth decay

'I think the next century will be the century of complexity' (Stephen Hawking – 1942-2018)

5.1 Introduction

As mentioned in Chapter 2, previous work has used a combination of spatial microsimulation modelling and agent-based models to simulate population health and demographic trends, including an investigation of mortality projections in Leeds, UK (Wu and Birkin, 2012), and a simulation of student migration patterns within the same city (Wu et al, 2008). Similarly, agent-based models have been combined with theoretical concepts, including the work of Auchincloss et al. (2011), who explored the role of economic segregation on income differences in healthy eating, using hypothesised relationships from literature to design their interactions. Research by Yang et al. (2011) also used survey data to calibrate their model which investigated urban walking patterns. Thus while many models have used combinations of microsimulation, theory and individual-based approaches, none have combined a theoretical framework with a dynamic simulation model of an accurately created synthetic population within a dental context. This research therefore represents a new opportunity within Dental Public Health research.

Some agent-based models have been concerned with the prediction and evaluation of certain future events and scenarios, particularly related to the spread of disease (Potter et al, 2012; Merler et al, 2013), and in some historical evaluations have had access to past data to help parameterise models (O'Neil and Sattenspiel, 2010). This research is more concerned with exploratory analysis however, given the difficulties in parameterising these models, and the lack of available data to validate any findings. Therefore, the trends in the data will be the focus of this research, rather than the absolute values from the simulation outputs. This was seen as the most appropriate use of the models, given that results of agent-based models should be interpreted conservatively, since they do not represent empirical tests but rather 'explore the extent to which a theory is plausible' (Johnson and Groff, 2014 – p.514).

Previous chapters have outlined theoretical pathways by which features of neighbourhoods may influence levels of tooth decay (see Chapter 3), as well as ways to create accurate populations to test such theories on (see Chapter 4). This chapter will now tie these elements together along with the introduction of the agent-based models, and provide an overview of the creation of the final simulation models used in this research. The following chapter is split into two sections. Section 5.2 will outline the design of the agent-based models used in this research. This section is presented in the style of the ODD protocol (Grimm et al, 2006), which was designed as a standardised approach for presenting the characteristics and setup of agent-based models, with the aim of aiding clarity and reproducibility. Section 5.3 will then present the simulation experiments, including details of relevant validation processes, as well as the results of the simulation models and a brief discussion of these. Results and validation are presented separately as they do not fall within the remit of the ODD protocol, which Grimm et al. (2006) have described as being ‘designed to describe the basic model’ (p.123).

The agent-based modelling was conducted in the multi-agent programming environment, NetLogo (Wilensky, 1999). NetLogo is object oriented, meaning a system can be modelled as a set of objects, which can be controlled and manipulated in a certain way depending on the purpose of the model or system. This software has been described as being ‘low threshold, no ceiling’ (Tisue and Wilensky, 2004), a central principle carried forward from modelling conducted in the original Logo language (Logo foundation, 2016).

5.2 The ODD protocol

The ODD protocol is split into three domains: ‘overview’, ‘design concepts’, and ‘details’. The ‘overview’ domain is divided into three further subsections. The first of these (‘purpose’) describes the purpose and context of a model, in order to provide ‘a guide for what to expect in the model description that follows’ (p.117). The second overview subsection is ‘state variables and scales’, which describes the structure of the model, the entities included in it, as well as how they are displayed, and which hierarchical levels may exist. The extent of the model’s spatial and temporal resolution should also be described here. The final overview subsection is ‘process overview and scheduling', which is concerned with listing the processes, both individual and environmental, that have been built into the models.

The 'design concepts' domain covers a wide variety of factors, including: emergence, adaptation, fitness, prediction, sensing, interaction, stochasticity, collectives, and observation. It is not compulsory to include all of these, but elements should be included where relevant to the model at hand. The 'details' domain is also split into three further subsections (similar to the 'overview' domain). The first of these, 'initialisation', details how relevant entities (i.e. individuals or the environments in the model) are setup at the start of each simulation, as well as the initial parameters of the state variables, and whether these are setup the same way each time. How initial values were chosen (i.e. arbitrarily, or based on theory or data) is also covered. The second subsection in the 'details' domain, 'input', is concerned with environmental conditions that change over time and space, which are dynamics imposed on some of the state variables. Details on how these are generated or obtained is also important, as Grimm et al. (2006) state that 'the model output gives the response of the model to the input' (p.119). Finally, the third subsection ('submodels') presents and explains the processes from the 'process overview and scheduling' section in more detail, including model parameterisation. 'Mathematical skeletons' (p.119) of models or full model descriptions can be given, and cover assumptions made by underlying rules and equations, how parameters values were chosen, and how the models were tested (i.e. verified) and calibrated.

Figure 40 shows the basic structure of the ODD protocol, and each of its sections will be discussed in order below.

Figure 40 – ODD protocol outline (Grimm et al, 2006)

Overview	Purpose
	State variables and scales
	Process overview and scheduling
Design concepts	Design concepts
Details	Initialization
	Input
	Submodels

5.2.1. Overview - purpose

The purpose of the agent-based models created in this thesis was to understand how (if at all) the local environment and contextual features of neighbourhoods (as defined by the dental, and health inequalities literature) influence spatial inequalities in tooth decay in an adult population, using an exploratory approach. The main aim of using this method was to find the most influential theoretical pathways with regard to tooth decay within different neighbourhoods in Sheffield.

The need for the use of these models stems from a number of reasons. Traditional statistical techniques such as regression are not suited to complex scenarios involving adaptations, interactions, or feedback mechanisms (Auchincloss and Diez Roux, 2008). While Gelman (2005) illustrates the benefits of multilevel modelling, particularly being able to estimate the predictive effects of an independent variable separately from others, while also being able to control for correlation and dependence between the individual and area level variables, it has also been stated that ‘these models necessarily simplify complex interrelations’ (Auchincloss and Diez Roux, 2008 - p.2). As a result, models tend to segregate different elements from others, and fail to take feedback loops and adaptation mechanisms into account, although the complimentary nature of this method to agent-based models is also noted. Spatial microsimulation, while an important part of this research, has also been described as not suitable for the study of long term behavioural trends (Ballas et al, 2005a).

Agent-based models, however, have the ability to track agent characteristics of interest as simulations progress and interactions occur (Gorman et al, 2006). Through this it may be possible to gain better knowledge of the processes occurring at the small area level that affect tooth decay. Agent-based models can help research move beyond descriptive analysis, in order to test theoretical hypotheses which in turn may offer a better insight into the problem at hand (Johnson and Groff, 2014; Cerda et al, 2014).

5.2.2 Overview – state variables and scales

The models consisted of two hierarchical levels: the individual level, and the neighbourhood environment. Individual agents comprised a number of state variables derived from the spatial microsimulation models, including values for the following census variables: gender, age, education level, quality of health, NS-SEC status, and car ownership. The following oral health related variables from the ADHS were also assigned via the spatial microsimulation modelling: cost affecting delays in treatment, dental hygiene product use, cost affecting treatment type, psychological discomfort, advice on smoking, high sugar intake, general dental attendance, fluoride intake, sweet consumption, difficulty going out, cake consumption, tooth brushing frequency, advice on dentist visits, access to dentists, and a total OHIP score.

The software used in this research, NetLogo, can read in a variety of external file types, including comma separated values (.csv) files, through the use of the CSV extension. This was used to import the data outputs of the spatial microsimulation modelling conducted in R, with each line of the file assigned to individual agents systematically, creating a population of agents with the demographic characteristics for the two study areas in Sheffield. Thus additional data (i.e. the oral health related characteristics) created for agents through the spatial microsimulation process that were not available in the Census were also imported into the agents. This allowed agents to have attributes relevant to the theory driving the research, which could be updated and referred to as the theory based rules of the simulation were run. These processes will be described in more detail in Section 5.2.7. The attributes of the individual agents are listed in Table 18. Figure 41 gives an example of how the characteristics were stored 'within' each agent in NetLogo.

Table 18 – Overview of individual agent characteristics

Variable	Meaning	Parameter value range	Source
Gender	The assigned gender of each individual	Male; female	ADHS/Census 2011
Age	Grouped age bands at 10 year intervals	16-24; 25-34; 35-44; 45-54; 55-64; 65-74; 75-84; 85+	ADHS/Census 2011
Perceived quality of health	Self-assessed quality of health – Likert scale	Very good; good; fair; bad; very bad	ADHS/Census 2011
Car/van ownership	Whether individuals owned a car/van of not	Yes; no	ADHS/Census 2011
Education level	Whether individuals were educated to degree level, or below	Degree or higher; below degree level	ADHS/Census 2011
NS-SEC classification	The National Statistics Socio-economic Classification of each individual	Large employer/higher managerial; higher professional occupation; lower managerial/professional occupation; intermediate occupation; small employers/own account workers; lower supervisory/technical occupation; semi-routine; routine	ADHS/Census 2011

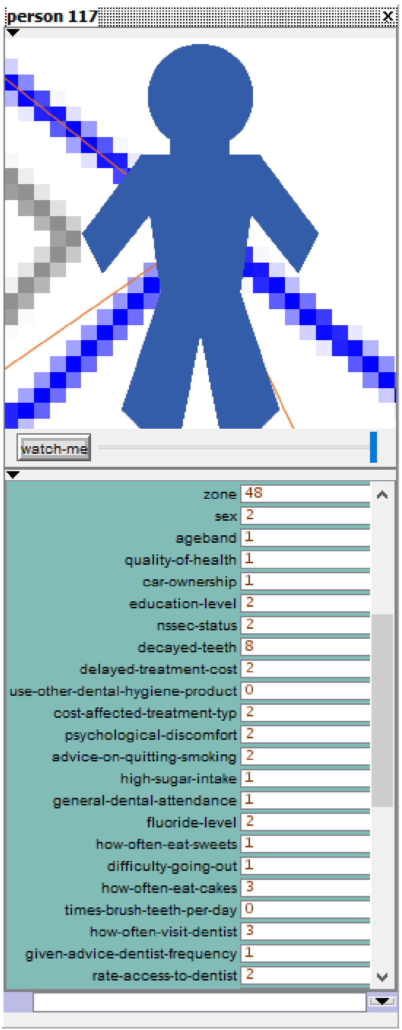
Variable	Meaning	Parameter value range	Source
		occupations; never worked and long term unemployed	
CostDly	Whether an individual had to delay treatment due to cost	Yes; no	ADHS
TPaste	Whether individuals used other dental hygiene products	Yes; no; I don’t have a toothbrush and/or toothpaste	ADHS
CostTyp	Whether cost affected the type of treatment or care an individual received	Yes; no	ADHS
PsycDisc	Whether an individual felt psychological discomfort in the form of feeling tense	Never; hardly ever; occasionally; fairly often; very often	ADHS
EvrAdSm	Whether individuals had ever been given advice about giving up smoking	Yes; no; I have never smoked	ADHS
HighSug	Whether an individual had a high sugar intake*	Yes; no	ADHS

Variable	Meaning	Parameter value range	Source
Regular	General dental attendance	Regular check-up; occasional check-up; only when having trouble; never been to the dentist	ADHS
Fluoride	Fluoride level (intake in parts per million)	1350-1500ppm; 1000-1300ppm; 550ppm or less; no fluoride	ADHS
Sweets	How often an individual ate sweets	6 or more times a week; 3-5 times a week; 1-2 times a week; less than once a week; rarely or never	ADHS
IpOut	Severity of difficulty going out	Continuous scale	ADHS
NCakes	How often an individual ate cakes	6 or more times a week; 3-5 times a week; 1-2 times a week; less than once a week; rarely or never	ADHS
ClnTthG3	How many times an individual brushed their teeth per day	Twice a day or more; once a day; never, less than once a day	ADHS
FreqDen	How often an individual went to the dentist	At least every 6 months; at least once every year; at least once every two years; less frequently than	ADHS

Variable	Meaning	Parameter value range	Source
		every two years; only when having trouble	
EvrFrqy	Whether an individual had ever been given advice about frequency of visits to the dentist	Yes; no	ADHS
Rteacc	Rate access to dental services	Very good; good; fair; poor; very poor	ADHS
TotOHIP	Total OHIP score	Continuous scale	ADHS

*Has cakes, biscuits, puddings or pastries, sweets or chocolate or fizzy drinks 6 or more times a week

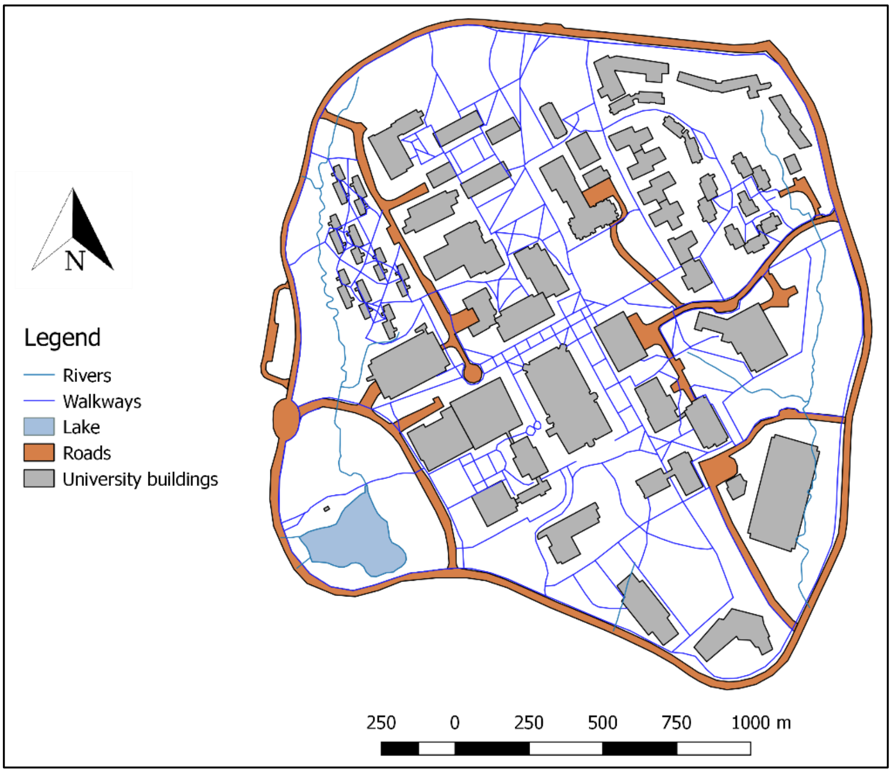
Figure 41 - Example of the characteristics given to agents, and how they were stored in the simulations



The environments in the models were occupied by these individuals, however individuals did not form larger social groups or collectives at any point. The environments themselves were created through the NetLogo GIS extension which is able to read in spatial data (both vector and raster), allowing for the design of specific landscapes, features and contexts within models. An example of the use of both types of GIS data in a NetLogo simulation can be seen in the work of Dawson et al. (2011), who designed a model investigating coastal flood management in the Welsh seaside resort of Towyn. In this research the boundaries of the study areas in Sheffield were set through the use of administrative shapefiles (vector data), which took the form of clusters of LSOA boundaries (as discussed in Chapter 4) gathered from the UK Borders Boundary Data Selector, accessed through the UK Data Service's Census boundary data collection (UK Data Service, 2011b). The agent-based modelling environment was set so that the 'edges' of the environments acted as boundaries which could not be passed, rather than having a wrap-around effect, where agents could reappear on the other side of the environment. This would not be a realistic portrayal of individual movement, and is an approach taken in other geographical agent-based models (Crooks, 2008). Agents could move between any of the LSOAs in the environment though. Chapter 2 highlighted the significant debate about the definition of neighbourhoods, and the difficulties in using administrative data to delineate these. Unfortunately, when conducting quantitative analysis this issue is hard to avoid, and the many methods for creating custom boundaries (Haynes et al, 2007; Vallee et al, 2015) were beyond the scope of this research. LSOAs were chosen as they represent small geographical areas that are also large enough to include facilities that may be of regular use to locals, particularly if combined in clusters. This represents the type of trade-off that can occur in such research.

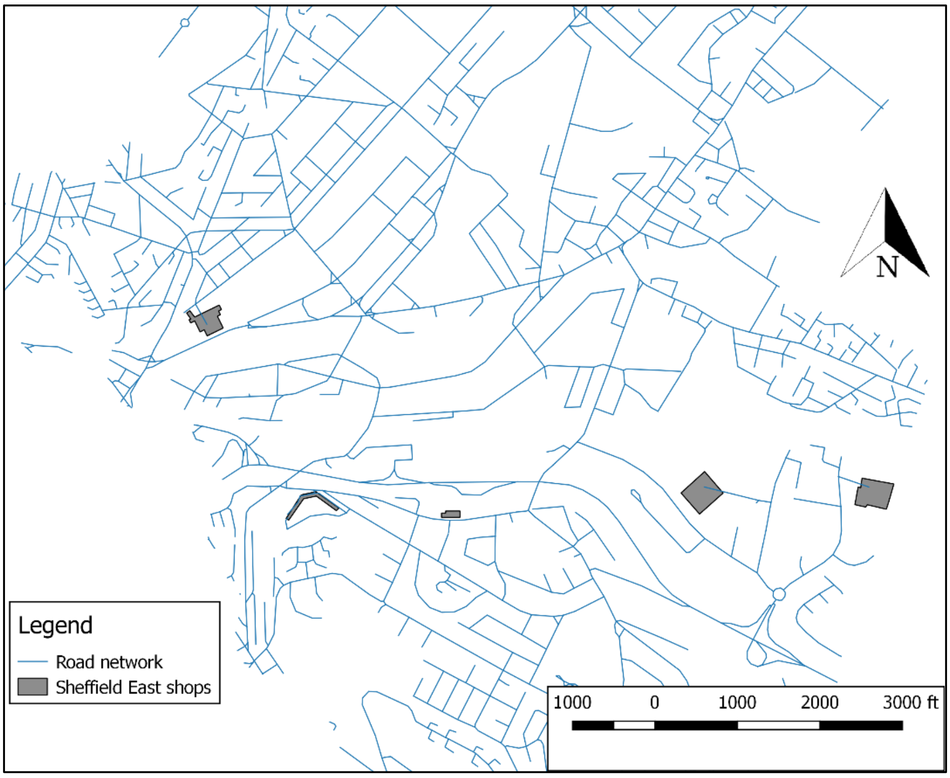
As mentioned in Section 4.11, further shapefiles were included which detailed the road network for the two study areas, in the form of Ordnance Survey stylesheets obtained through the Edina mapping service (Digimap, 2016). This allowed for the inclusion of 'geometry' in the model, an idea suggested by Crooks (2008) which includes point, line or polygon based features that allow for more realistic urban environments, and the study of the impact of geometry in potentially influencing the outcome of certain processes. These layers were subsequently clipped to fit the boundaries of the two study areas in QGIS. The aim was that the road networks would guide the movement of the agents, rather than having them move around freely across open space. Further shapefiles (also discussed in Chapter 4) were included which contained geocoded point data on the locations of shops, dentists and further education facilities within each of the study areas. It should be noted that no further education facilities were located in the LSOAs that made up the Sheffield West study area, so these were not present within this area in the modelling. The rest of the buildings in the respective areas were also included in an additional layer to create a realistic looking urban environment. The geographical elements of the model were made accessible to the agents through the use of a previous GIS based NetLogo model, a path finding agent-based model set on the campus of George Mason University (GMU), Virginia, USA (Zhou, 2016). An example of the layout of this model can be seen in Figure 42, and this process will be explained in full in the 'submodels' section of the protocol (see Section 5.2.7).

Figure 42 – Layout of the A-star path finding algorithm model for George Mason University (Zhou, 2016)



Additional shapefiles that contained only the location of the buildings which would act as destinations (identified using the point data for the locations of dentists, shops and educational facilities) were also required. This was because the original GMU model used the building layer shapefile to identify the destinations for agents. In the example of a university campus, all of the buildings acted as destinations for students at some point during the simulation, which is logical as they would all have some purpose for at least a sample of the student population. However, this is not practical when simulating larger (relative to a university campus) urban areas with hundreds of buildings, as the code in the model would assign each of these as destinations. Therefore, a new shapefile layer was created from the original building layer that contained only the buildings of interest (e.g. shops) for this research (see Figure 43). This was included in the simulation along with the original buildings layer, effectively overlaying it. This allowed the code to recognise only the buildings that were of relevance, and define these as the destinations for the agents.

Figure 43 - Shapefile containing only (a sample of) buildings of interest (in grey) in the Sheffield East study area



This was problematic when two destination buildings occupied the same patch in the model. One example in the Sheffield West study area occurred with a dentist and a supermarket being located on opposite sides of the same road. In the Sheffield East model a similar problem was encountered. As only one destination could be placed in each patch, on these occasions one of the destinations was moved slightly so that it fell in an adjacent patch. Thus, two of the destinations were slightly removed from their original coordinates, but still near enough to be geographically accurate. Further customisation of the code allowed for buildings to be designated as different types of destinations, allowing agents to 'know' the difference between shops, dentists and FE facilities. These were represented with a star, and also colour coded to make them easier to identify in the GUI. Locations within the model did not differ by type. For example, dentists were not differentiated by whether they offered NHS or private services, and this is a potential area for future research.

In addition to including relevant buildings, other neighbourhood based variables were included as 'background data'. These did not represent physical locations like the aforementioned buildings, but rather important neighbourhood level characteristics of the local environment. These included material circumstances, employment, health behaviours and social capital. As alluded to by the work of Cummins et al. (2007), in some situations it is preferable for neighbourhood level variables to not represent aggregated individual level data, and this approach was followed as much as possible in this research. Material circumstances were represented by data on median house prices (ONS, 2015) and Section 3.5 demonstrated that this variable should act as a reasonable proxy (Nkosi et al, 2011; Tunstall et al, 2013). Employment was represented by estimated weekly income data (ONS, 2009), which was seen as a suitable proxy as employment will directly influence income. While income data could have been estimated for individuals using the spatial microsimulation modelling, the income data in the ADHS had very few responses for both individual income (n=910) and household income (n=145). This would have likely reduced the sample size from the ADHS, which could have adversely affected the accuracy of the spatial microsimulation.

Data on years of potential lost life were taken from the IMD health domain (IMD, 2015a) to act as a proxy for health behaviours, which seemed a reasonable assumption given the links between oral health outcomes and other types of illnesses (Bailey et al, 2004; Chroinin et al, 2016). Finally, social capital was represented by crime data, taken from the IMD crime domain (IMD, 2015b). Crime was taken as a proxy for community cohesion, as similar measures of social disorder including homicide rate have been used in an oral health context for this concept previously (Pattussi et al, 2001). The latter two variables were available at the LSOA level, which meant that this data could easily be assigned to each LSOA using QGIS. However, the former two variables were only available at the MSOA level, and due to data shortages, this information was manually added to the relevant LSOAs. This meant that some neighbouring LSOAs shared the same score for some variables, however there was still enough difference across the two study areas in these variables to make the exercise worthwhile. The background variables and their data sources are presented in Table 19.

It should be noted that the four background variables were also assigned to individuals. This was a practical decision, as during the simulation it was far less computationally intensive for agents to check their own characteristics than to stop to check the characteristics of the patches of the model at every step.

Table 19 – Background variables used in the simulations

Variable	Meaning	Parameter value range	Source
Material circumstances	Median house price statistics for small areas (MSOA)	£63,000 - £256,500 (continuous scale)	Office for National Statistics (2015) - Neighbourhood Statistics
Employment	Model based income estimates - per week (MSOA)	£449.62 - £1321.69 (continuous scale)	Office for National Statistics (2009) – Neighbourhood Statistics
Health behaviours	Years of potential life lost – age and sex standardised measure of premature death (i.e. before 75 - LSOA)	49.192 - 79.736 (continuous scale)	Indices of Multiple Deprivation (2015) – Health domain
Social capital	Crime score - combined data on violence, burglary, theft and criminal damage per 1000 individuals (LSOA)	-1.302 - 0.637 (continuous scale)	Indices of Multiple Deprivation (2015) – Crime domain

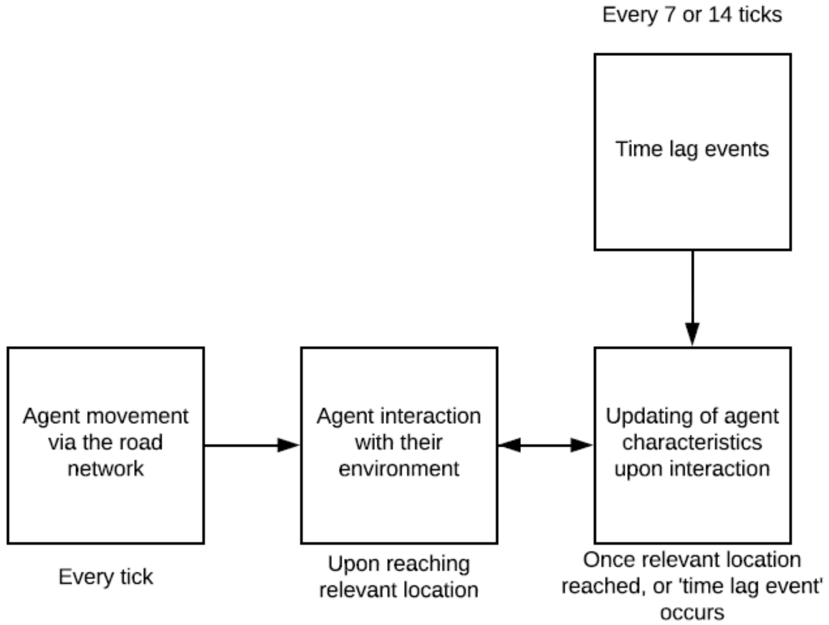
The final point in this section regards the temporal element of the models. In NetLogo models proceed in 'ticks', which are arbitrary time units. These ticks can be assigned values depending on what type of system is being studied, and the timescale over which this study takes place. In this research the ticks refer to days, with models running for 730 ticks to represent a timescale of 2 years. This was chosen based on advice from

Professor Zoe Marshman, from the Unit of Dental Public Health in the School of Clinical Dentistry, specifically that (while difficult to define an exact time period) tooth decay often takes 2-3 years to develop from start to finish. The lower bound of this (two years) was chosen for practical reasons, as adding another 365 ticks to each model run would have greatly increased the computational time to run the models.

5.2.3 Overview – process overview and scheduling

As mentioned in the previous paragraph, the models proceeded in daily time steps. Within each time step, three processes occurred in the following order: agent movement, agent interactions with the environment, and updating of agent characteristics. It was also possible for the 'updating of agent characteristics' to influence 'agent interactions with the environment', and this will be explained in more detail in Section 5.2.7. A fourth process, time lag events affecting agents, was separate and only related to the 'updating of agent characteristics' process. Agent movement concerned the way agents interacted with, and used the road network to navigate the environment in the model, which occurred every tick (or day). Agent interactions with the environment occurred when agents reached a certain point on the road network (i.e. relevant buildings), and as such there was not a defined time point, or tick on which this happened. The updating of agent characteristics occurred either when an agent interacted with an environmental location, or, as part of the fourth process in the model, when a time lag event influenced an agent's characteristics in some way. These time lags processes differed depending on the process in question, and were either at 7 or 14 day intervals. As per the guidelines set out in Grimm et al. (2006), the above four processes will be discussed in full detail in the 'submodels' section of the 'details' domain. The processes listed above and their scheduling are depicted visually in Figure 44.

Figure 44 – Diagram representing the scheduling of events in the agent-based models



5.2.4 Design concepts

Emergence: Aggregate tooth decay scores for each of the study areas emerged from the tooth decay scores of individual agents, which were in turn influenced by agent behaviours and interactions with their environment. These behaviours and interactions were represented by theoretical rules, with trends taken from the Dental Public Health literature used to define rules for tooth-brushing, diet, sugar intake, attendance rates, material circumstances, smoking, stress, purchasing power, financial constraints, dental knowledge, social capital, employment, social hierarchy position, education, and interactions with shops, dentists and further education facilities,

Adaptation: Agent characteristics could 'adapt' as the model progressed, as the scores associated with certain variables changed over time based on their own characteristics, as well as interactions with their environment. The changing of these scores had the potential to push them over/under a given threshold, which may change how processes affected them later in the simulation.

Sensing: Individuals were assumed to know their own socio-demographic characteristics (i.e. age, NS-SEC category, education, income), as these influenced the processes that affected them over the course of the simulation. Agents were also aware of the neighbourhood characteristics of where they were from (i.e. social capital, material circumstances), the values of which also affected them over the course of the simulations.

Interaction: Agents interacted directly with their environment through three mechanisms: interactions with the road network, interactions with point based locations, and interactions with environmental ('background') variables. Time lag processes also interacted with agent characteristics as the simulations progressed. These interactions were modelled explicitly. Agents did not interact with each other within the simulations.

Stochasticity: The behaviours of agents when visiting point based locations were based on probabilities. These were used firstly to differentiate the probability of visiting a given location based on differences in educational attainment (as per the literature), and also to allow for an element of randomness in the model. This was important when modelling reactive agents, where it would be a mistake to assume that expected behaviours would occur on every occasion. The use of attendance probabilities and a random number generator to determine visiting patterns meant that these could differ from their expected distributions.

Collectives: Individuals were not grouped into collectives. Individuals differed in their behaviours and interactions in some processes (i.e. visiting neighbourhood based locations) based on certain shared characteristics, but were not grouped by these variables explicitly.

Observation: For the testing of the models, the effects of each pathway as well as agent movement were observed process by process in a simplified environment. For the analysis of the final model only one population level variable was observed, this being the simulation of the tooth decay scores for each individual which was summed to create an aggregated decay score for each study area. This was used as the final outcome variable for each of the different simulations. These scores were compared across simulations to assess the effects of different combinations of theoretical pathways on the outcome variable.

The 'fitness' design concept was not modelled, as the research was more concerned with the overall trends in the data, and how these changed between simulations, rather than with the exact values emerging from each simulation. There was also no empirical data available to fit the data against. The 'prediction' design concept was also not modelled, as agents did not make predictions about their future state, instead acting on their characteristics at a given point in time.

5.2.5 Details – initialisation

Upon initialisation, the Sheffield East study area was occupied by 8,524 agents, while the Sheffield West study area was occupied by 8,644 agents. Each simulation run started at the beginning of the study time period, which equated to 2011 based on the use of Census data from this year to create the agent population (through the spatial microsimulation modelling). For this exploratory research, the initialisation of the models was influenced by the two experiments that were undertaken. The first experiment involved adding the theoretical pathways one at a time to consecutive iterations of the model, leading to each new initialisation of the model containing an extra pathway to the previous one. Due to this approach, in the first experiment the built environment did not change notably with the initialisation of each new iteration of the model, while the theoretical rules applied to the agents and the environment did.

While the ability to add constructs or variables of interest one at a time is similar to approaches that regular multiple regression models could run, within the agent-based models the agents and environments were dynamic, interactive and ever changing, adding an element that cannot easily be replicated using conventional statistical approaches. By adding the pathways one at a time, the overall effect on the model was assessed after each model run, with the idea being to identify which of the pathways influenced and changed the outcome score the most. This also allowed for analysis between the two study areas, to see if these effects were the same in both. The process of sequentially adding the pathways is represented in Table 20. Similar approaches have been taken before within geographical agent-based models, such as Crooks' (2008) work, which iteratively increased the percentage of preference for living with similar types of individuals in a model of residential segregation in London, finding that the most notable variation occurred at 50%, when segregation rose the most. The approach taken in this research was not too dissimilar to this, in that pathways were added iteratively to assess the impact on the outcome variable from these additions.

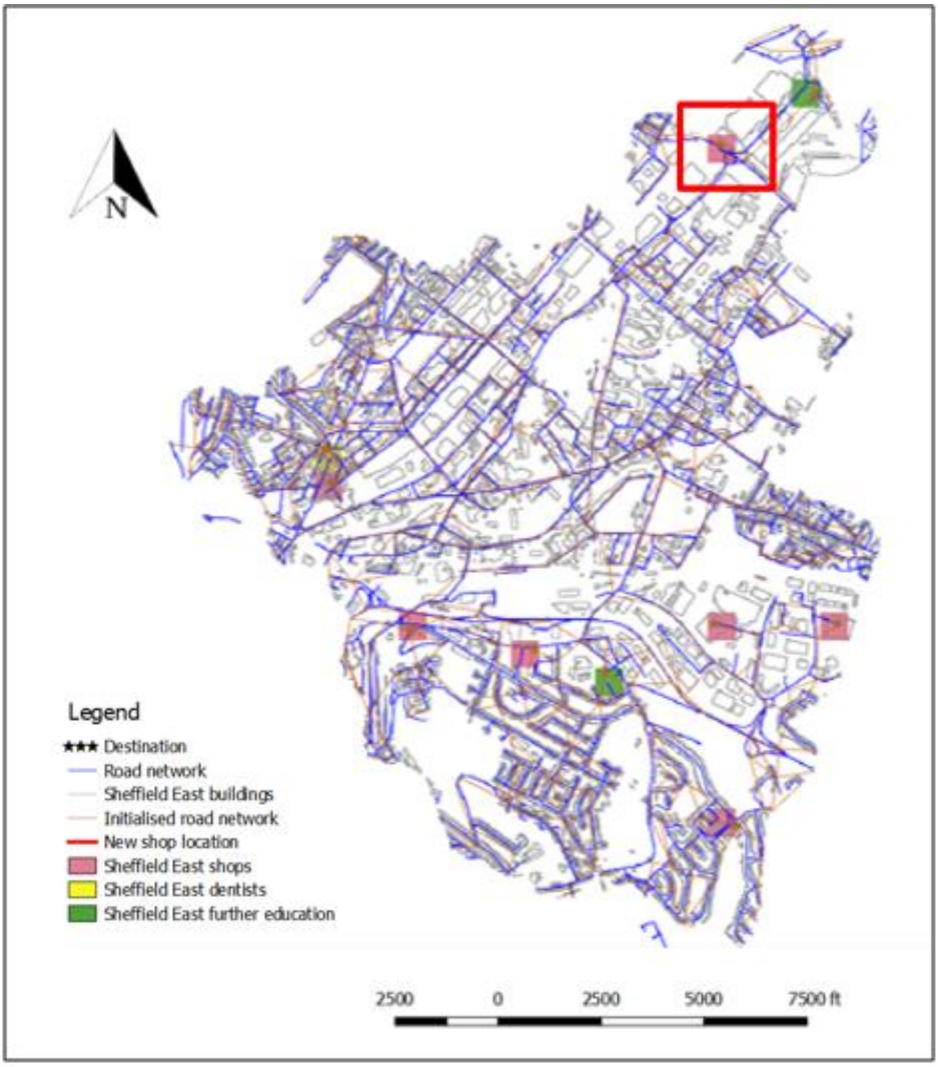
Table 20 – Simulation runs and included pathways in each iteration

Model run	Pathway 2.1	Pathway 2.2	Pathway 2.3	Pathway 3.1	Pathway 3.2	Pathway 3.3	Pathway 3.4	Pathway 3.5	Pathway 4.1	Pathway 4.2	Pathway 4.3	Pathway 4.4	Pathway 4.5	Pathway 4.6
1	X													
2	X	X												
3	X	X	X											
4	X	X	X	X										
5	X	X	X	X	X									
6	X	X	X	X	X	X								
7	X	X	X	X	X	X	X							
8	X	X	X	X	X	X	X	X						
9	X	X	X	X	X	X	X	X	X					
10	X	X	X	X	X	X	X	X	X	X				
11	X	X	X	X	X	X	X	X	X	X	X			
12	X	X	X	X	X	X	X	X	X	X	X	X		
13	X	X	X	X	X	X	X	X	X	X	X	X	X	
14	X	X	X	X	X	X	X	X	X	X	X	X	X	X

The second experiment could not be as easily replicated by traditional statistical models, and involved adding new physical attributes to each of the study areas which agents could interact with. This highlights the exciting potential of these models. In theory these physical attributes could include a range of additional buildings such as supermarkets and shops, or more abstract concepts such as advertising and marketing campaigns. Marketing campaigns promoting beneficial oral health practices could be placed at bus stops for example, to see what the effect on the model (compared to a base simulation) might be. In this research an additional shop was added to each of the study areas to investigate if, hypothetically, more shops in an area would impact on the tooth decay scores of residents. This approach was beneficial in that the theory surrounding shops and the effects of visiting these had already been programmed into the model. Thus the second experiment resulted in a change to the built environment from that seen in previous iterations of the model.

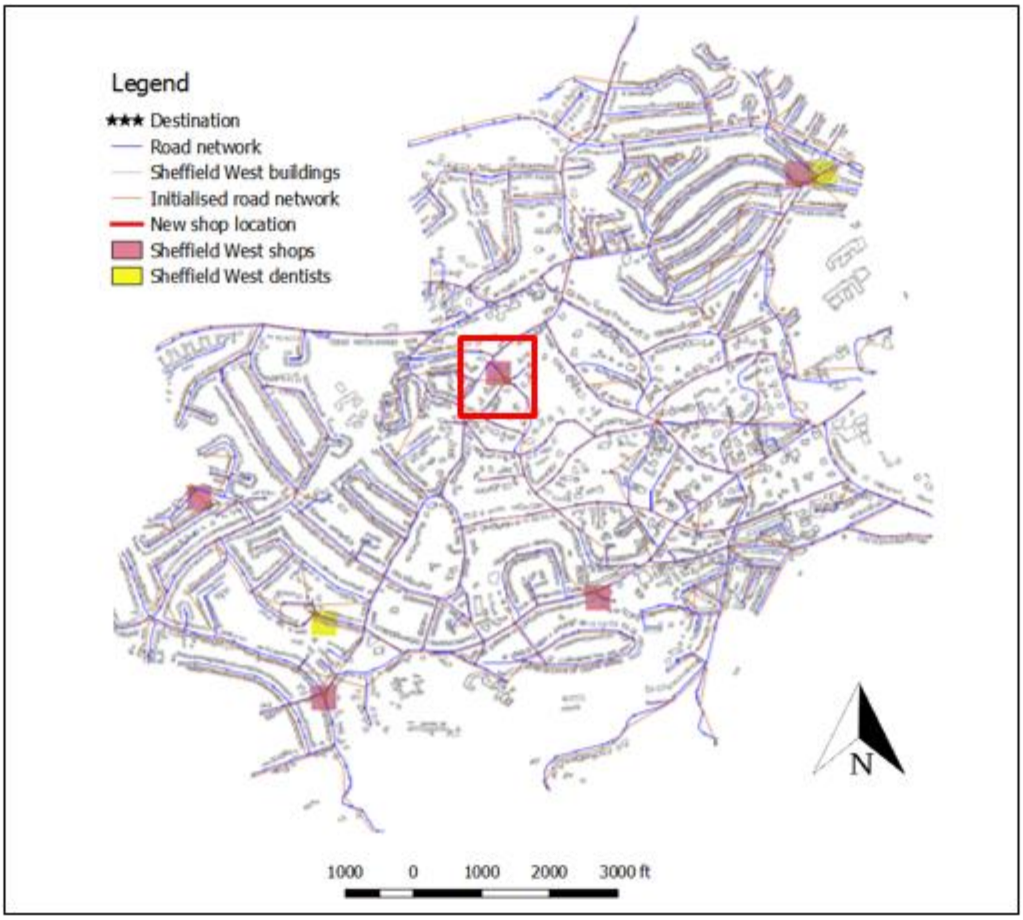
For the Sheffield East study region Google Maps (Google Maps, 2017a) was consulted alongside the extend of the study area in QGIS. Building locations were assessed from an aerial view to judge their suitability as a potential location for a new shop. This part of Sheffield contains large areas of industrial buildings, which were perceived as large enough spaces to accommodate a new shop. An area to the north-east of the study region was identified with an abandoned building which was chosen as the site for the new shop. This is shown in Figure 45.

Figure 45 – Location of new shop within the Sheffield East study area



While this location may seem a little arbitrary, the focus of this research was exploratory and therefore more concerned with demonstrating what the method can do, and help with, than applying the exact methods associated with location allocation models which may be used to assess suitable locations for new shops. Such methods have been studied in a dental context before (Horner et al, 2007), and could be added to future analyses. The new shop locations (along with the existing shops in the study area) were identified in the existing shapefiles in QGIS, and a new shapefile layer created to represent the (now) seven shops in the study area. This was loaded into NetLogo in place of the original shop location shapefile. A similar process was undertaken for the Sheffield West study area using Google Maps (Google Maps, 2017b), although as this area was almost exclusively residential a random building in the middle of the simulation study area was selected (most of the other destination buildings were located towards the periphery of the study area). This can be seen in Figure 46.

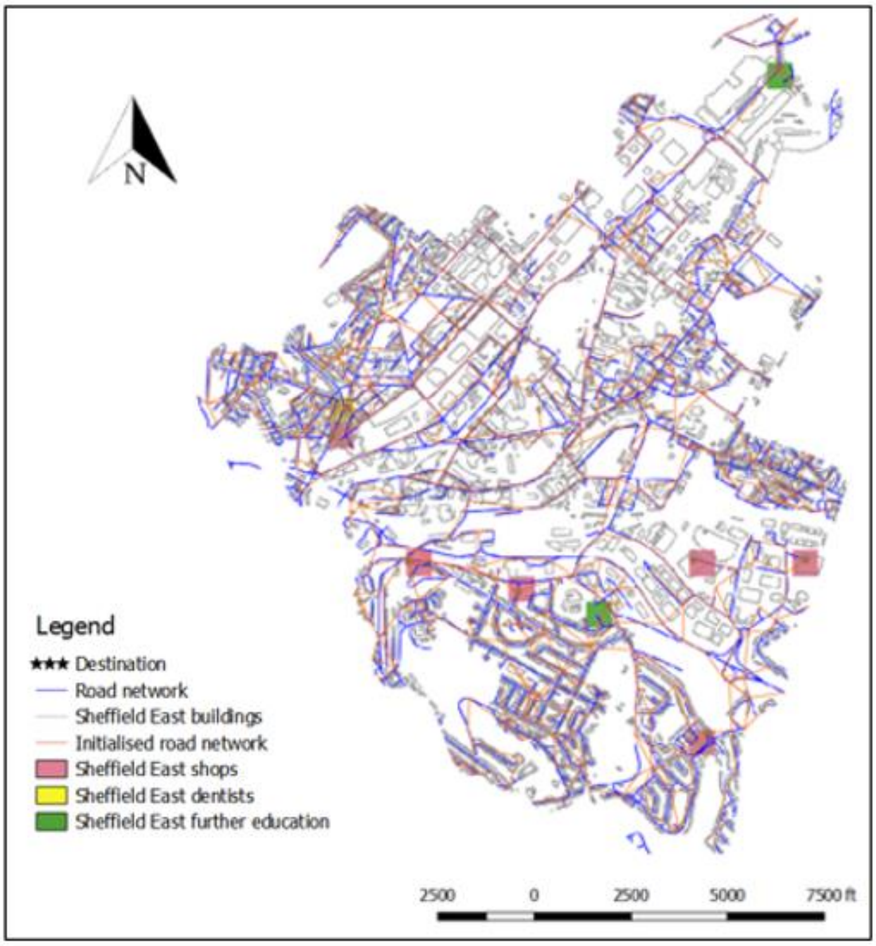
Figure 46 – Location of new shop within the Sheffield West study area



In addition to changes in the theoretical rules as part of the first experiment, and the physical environment as part of the second experiment, the road networks in the model also varied slightly in their setup each time the model was re-initialised. This was because in the GMU model the layout of the nodes and connections between them changed slightly each time the model was setup. As a result of this, the nodes and links created in the simulation did not always accurately reflect the road and path networks of the university. When tested on the models for the two Sheffield study areas this had the effect of creating slightly inaccurate road networks (see Figure 47), with nodes and connections moving through buildings at certain points.

With the exception of the new locations of the additional shops added for the second simulation experiment, the locations of all other shops, dentists and educational facilities, and the coloured patches associated with these, were consistent in their initialisation location in each iteration of the model. Finally, the starting location of each agent differed each time a different iteration of the model was initialised, with agents not always starting the simulation in their 'home' LSOA (i.e. the LSOA for which they had been created in the spatial microsimulation modelling).

Figure 47 – Example of the mismatch between the road network (blue) and the nodes/links (orange) in the Sheffield East study area



5.2.6. Details – input

Grimm et al. (2006) state that 'the dynamics of many IBMs are driven by some environmental conditions which change over space and time' (p.119), when referring to the effects of 'input' on agent-based models. The authors also state that model inputs may vary over space and time, and are 'imposed dynamics of certain state variables' (p.119), using the examples of precipitation and harvesting regimes in their document. The agent-based models used in this research did not include input features of this nature. While the addition of pathways one at a time changed the theoretical background of the models, and the addition of extra shops changed the built environment, these features did not change over the course of a model run. In other words, once a model run had started, the environmental conditions remained constant throughout the rest of the simulation. This is similar to other agent-based models that did not have input variables of this kind, such as the SLUDGE (Simulated Lane Use Dependent on eDGe Effect externalities) model described by Polhill et al. (2008). While it was noted in Section 5.2.5 ('Initialisation') that the road network varied slightly each time the models were initialised, this was more a quirk of the model design, rather than something that naturally changed or was designed as a variable environmental input.

5.2.7. Details – submodels

There were four submodels covered in the 'process overview and scheduling' section of the 'overview' domain, and these will be described in more detail in this section. Firstly, agent movement within the models will be described, including the probabilities associated with movement and the testing of the method for implementing movement within the models. Grimm et al. (2006) have stated that the 'submodels' section of the ODD protocol should cover how models were tested and calibrated, so the verification of agent movement will also be covered. Secondly, agent interactions will be detailed, including a summary of each interaction between an agent and its environment. Thirdly, the updating of agent characteristics will be detailed, including how agent characteristics could change, and how the theoretical pathways were incorporated into the models to do this over the course of the simulations. The verification and calibration of these pathways will also be covered. Finally, the time lag events that affect agents will be described, with an explanation of the different time lags involved.

Agent movement within the models

Movement in the agent-based models was based on a previous NetLogo model of student movement at George Mason University (Zhou, 2016 – see Section 5.2.2). This process began with agents selecting a random destination at the start of the simulation and calculating the shortest path to this destination. The agents then navigated the nodes and links, which were created based on the road and path data obtained from the university, in order to reach their chosen destination. The model used the A-star path finding algorithm to define this movement, which was developed by Hart et al. (1968), building on earlier work from Edsger Dijkstra (1959). The algorithm works by searching through all possible paths in the model that lead to the desired destination, aiming to find the path that results in the smallest 'cost'. This could be the least amount of time taken to reach the destination, or the shortest distance covered to reach a destination. In this case the model was based around finding the shortest distance, rather than saving time. Within Zhou's example NetLogo model, once an agent had reached its destination it stayed there for one step (or 'tick') in the model, before moving onto the next randomly selected destination. This was deemed an appropriate approach for this research, as it would allow agents to move between and visit different destinations based on a road network or grid. As highlighted in Section 5.2.2, extra geographical data were included (with edits made to the model's code) so that agents could tell destinations apart, and aim for a destination different to the one they were previously at.

Due to the issues associated with the initialisation of road networks in the GMU model (see Section 5.2.5), a second GIS based NetLogo example was also considered. The Venice model (Densmore, 2010) was created as part of a collaborative project with the Venice City Water Traffic Department, and used nodes and links to recreate the waterways of Venice, while agents in the form of boats navigated the city's canal network (see Figure 48). Figure 49 offers a 'birds-eye view' of the model using NetLogo's 3D viewing feature.

Figure 48 – Areal contextual view of the Venice model (Densmore, 2010)

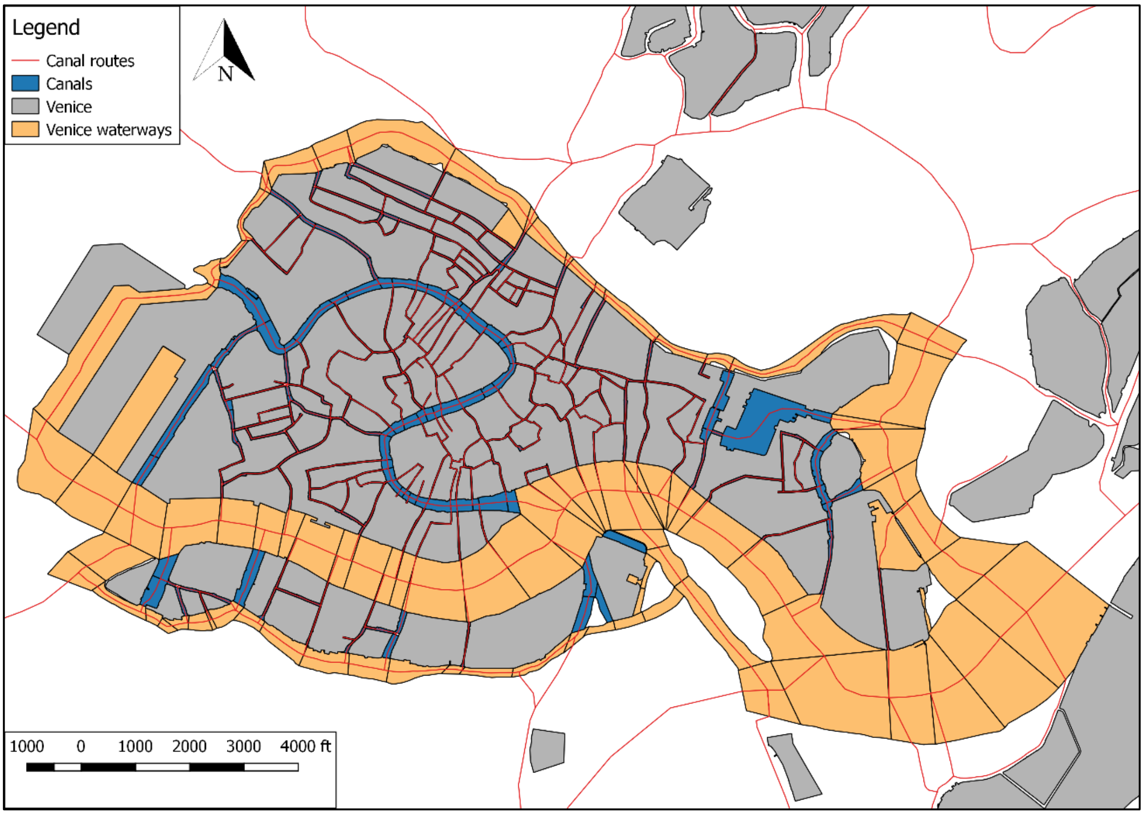
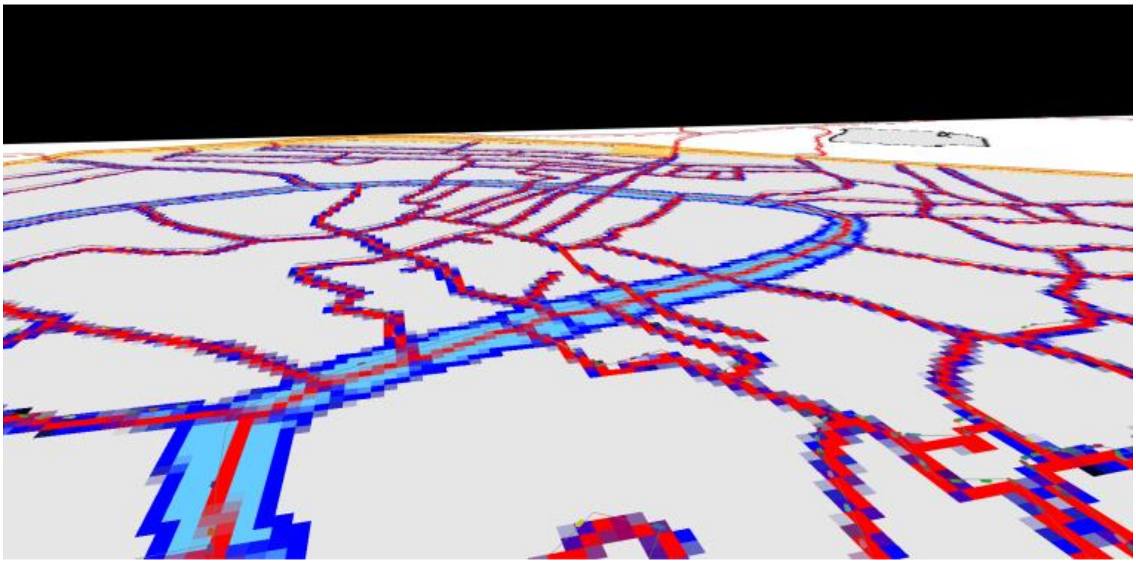


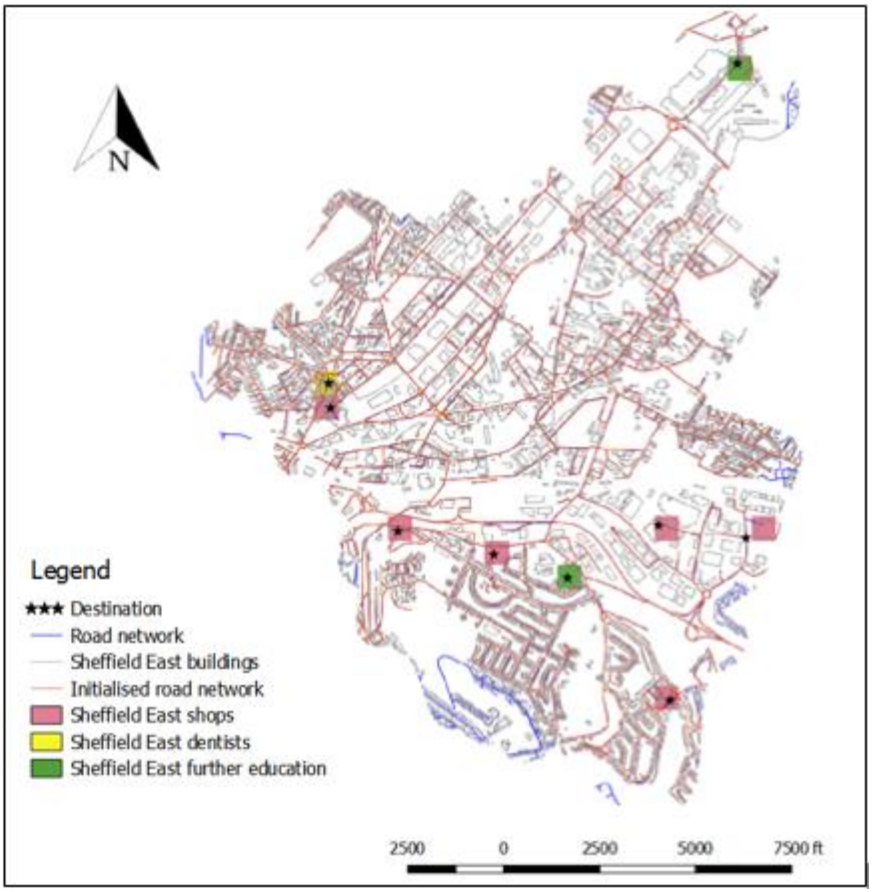
Figure 49 – Closer view of the Venice model in NetLogo's 3D viewer (Densmore, 2010)



This was similar to the GMU model in that it used links to allow agents to traverse the city's waterways, however the advantage of the Venice model was that the links and nodes represented the exact positioning of the canal network, and initialised in the same way each time. This allowed for spatial explicitness to be included within the simulation. One downside to the model was that (at least in the prototype model available online - http://backspaces.net/wiki/NetLogo_Bag_of_Tricks) the agents navigated random routes through the canals, with no obvious goal or destination. This is acknowledged by the model's creators.

A test simulation combining elements of the two models was therefore analysed to assess its usefulness. The path finding capabilities of the GMU model were combined with the methods for creating the canals in the Venice model, creating a geographically explicit model allowing agents to move between destinations. However, several downsides were noted. Firstly, some roads in the study areas were 'cut-off' from the simulation as they were not connected to the main road network. This was due to the extent of the LSOA boundaries, which are designed to capture population rather than physical features. This can be seen in Figure 50, which visually represents the Sheffield East study area after the aforementioned changes. A further increase to the buffers used to clip the road network was considered to counter this, however this was not pursued as it would likely have led to at least a small sample of the roads being 'cut-off' again in a different location. Beyond modelling the whole of the UK it would be hard to create any simulation that would not cut roads off at some point.

Figure 50 – New layout of the roads (orange) in the Sheffield East model, with some (blue) cut off



A second concern was the length of time this new simulation setup took to run. Due to nodes and links being set up in the exact positioning of the roads there was an increase in the number of these that agents had to traverse, which led to the model being far slower than either the GMU or Venice models individually. Some took a week or more to run for a relatively low number of ticks, compared to equivalent models using the GMU set up, which finished in less than a day. A third issue concerned agent movement within the models. Both models were tested to check the frequencies of visits to certain facilities compared with pre-defined parameters (this will be explained in greater detail on pages 187 and 189), for which the GMU model showed a greater consistency than the combined model. Finally, the number of visits to certain facilities was far lower in the combined model than in the GMU model, which meant that an already more computationally intense model would need to run for even longer to mimic the same time period as the GMU model. Given these concerns over speed and accuracy, it was decided that the GMU setup procedures would be used for the final simulations. Further testing of the GMU model showed that increasing the size of the model made the links and nodes more geographically accurate. This can be seen in Figures 51 and 52.

Figure 51 – The 16x16 grid GMU model for Sheffield East

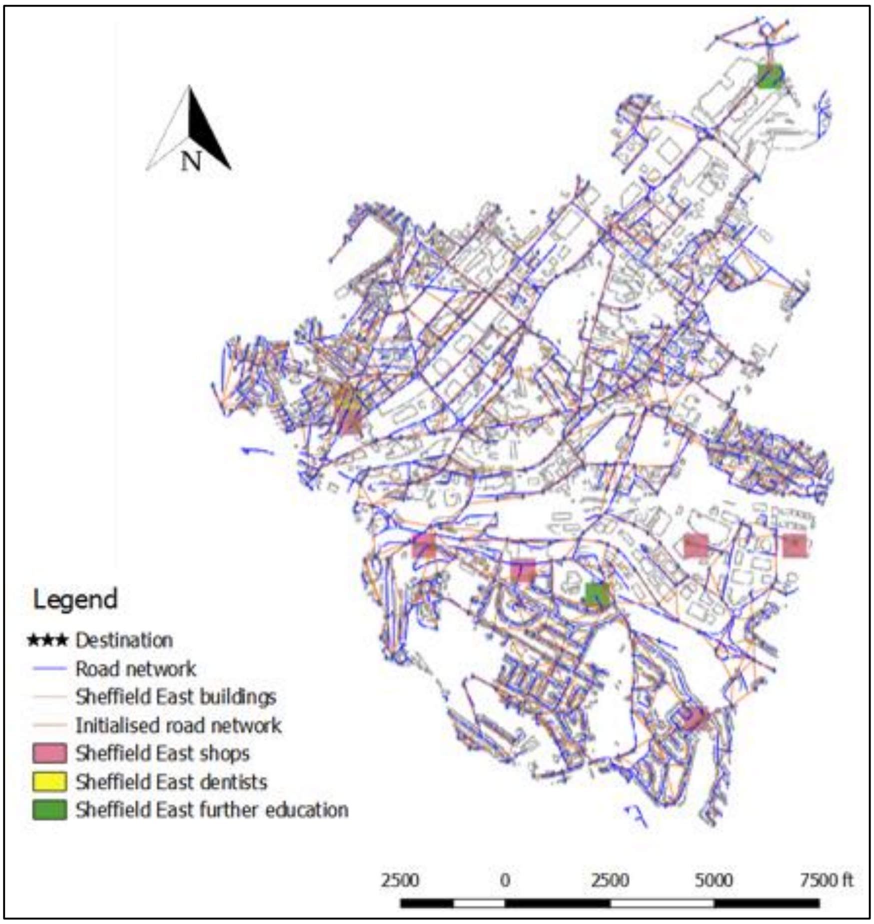
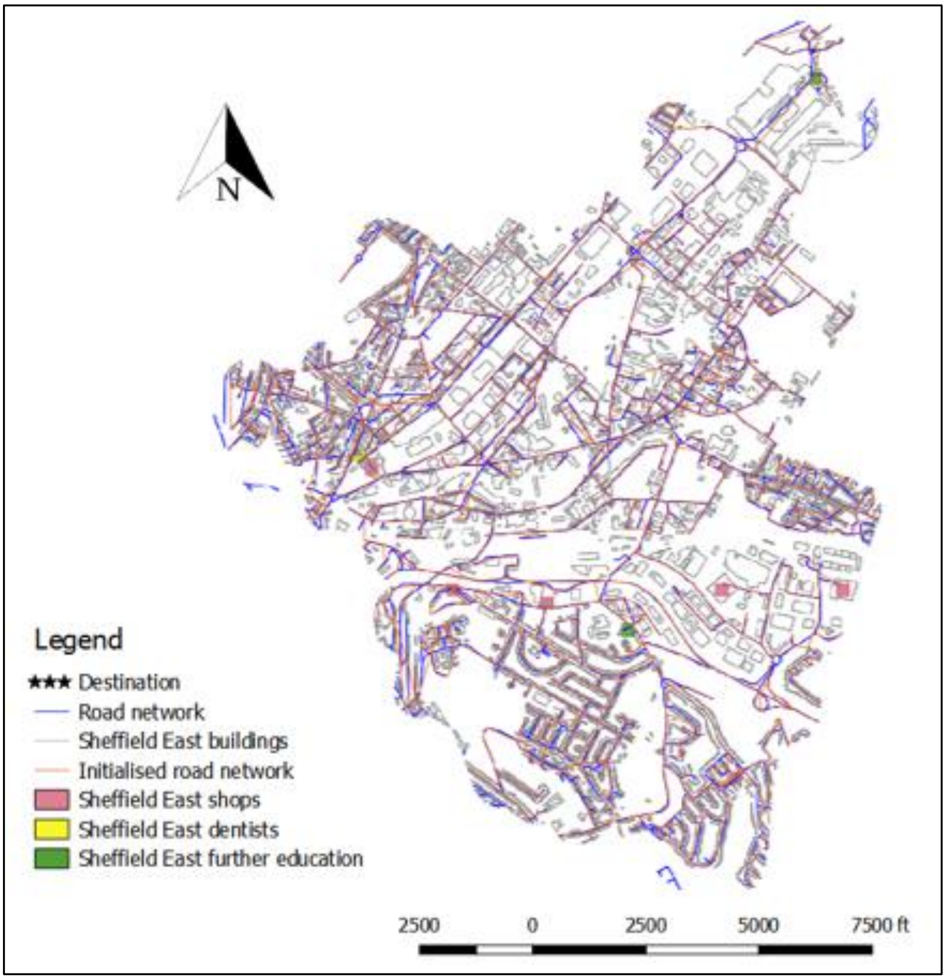


Figure 52 – The 32x32 grid GMU model for Sheffield East



This is perhaps logical as more patches allow for a finer level of detail to be included in the model; however, this again led to significant reductions in speed when running the model, so the 16x16 grid was used for the final simulations. This demonstrates another example of the types of compromise that exist in this type of research, which essentially represented a trade-off between speed and geographical accuracy. The combination of the Venice and GMU models would have been far too computationally intensive, and taken far too long for this research, and while increasing the grid size of the final model allowed for a more geographically accurate environment than the original GMU model, this again led to difficulties in running the simulation. Ideally future research should look to explore the possibility of geographically explicit models, and making these computationally viable. The 16x16 model still covered all of the main transport routes in each of the study areas, and agents were still able to traverse these to get to the correct locations, so the model was deemed fit for use.

Having finalised the method of movement, agents were set upon an initial course at the start of the model to a nearby facility based on probabilities. These probabilities differed based on the characteristics of the individuals, and are listed below in Table 21. The conceptual idea behind these probabilities was that different groups in society have different patterns of visiting the destinations in the model, and these probabilities represent a crude way of differentiating these behaviours. For example, attendance is known to vary by socio-economic status (Eckersley and Blinkhorn, 2001), which in this research was represented by the NS-SEC variable.

However, the use of the NS-SEC variable caused concern given the structure of the data, which is not intended to be hierarchical in nature (ONS, 2010), a rule which relates to both the 8 group and 5 group classifications. The 3 group classification could be considered hierarchical, however this is still conceptually problematic and less than ideal (ONS, 2010). Therefore, the NS-SEC data was not used as a way to differentiate agent behaviours, with education data used as a proxy for this instead. Education is also a theoretically relevant variable, and has been shown to be related to oral health related behaviours by a number of previous investigations (Chu et al, 1999; Williams et al, 2002; Singh et al, 2013). Riley et al. (2006) found positive associations between education and dental attendance, so it was therefore assumed to be a suitable proxy for driving the visiting probabilities of the agents in this research,.

Table 21 - Probability of movement to destinations, based on educational attainment (Sheffield East)

Agent characteristic	Action	Probability of action
Has degree	Movement to dentist	20%
Has degree	Movement to shop	60%
Has degree	Movement to FE	20%
No degree	Movement to dentist	10%
No degree	Movement to shop	80%
No degree	Movement to FE	10%

Probabilities have been used to drive the dynamics of agent-based models on numerous occasions (Grimm et al, 2006; Tracy et al, 2014; Olivella-Rosell et al, 2015), and are often informed by mathematical formulas based on known characteristics of agents or their environments. However, defining probabilities is a difficult task. None of the previous oral health focused agent-based models used probabilities for dental visits in relation to other neighbourhood based features, so there was no previous work on which to base such probabilities for this research. From studying the literature, it became clear that attendance patterns varied by deprivation, however this was not a quantifiable effect, either individually or relative to visiting patterns of other locations. There is also a lack of official data with which to compare attendance rates between different locations. The United Kingdom Time Use Survey (Gershuny and Sullivan, 2017) was investigated for its potential use, however this contained no specific data on visiting patterns to dentists.

The probabilities in Table 21 therefore represent an attempt to quantify, as logically as possible, the visiting patterns for both groups, based on education acting as the differentiating variable. Conceptually shops were assumed to be the most likely buildings to be visited, given that shopping likely occurs more often than visits to dentists and further education centres. Therefore, both education groups had their highest visiting probabilities for these locations. Dentists and FE locations were then given equal weighting in both areas as they were presumed to be less frequently attended. Based on the literature however, those with higher levels of education were given a higher probability of attending the dentist and FE locations. It is important to remember that this was exploratory research, and due to the lack of guidance in previous literature and official data these experimental probabilities represented a first attempt to quantify such interactions, and as such were deemed appropriate for use in this analysis. These probabilities could be changed in future research.

The probabilities were implemented using the ‘random-float’ feature in NetLogo. This function produces a random number between zero and a defined ceiling, which in this case was set to 1. For agents with a degree living in the Sheffield East study area, if the random number fell between 0 and 0.2 then the agent would head to the nearest dentist. If the number generated by the random float function was over 0.2 and below 0.8 they would head to the nearest shop, and if the number was equal to or over 0.8 the individual would head to the nearest FE facility. Once an agent had arrived their next location would be chosen based on the probabilities of attending one of the other two location types. In other words, if an agent went to the dentist, their next destination would be either a shop or FE facility (although FE facilities were only present in the Sheffield East study area). This was an attempt to mimic real life behaviour, by making the assumption that it is highly unlikely that individuals would visit the same type of facility twice in a row.

As there were no educational facilities present in the Sheffield West study area, having instructions to send agents to such facilities would be meaningless if the locations did not exist. For the Sheffield West study area, the probabilities were therefore edited from those in Table 21 to reflect the fact that only shops and dentists were present in this area. These probabilities can be seen in Table 22 below. In order to maintain the educational difference between the groups, the probability of attending a dentist was changed to 25% for those with a degree, and 15% for those without.

Table 22 - Probability of movement to destinations, based on educational attainment (Sheffield West)

Agent characteristic	Action	Probability of action
Has degree	Movement to dentist	25%
Has degree	Movement to shops	75%
No degree	Movement to dentist	15%
No degree	Movement to shops	85%

The probabilistic nature of the model therefore meant that an element of randomness was included, which was important when trying to model reactive agents, where it would perhaps be a mistake to assume that expected behaviours would occur on every single occasion. While agents in theory should attend a given facility a certain number of times, this could deviate slightly, leading to changes in agent characteristics. As it was not known if such changes would affect the outcome of the models, each simulation was run 10 times, and an average score taken from these. Taking the average of multiple runs has been used before in geographical agent-based modelling (Crooks, 2008), although 50 runs has been identified as an appropriate number based on previous research (Malleson et al, 2010). The computational intensity and timescales involved meant that this was not possible when running as many models as were needed to test the theoretical pathways in this research.

Before running the simulations, the agent movement functions needed to be tested, or verified. Verification forms an important first stage in the process of evaluating an agent-based model. Castle and Crooks (2006) suggest this process should involve 'checking the model behaves as expected' (P.35), or inner validity (Brown, 2005). Xiang et al. (2005) add that verification involves checking that the 'programming implementation of the abstract or conceptual model is correct' (p.47), and involves steps such as debugging code, identifying incorrect theoretical model implementation, and verifying calculations. Agent-based models can be very complicated in nature, often with many interacting sections of code running simultaneously. Therefore, it was vitally important to thoroughly check each element of the model to make sure that it was acting as it should. Although verification is sometimes confused with validation, it is a separate step, as validation involves analysing the extent to which a model represents the system being studied (Castle and Crooks, 2006). For the agent-based models of the two Sheffield study areas the process of verification involved separating the model out into its more distinct parts, and testing these separately, so that only the process of interest could be studied. It was important to confirm that individual processes were working correctly, as these would combine with the other processes to contribute to the emergent phenomena of the model, and its overall outcome.

The aim when verifying the code associated with agent movement was to make sure that, roughly speaking, the different groups were visiting the three types of physical locations in the model in proportion to their expected probabilities. This verification test took place in a simplified version of the Sheffield East model (no theory code was included), due to the lack of FE locations in the Sheffield West study area. Monitors were included within the NetLogo GUI to count the number, and type of agents arriving at each destination, with the model running for 100 ticks with 100 agents. The results of this are displayed below in Table 23.

Table 23 - Expected versus actual visits to destinations, by educational attainment

Variable	Expected visits (%)	Actual visits in test simulation (%)
Dentist visits – degree (Got_to_dentist_D)	20	20.4
Dentist visits – other qual (Got_to_dentist_O)	10	8.8
Shop visits – degree (Got_to_shop_D)	60	57.4
Shop visits – other qual (Got_to_shop_O)	80	81.1
FE visits – degree (Got_to_FE_D)	20	22.2
FE visits – other qual (Got_to_FE_O)	10	10.1

As can be seen from Table 23, the percentage of individuals from each educational group matches roughly (given the potential for variation within the model) what would be expected based on the allocated probabilities. As such this section of code was deemed fit for use within the final model.

Agent interactions with the environment

As mentioned in Section 5.2.2, the locations of shops, dentists and FE facilities were marked by colour coded patches. This served a practical purpose as it allowed instructions to be given to agents based on the colour of the patches they were on at a given time. For example, if an agent crossed a pink patch, instructions based on behaviours in a shop were carried out by the agent. Within NetLogo the 'gis:intersects?' function allows users to define certain patches depending on whether or not they intercept features within shapefiles. This was considered as an alternative to allow agents to identify different building types, however due to the size of the patches this function often led to much larger areas around the destinations being incorrectly assigned as buildings, so was not used.

Within this research the agent interactions were based on the idea of thresholds, or the idea that an agent's behaviour, or the effects on an agent from a certain interaction, may differ depending on whether the score for certain agent variables were above or below predefined thresholds. As models progress an agent's score may change, pushing them to the other side of this threshold, hence changing their behaviour or outcome. Such approaches can be advantageous to agent-based models when associated behaviours are well known and documented (Heppenstall et al, 2016), with Wang et al. (2016) using thresholds based on the oral health status of agents to determine referral patterns, as well as to designate more influential members of peer groups. Thresholds have also been used in agent-based models of land use, such as aspirational break-even thresholds for activities on parcels of land (Matthews et al, 2007).

As mentioned in the previous section on agent movement, there were concerns over the use of the NS-SEC variable due to its non-hierarchical nature. Some agent interactions used social hierarchy position (represented by the NS-SEC classification) as a differentiating concept, so use of this variable was therefore necessary. To avoid violating the data structure of this variable one group (group 8 – 'long term unemployed and never worked') was chosen to act as the reference category throughout the research, while the other seven groups felt the opposite effect to this class. While this was a slightly crude assumption, it avoided violating the conceptual structure of the data, and matches the dichotomous way that employment and unemployment have been studied in previous oral health related research (Roberts-Thomson and Stewart, 2008; Costa et al, 2012). A full list of the agent interactions present in this research is provided below.

- Agent-dentist interactions: Upon encountering a patch containing a dentist, agents either derived benefit or saw a negative effect on their health depending on their individual characteristics. For example, agents who had the characteristic of attending the dentists every six months found that their dental knowledge increased (through increased fluoride intake), which could have a positive effect on their teeth if this pattern continued. Conversely, those with less regular attendance patterns saw their dental knowledge decrease, which in the long term could adversely affect their tooth decay score.

- Agent-shop interactions: Once an agent had arrived on a patch containing a shop, if their NS-SEC score was equal to 8 their sugar intake increased. In the long term this would mean they were more likely to experience adverse effects of sugar intake through their tooth decay score. Conversely, those with an NS-SEC score between 1 and 7 saw their sugar intake decrease, which could decrease their tooth decay score over time.

- Agent-educational facility interactions: Local educational facilities were included that provided apprenticeships, further training or further education for adults. Adults without educational qualifications and aged 16-24 were attracted to these facilities, which were only present in the Sheffield East simulation. Attendance at these locations increased the individual's dental knowledge (a proxy for dental education), which in the long term could have a positive effect on the individuals' tooth decay variable.

- Agent-material circumstances interactions: Material circumstances were represented by background data in the model (and subsequently within the agents as well), in the form of median house price statistics (ONS, 2015). Agents assessed this data at various time points throughout the model, and depending on the value associated with their area, would experience a different effect. The mean of the house price data for each of the two areas was used as a threshold to differentiate between individuals. House price values were checked, and if the house price for a given agent's MSOA was above the mean threshold they derived a positive effect, with adverse effects for those whose value fell below the mean threshold. For example, if the house price for an area was above the mean threshold the corresponding agents would see their psychological stress variable reduce, and vice versa.

- Agent-employment interactions: Similar to the 'agent-material circumstances' interaction, employment data were stored as background and individual data within the model, this time in the form of model based estimates of income for MSOAs (ONS, 2009). Those with a value above the mean threshold would derive benefit, while those with a score below it experienced a negative effect. For example, MSOAs with average income above the mean threshold would see the tooth decay score of corresponding agents decrease, and vice versa.

- Agent-health behaviours interactions: As with the previous two interactions, the data for health behaviours were stored as background data in the model as well as within the individual agents, and checked periodically. Health behaviour data were represented by the average years of lost life per LSOA (IMD, 2015a). Those with a score above the mean threshold derived a positive effect, while those with a score below it experienced a negative effect. In this case LSOAs with a higher value for years of lost life could see the fluoride intake scores of corresponding agents decrease, and vice versa.

- Agent-social capital interactions: Social capital was also included as a background variable and an individual characteristic, and checked periodically. The mean value again acted as the effect threshold, with those scoring above this threshold deriving a positive effect, and those scoring below the threshold experiencing a negative effect. The data for social capital were represented by data from the 2015 Indices of Multiple Deprivation crime domain for each LSOA (IMD, 2015b). For example, this could entail an increase in psychological stress in agents from LSOAs with higher crime scores, and vice versa.

- Agent-road interactions: Agents within the model interacted with the road network as a way of moving around the study areas. As described in the previous section ('Agent movement within the models') this was accomplished by turning the road network into a series of links and nodes that the agents could travel on and between. While it was not possible to include all of the pedestrian walkways in the city, following the road network at least allowed agents to travel to and from locales in a logical manner, most of which were accessible by road (or pavements adjacent to these).

Updating agent characteristics

Both individual characteristics and tooth decay scores increased or decreased by 0.01 (depending on the characteristics of the individual) upon the implementation of the theoretical pathway rules. This allowed for an agent to cross at least one threshold during the 730 ticks of the models runtime with all of their associated variables, although there was no guarantee of this happening given the number of processes occurring at the same time (i.e. an increase in one variable could be cancelled out by a decrease due to another section of the code). The subtraction as well as the addition to the tooth decay scores was designed as an attempt to keep the tooth decay variable more balanced in the populations, so that it did not increase too much.

Within the models there needed to be some way of increasing or decreasing the tooth decay variable over time, as it would seem less appropriate to simply reach a certain point in the model and have an agent's tooth decay score change by +/- 1. A seemingly more realistic approach would be the build-up of conditions in the agents over time that lead to their tooth decay scores increasing or decreasing. According to the life course perspective, specifically the accumulation model, some health related processes take longer to occur due to exposures over the life course gradually accumulating, which can cause long term damage to health (Sisson, 2007). As mentioned before, previous literature would seem to support this view, with Mitchell et al. (2002) citing the longer term effects of unemployment on health, while specific examples of the longitudinal effects of earlier life conditions have been demonstrated in relation to tooth decay (Poulton et al, 2002; Thomson et al, 2004). Despite some authors arguing that mortality and morbidity rates can change suddenly (Mackenbach, 2012), the accumulation model approach was deemed appropriate to base the simulations on. While the critical periods life course model is equally valid, this would be harder to implement programmatically, as it is harder to conceptually define when these periods would be.

The changing of tooth decay scores and other agent characteristics worked as follows. Table 18 showed that individuals have a 'sugar' variable, containing information on sugar intake. If testing Pathway 3 (see Section 3.5), the sugar variable could be updated after an agent had visited a shop, increasing or decreasing the variable to represent higher or lower intake through their shopping choices. This could then be used as a reference point later in the simulation when judging the health damaging behaviours of agents. The background, or neighbourhood based data presented in Table 19 were utilised in a similar way. The 'material circumstances' variable for example could be used to influence the decision of an agent in a given area. Depending on the material circumstances of an agent's neighbourhood, variables such as an individual's 'purchasing power' could be influenced in the model, which could have knock on effects on their shopping habits, represented by their sugar intake later in the pathway.

The continual reference to, and updating of individual characteristics within the model allowed it to be truly dynamic in nature over time. In line with this, individual characteristics were stored as integers rather than factors, to allow addition and subtraction to occur.

In order for agent characteristics to be updated, the theoretical rules of the model needed to be defined. This presented an issue with regard to implementing theory in an applied simulation model. Theoretical concepts cannot simply be transferred into an agent-based model, and there needed to be an intervening stage for this to occur. For example, conceptual ideas need some form of hard data in order to be operationalised and represented in quantitative or simulation modelling. This was covered in Chapter 3, where the theoretical concepts from the pathways were operationalised, with these data acting as proxies for the theoretical concepts within the agent-based models.

A further issue concerns the overall structure of the pathways. The visualisations in Chapter 3 (see Figures 3-7) show these pathways are interlinked and at times complex in nature. This would be complicated to incorporate into a simulation model in its entirety, so an alternative approach was taken. Each of the features of the framework were broken down further into individual pathways, established through manually checking the original theoretical structures, and creating new pathways whenever a different route to inequalities in tooth decay was identified. This approach simplified the theoretical models and allowed them to be included in the simulations, while also maintaining all of the theorised pathways by which neighbourhoods may influence tooth decay. Despite the individual pathways not being in the same collective structure that was shown in Chapter 3, all of the pathways could still be run at the same time within the simulation, allowing for the dynamic interactions of the model to take place, and for multiple variables to be influenced at once by others.

Table 24 shows each of the individual pathways taken from Figures 3-7. Pathways 1 (Figure 3) and 5 (Figure 7) were not included. Pathway 1 (‘physical features of the environment shared by all residents in a locality’) only contained water fluoridation, and given that neither of the study areas had either artificially or naturally fluoridated water supplies it was seen as unnecessary to model this. This is, however, one of the many important issues within dentistry that agent-based modelling can assist in analysing, so future research should look to include this where appropriate. Pathway 5 (‘the reputation of an area’) was so closely related to features of other domains (particularly

Pathway 3) that reproducing this a second time was seen as unnecessary, and could in fact have influenced the model in a negative way, if an effect was being applied to an agent multiple times when it should not have been.

In some cases, the pathways were further simplified. For example, in Pathway 2 the ‘nutrition’ variable was removed, with only the sugar variable remaining. This was due to data for the nutrition variable only being available from the Understanding Society survey, and due to difficulties incorporating data from a different sample alongside the ADHS. This would have been problematic with regards to matching the population demographics and socio-economic variables from one survey with those of another, to make sure that the data being incorporated into the ADHS sample from other surveys were from the same types of people. Thus only data from the ADHS was included. Attempting to combine survey data from the ADHS and other surveys (most likely Understanding Society) was beyond the scope of this research, and represents a limitation with such an approach, as it limits the pool of data that can be used. Techniques such as data linkage (Slack-Smith, 2012) and statistical matching (Rodgers, 1984; Moriarity and Scheuren, 2001) may provide a solution to this, and in future research it may be beneficial to combine multiple surveys, potentially allowing for the inclusion of a richer set of variables. Nutrition was also removed from Pathway 3, although ‘health damaging behaviours’ in the form of diet was still included, so this concept was not completely dropped from the model. Finally, nutrition was also removed from Pathway 4, but again (similar to Pathway 2) the ‘sugar’ variable remained.

In some cases, theoretical constructs with the same name were represented in different ways within the pathways. This was in order to differentiate the processes occurring in different pathways. Education was a good example of this, as it was represented as an individual variable in one pathway looking at social hierarchy position, while it was represented as point data to represent the physical location of education centres in another pathway. Other examples of this included diet, which was represented by sugar intake in Pathways 2.2 (material circumstances -> financial constraints -> diet/sugar intake -> tooth decay) and 3.4 (shop -> diet/sugar intake -> damaging behaviours -> tooth decay), while it was represented by cake consumption in Pathway 4.1 (Health behaviours -> diet-> sugar/nutrition -> tooth decay). This was due to the conceptual definitions of diet associated with each pathway, and also because Pathways 2.2 and 3.4

were simplified due to data from Understanding Society being removed. In Pathway 4.1 sugar intake was a result of cake consumption, whereas in Pathways 2.2 and 3.4 it represented a lack of affordable nutritional food.

On occasions certain variables were combined in one section of the code. For example, the second domain contained material circumstances ('h_price') and financial constraints ('costdly'). These are very closely linked conceptually, so these variables were considered together when influencing the effect on fluoride levels (in Pathway 2.3 - material circumstances -> financial constraints -> (dental) knowledge -> health habits). This can be seen in the third line of code in Figure 53 below. The combining of two variables also had the effect on taking an additional step out of the code, which was helpful when debugging. This approach has been used before in such modelling, with Grimm et al. (2006) demonstrating how the winter morality rates of marmots can be based on a combination of age, winter strength, and number of young in the group. The idea throughout this research was to treat variables and data with respect with regard to using them in an appropriate way, rather than simply plugging them in arbitrarily, or editing numbers in a way which was contrary to the scale and design of the data source.

Figure 53 – Example of two variables being combined in the theory code

```
to d2p3
    if ticks mod 50 = 0[
      ifelse h_price < 88564 and costdly = 1
      [set fluoride fluoride + 0.01]
      [set fluoride fluoride - 0.01]
    ]
    d2p3-effect
end
```

Table 24 outlines the individual pathways, how the theoretical concepts for the pathways were represented in the model, as well as the parameter values that were applied to these. In some cases, the parameters applied equally (according to theory) across the two study areas. However, in some cases parameters differed, usually in the case of the neighbourhood based variables, and in these cases the parameters for the Sheffield East and West study areas are stated separately. This is because these background variables are likely to differ depending on the context of the areas. The divide between the east and west of Sheffield is well known (Thomas et al, 2009), so using an average house price from the Sheffield East study area would likely be too low for the context of the Sheffield West study area, and vice versa. On such occasions averages were taken in each area to represent the local context.

In order to assist with reading Table 24, the parameter scores have been standardised so that they represent the end of the scale which leads to negative effects on the agents. Thus it can be assumed that scores above or below those in the table (depending on the data direction for each parameter) will have the opposite effect. The final variable for each pathway is the one that will have the final effect on the tooth decay score. Where the 'effect on model' column has been extended to cover more than one variable (as in the case of Pathways 2.1-2.3), this represents occasions where multiple variables were used at the same time in the code, much like the example shown in Figure 53.

Table 24 - Theoretical pathways, concepts, indicators and parameters

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
2.1: Material circumstances -> financial constraints -> stress -> smoking -> tooth decay	Material circumstances	House price data (MSOAs)	Sheffield East: <= 88,654 Sheffield West: <= 251,636	Stress variable increased
	Financial constraints	Whether delayed dental treatment due to cost	1	
	Stress	Psychological discomfort	> 3	Smoking variable increased
	Smoking	Ever been given advice on giving up smoking	< 2	Tooth decay variable increased

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
2.2: Material circumstances -> financial constraints -> diet/sugar intake -> tooth decay	Material circumstances	House price data (MSOAs)	Sheffield East: <= 88,654 Sheffield West: <= 251,636	Purchasing power variable decreased
	Financial constraints	Whether delayed dental treatment due to cost	1	
	Purchasing power	Whether cost affected type of dental care/treatment	< 2	Diet variable increased
	Diet	Number of cakes eaten per week	< 3	Sugar variable increased
	Sugar intake	High sugar intake	< 2	Tooth decay variable increased
2.3: Material circumstances -> financial constraints -> (dental) knowledge -> health habits -> tooth decay	Material circumstances	House price data (MSOAs)	Sheffield East: <= 88,654 Sheffield West: <= 251,636	Dental knowledge variable decreased

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
	Financial constraints	Whether delayed dental treatment due to cost	1	
	Dental knowledge	Fluoride level	> 2	Healthy habits variable decreased
	Healthy habits	General dental attendance	1	Tooth decay increased
3.1: Employment -> social hierarchy position -> tooth decay	Employment	Model based income estimates (MSOAs)	Sheffield East: <= 491.8 Sheffield West: <= 1056.1	Tooth decay variable increased
	Social hierarchy position	NS-SEC classification	8 vs < 8	
3.2: Education -> social hierarchy positions -> tooth decay	Education	Highest qualification above or below degree level	Above vs below degree level	Tooth decay variable increased
3.3: Education -> dental knowledge -> damaging	Education	Location of further education	N/A	Dental knowledge

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
behaviours -> tooth decay		providers in Sheffield		variable decreased
	Dental knowledge	Fluoride level	< 2	Health damaging behaviours variable increased
	Health damaging behaviours	Consumption of sweets	< 3	Tooth decay variable increased
3.4: Shop -> diet/sugar intake -> damaging behaviours -> tooth decay	Shop	Location of shops and supermarkets within Sheffield	N/A	Sugar intake variable increased
	Diet/sugar intake	High sugar intake	< 2	Health damaging behaviours variable increased
	Health damaging behaviours	Consumption of sweets	< 3	Tooth decay variable increased
3.5: Dental service usage -> associated	Dental attendance	How often individuals go to the dentist	1	Dental knowledge

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
benefits/knowledge -> tooth decay				variable decreased
	Dental knowledge	Fluoride level	< 2	Tooth decay variable increased
4.1: Health behaviours -> diet-> sugar/nutrition -> tooth decay	Health behaviours	IMD health domain - Years of potential lost life (LSOAs)	Sheffield East: > 75.09 Sheffield West: > 52.94	Diet variable increased
	Diet	Number of cakes eaten per week	< 3	Sugar intake variable increased
	Sugar/nutrition	High sugar intake	< 2	Tooth decay variable increased
4.2: Health behaviours -> oral health habits -> tooth decay	Health behaviours	IMD health domain - Years of potential lost life (LSOAs)	Sheffield East: > 75.09 Sheffield West: > 52.94	Oral health habits variable decreased

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
	Oral health habits	Tooth brushing habits	> 1	Tooth decay variable increased
4.3: Health behaviours -> attendance -> knowledge -> tooth decay	Health behaviours	IMD health domain - Years of potential lost life (LSOAs)	Sheffield East: > 75.09 Sheffield West: > 52.94	Dental attendance variable decreased
	Dental attendance	How often individuals go to the dentist	1	Dental knowledge variable decreased
	Dental knowledge	Fluoride level	< 2	Oral health habits variable decreased
	Oral health habits	Tooth brushing habits	> 1	Tooth decay variable increased
4.4: Social capital -> acquired dental knowledge ->tooth decay	Social capital	IMD (2015) crime domain (LSOAs)	Sheffield East: > 0.24 Sheffield West: > -1	Dental knowledge variable decreased

Pathway	Theoretical concept	Variable in model	Parameter in model	Effect on agent
	Dental knowledge	Fluoride level	< 2	Tooth decay variable increased
4.5: Social capital -> Healthy behavioural norms -> tooth decay	Social capital	IMD (2015) crime domain (LSOAs)	Sheffield East: > 0.24 Sheffield West: > -1	Healthy behavioural norms variable decreased
	Healthy behavioural norms	General dental attendance	< 2	Tooth decay variable increased
4.6: Social capital -> stress -> smoking -> tooth decay	Social capital	IMD (2015) crime domain (LSOAs)	Sheffield East: > 0.24 Sheffield West: > -1	Stress variable increased
	Stress	Psychological discomfort	> 2	Smoking variable increased
	Smoking	Ever been given advice on giving up smoking	< 2	Tooth decay variable increased

Having created the individual pathways, the process of coding these into NetLogo was much simpler. Variables in the pathways could now be dealt with one at a time in a more straightforward manner using 'ifelse' statements. This essentially worked by saying ‘if A is X, do this to B, if A is Y, do this to B’. An example of this can be seen below in the example code for Pathway 4.6 (Figure 54: Social capital -> stress -> smoking). These functions apply an effect to an agent if they match a certain criterion, and a different effect to those that do not, thus making use of the thresholds shown in Table 24.

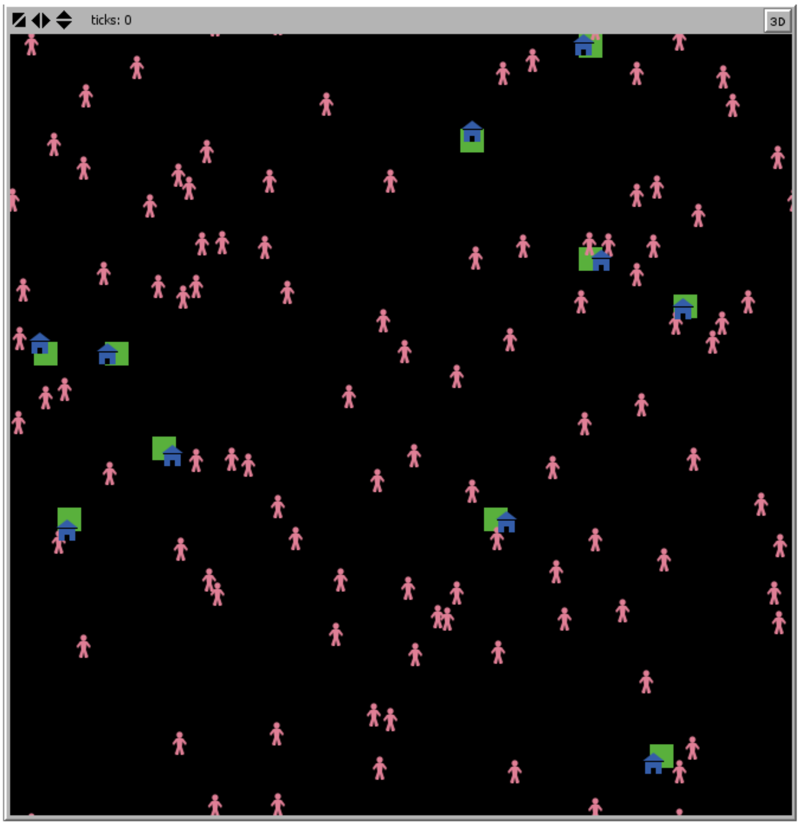
Figure 54 – Example of theory converted into code in NetLogo (Pathway 4.6)

```
;;;;; D4P6 - Social capital -> stress -> smoking -> OHI ;;;;;

to d4p6
    if ticks mod 50 = 0[
    ifelse crime_score > 0.24
    [ set psycdisc psycdisc + 0.01]
    [ set psycdisc psycdisc - 0.01]
    ]
    d4p6-smoking
  d4p6-numdu98
end

to d4p6-smoking
    if ticks mod 50 = 0[
      ifelse psycdisc > 2
      [set everadsm everadsm - 0.01]
      [set everadsm everadsm + 0.01]
    ]
end

to d4p6-numdu98
    if ticks mod 50 = 0[
      ifelse everadsm < 2
      [set numdu98 numdu98 + 0.01]
      [set numdu98 numdu98 - 0.01]
    ]
end
```

Again it is worth emphasising that while these pathways were set up individually, when run at the same time they were effectively interacting, as multiple pathways influenced the agents, and their characteristics, at the same time. This approach also had the advantage of being able to test combinations of pathways depending on the processes of interest. For example, if looking specifically at pathways concerned with sugar consumption, it would be possible to test Pathways 2.2 (material circumstances -> financial constraints -> diet/sugar intake), 3.4 (shop -> diet/sugar intake -> damaging behaviours) and 4.1 (health behaviours -> diet-> sugar/nutrition) separately to see which had the greatest effect by selecting these from the list of available pathways at the start of the simulation. Other pathways could then be included in an additive manner in the next iteration. This is a simplified example, but demonstrates the flexible approach that was available.

Through the approach outlined above it was possible to update agent characteristics using the rules from the theoretical pathways. However, the code representing the theory still needed verifying, and additionally, calibrating. For the verification, each pathway and its associated section of code was put into its own, non-spatial agent-based model to be run separately from the rest of the procedures. The aim was to test whether the various processes worked as would be expected. These non-spatial environments consisted of a 16x16 grid with a series of random patches set to different colours, and buildings on top of these patches where destinations were relevant to the code. In the verification models that did not involve an agent arriving at a physical location these coloured patches were set to have no effect when agents landed on them. Some of the pathways only included the lag effects of certain variables over time for example. Figure 55 shows an example of what these test environments looked like.

Figure 55 – Example NetLogo environment for model verification



Agents had simple movement instructions (e.g. forward 1, right 1), as the actual directions they took were not important in this context. These non-spatial environments allowed the agents to be monitored more easily, based on their simple and predictable movement as well as the lack of detail (e.g. the road network and buildings) to distract from following the agents. This process was similar to 'tracing', a verification process suggested by Xiang et al. (2005) which involves following entities within the model, and their subsequent behaviours, to ensure that the models logic is correct. While the authors state that this process can be quite computationally intensive, this was addressed through the use of the smaller, non-GIS based verification simulations.

The effects of the theoretical pathways were set to kick in every 50 ticks. This was an arbitrary number, but given the non-spatial and non-temporal nature of the verification models this was not an issue. All sections of the individual pathways were manually verified, and any corrections that were required noted. The results of the verification models for each of the pathways are presented in Tables 25-27. It is worth remembering that when a rule, and the effect of the rule is stated, the opposite effect will apply to individuals who do not meet that criteria. To give a hypothetical example, if tooth decay increased for individuals with below average income, it would therefore decrease for those with above average income.

Table 25 - Verification of Pathways 2.1 - 2.3

Pathway	Rules of pathway (in chronological order)	Corrections
2.1	Every 50 ticks, those in MSOAs with below average house prices, and who delayed treatment due to cost saw their stress score increase	None
	Every 50 ticks, those with a stress score over 2 saw their smoking score increase	None
	Every 50 ticks, those with smoking scores over 1 saw their tooth decay score increase	None
2.2	Every 50 ticks, those in areas with below average house prices, and who delayed treatment due to cost saw their score for delaying the type of treatment they sought increase	None
	Every 50 ticks, those with a score above 2 for delaying the type of treatment they sought saw their cake consumption score increase	None

	Every 50 ticks, those with a cake consumption score above 3 saw their sugar intake score increase	None
	Every 50 ticks, those with a sugar intake score below 2 saw their tooth decay score increase	None
2.3	Every 50 ticks, those in MSOAs with below average house prices, and who delayed treatment due to cost saw their dental knowledge score increase	None
	Those with a dental knowledge score above 2, and a healthy behavioural norms score of 1 saw their tooth decay score increase	None

None of the pathways in Table 25 required corrections. As previously mentioned, in Pathway 2.1 household price (representing material circumstances) and cost delaying treatment (representing financial constraints) were considered together, as the two constructs are conceptually very similar, and are likely to impact on individuals/households in a similar way. It therefore made little sense to add an extra section of code by splitting these. In Pathway 2.3 dental knowledge (represented by fluoride levels) was considered alongside the healthy behavioural norms variable (represented by regularity of attendance) as it was not deemed appropriate to adjust a person's regularity of attendance at the same intervals as other variables. An individual's attendance score could not really be increased or decreased throughout the model in the same way as other variables, as the model effects kicked in at time periods of far less than every six months. Therefore, such an increase or decrease would be fairly unrealistic.

Table 26 - Verification of Pathway 3.1 - 3.5

Pathway	Rules of pathway (in sequence)	Corrections
3.1	Every 50 ticks, those with an NS-SEC classification of 8 and below average income levels saw their tooth decay score increase	None
3.2	Every 50 ticks, those with a qualification that was below degree level saw their tooth decay score increase	None
3.3	When an agent who had a degree landed on a yellow patch (dentist) they saw their dental knowledge score increase	Extra 'ask' command removed due to effect on model
	Every 50 ticks, those with dental knowledge scores above 2 saw their sweet consumption score decrease	None
	Every 50 ticks, those with a sweet consumption score above 3 saw their tooth decay score increase	None
3.4	If a person with an NS-SEC classification of 8 landed on an pink patch (shop), their sweet consumption score would increase	None
	Every 50 ticks, if an individual had a sweet consumption score above 3, their sugar score would increase	None
	Every 50 ticks, if an individual had a sugar score over 5 their tooth decay score would increase	None
3.5	When landing on a yellow patch (dentist), if an individual attended the dentist every six months their dental knowledge score would increase	None
	Every 50 ticks, if fluoride levels were over 5 an individual's tooth decay score would increase	None

In Pathway 3.1 NS-SEC and income were considered together in a similar way to variables in Pathway 2.1. This was because NS-SEC and income are both indicators of social standing, and it was not considered worthwhile to have them as separate steps in the pathways. Pathway 3.3 threw up a problem that saw individual dental knowledge scores increase after every tick, which should not happen (it should have changed every 50 ticks). This meant debugging the code for this section, which found that certain command procedures (an ‘ask’ function, which gives instructions for the agents to act on) had been entered more times than were necessary. This was disrupting the code and causing it to apply certain effects more frequently than it should have. This was addressed and the model ran correctly. None of the other pathways required corrections.

Table 27 - Verification of Pathways 4.1 - 4.6

Pathway	Rules of pathway (in sequence)	Corrections
4.1	Every 50 ticks, if the neighbourhood score for years of lost life was above the average value, an individual’s cake consumption score would decrease	None
	Every 50 ticks, if an individual’s cake consumption score was below 3, their sugar consumption score would increase	None
	Every 50 ticks, if an individual’s sugar consumption score was above 2, their tooth decay score would increase	None
4.2	Every 50 ticks, if the neighbourhood score for years of lost life was above the average value, an individual's tooth brushing score would increase	None
	Every 50 ticks, if an individual’s tooth brushing score was over 1 their tooth decay score would increase	None
4.3	Every 50 ticks, if the neighbourhood score for years of lost life was above the average value, and an individual’s attendance score was over 2, their dental knowledge score would increase	None

Pathway	Rules of pathway (in sequence)	Corrections
	Every 50 ticks, if an individual's dental knowledge score was over 2 their toothbrusing score would increase	None
	Every 50 ticks, if an individual's tooth brushing score was over 2, their tooth decay score would increase	None
4.4	Every 50 ticks, if the neighbourhood crime score was above the mean value, corresponding individuals saw their fluoride score increase	None
	Every 50 ticks, if an individual's fluoride intake score was below 2, their tooth decay score would increase	None
4.5	Every 50 ticks, if the neighbourhood crime score was above the mean value, and an individual's attendance score was over 2, their tooth decay score would increase	None
4.6	Every 50 ticks, if the neighbourhood crime score was above the mean value, individual psychological discomfort scores would increase	None
	Every 50 ticks, if an individual's psychological discomfort score was above 2 their smoking variable would decrease	None
	Every 50 ticks, if an individual's smoking variable was below 2 their tooth decay score would increase	None

As with previous pathways, certain variables were combined for conceptual reasons. This was the case with Pathway 4.5 for example, where social capital and healthy behavioural norms were combined, as research has shown them to be closely related (Lida and Rozier, 2013). The tests for this set of pathways revealed that all of the code was acting as expected, and none required corrections.

Having verified each part of the model, the next step was to calibrate them. Manson et al. (2012) state that 'calibration is the adjustment of model parameters and specifications to fit certain theories or actual data' (p.130). This step is essential because while the code may do what is expected of it, this still does not mean that what it is doing is mimicking theory in an accurate way. In other words, the code works, but the parameters of the code need to be checked to make sure they accurately match the theory. The importance of this stage of the modelling is highlighted by Johnson and Groff (2014), who state that 'choices made about how to represent those parameter values and the associated condition action rules will affect the process and the outcomes observed' (p.514). It was therefore important that the calibration process was as accurate as possible.

This meant checking each of the parameters to make sure the values associated with these were sensible, and also matched the relevant data from the ADHS. For example, if the model was judging favourable patterns of attendance at dentists, then a value of 1 related to attendance every 6 months, whereas with each value above 1 attendance levels become less frequent. The values associated with the agents in the model needed to match the values in the Adult Dental Health Survey, or the processes taking place would not make sense, or be traceable. It is of course possible that if the model was run for long enough that the values associated with agents would increase or decrease to the extent that they were no longer on the original scale, however it would at least be evident which end of the scale their behaviour was associated with, and the direction the behaviour was heading in. This was mitigated to a certain degree by each process only adding or subtracting 0.01 from the agent's characteristics, which, while an arbitrary number, was also designed to mimic the long term nature of the development of tooth decay.

An example of parameter checking can be illustrated through the code used to test Pathway 2.1. Here agents with smoking scores of '>1' saw their tooth decay variable increase. However, this made little sense when compared to the survey data, where a score of '1' related to having coping mechanisms, while '2' and '3' represented 'no' and 'never smoked' respectively. Due to this, the threshold for the smoking score was set to '<2', in order to capture those who do smoke as a coping mechanism. People with a smoking score of '<2' subsequently saw their tooth decay score increase. Pathway 2.2 displayed a similar problem, where the sugar intake variable was originally set to '>2' for those experiencing an increase in their tooth decay score. In the ADHS data the high sugar variable was binary in nature, with '2' representing a 'not high' intake. Therefore, the threshold for this section of code was changed to '<2' so that individuals with a score of '1' (high intake) experienced increased tooth decay. Pathway 2.3 had no such issues. The re-calibration of Pathways 2.1-2.3 are summarised in Table 28 below.

Table 28 – Re-calibration of parameters for Pathways 2.1-2.3

Pathway	Variable	Original parameter	Re-calibrated parameter	Reason for re-calibration
2.1	Smoking	>1	<2	Incorrect data scale direction
2.2	Sugar intake	>2	<2	Incorrect data scale direction

The re-calibration of Pathway 3.1 led to the NS-SEC value changing due to the nature of the data, while Pathway 3.2 had no issues with data classification. The dichotomy of degree vs other qualification used for this pathway may have been a little crude, however this was due to the binary nature of the variable in the ADHS data. In the first section of Pathway 3.3 fluoride scores increased if education level was equal to '1' (degree or higher), however according to the ADHS data, increasing the score associated with fluoride consumption corresponded with a decreased intake of fluoride in real terms. Fluoride scores were subsequently changed to decrease in the model if an individual had an education score of '1'. This in turn meant the second part of the function changing, which involved increasing or decreasing an individual's sweet consumption score. The threshold for fluoride was therefore set to '<2' for those who would see their sweet consumption score increase, and vice versa. This meant that those with high dental knowledge (fluoride <2) saw their sweet score increase, which according to the ADHS data values meant lower sweet consumption. This example in particular highlights the confusing, and sometimes counter intuitive nature of aligning such data scales correctly.

Pathway 3.4 also included the NS-SEC variable, and this was dealt with in a similar way to Pathway 3.1 (i.e. group 8 vs groups 1-7). Other than this however Pathway 3.4 had no other parameters that needed re-calibrating. In the case of Pathway 3.5, as the ‘best’ behaviour for the ‘freqden’ variable was perceived to be a score of 1 (attendance at least every six months), individuals with this score saw their fluoride scores decrease to keep them closer to 1 (the highest levels of fluoride intake), while those who attended less regularly saw their scores increase, indicating a lower intake of fluoride. This also meant that the numbers associated with the second part of the function needed to change. Firstly, the threshold of ‘>5’ set to test the model for the fluoride variable was far too high, as the scale for fluoride data in the ADHS only went up to 4. This was adjusted so that those with a fluoride score of ‘<2’ would see their tooth decay scores decrease (and vice versa). The re-calibration of Pathways 3.1-3.5 are summarised below in Table 29.

Table 29 - Re-calibration of parameters for Pathways 3.1-3.5

Pathway	Variable	Original parameter	Re-calibrated parameter	Reason for re-calibration
3.1	NS-SEC	<4	8	Violates structure of the data
3.3	Education	Increase fluoride score if education = 1	Decrease fluoride score if education = 1	Incorrect data scale direction
3.3	Fluoride	>2	<2	Incorrect data scale direction
3.4	NS-SEC	<4	8	Violates structure of the data
3.5	Freqden	Increase fluoride if freqden = 1	Decrease fluoride if freqden = 1	Incorrect data scale direction
3.5	Fluoride	>5	<2	Incorrect data scale direction

The second section of Pathway 4.1 stated that those with a 'number of cakes per week' score over 3 ('less than once a week', 'rarely or never') would see their sugar variable decrease. This however failed to match the scale of the data from the ADHS. The sugar variable in this dataset was binary in nature, with '1' representing high intake. Therefore, those with high cake consumption ('< 3') saw their sugar scores decrease rather than increase, to bring their scores in line with the higher sugar intake variable. This also meant changing the last part of the function, as originally those with a sugar score of ‘>2’ would see their tooth decay scores increase. This was changed to ‘<2’, to capture all those who had a high level of sugar intake. Pathway 4.2 had no such issues so no re-calibration was required. For Pathway 4.3 the second section of the function included the attendance variable ‘freqden’, where individuals with a score ‘>3’ (’less frequently that two years’, ‘only when having trouble’) saw their fluoride scores increase. This was changed to ‘>2’ to also include the ‘at least once every two years’ response. The thresholds for the first section of Pathway 4.4 were deemed correct. The second section however stated that if an agent's fluoride score was below 2 their tooth decay score would increase. According to the ADHS data scales, this meant that those with higher fluoride intake saw their tooth decay scores increase, which is incorrect conceptually. This was changed to ‘>2’, so that those with lower fluoride intake saw their tooth decay scores increase. Pathways 4.5 and 4.6 were both judged to have appropriate thresholds. The re-calibration of Pathways 4.1-4.6 are listed below in Table 30.

Table 30 - Re-calibration of parameters for Pathways 4.1-4.6

Pathway	Variable	Original parameter	Re-calibrated parameter	Reason for re-calibration
4.1	Cake consumption	High cake consumption increased sugar score	High cake consumption decreased sugar score	Incorrect data scale direction
4.1	Sugar intake	>2	<2	Incorrect data scale direction
4.3	Freqden	>3	>2	Adjusted to include appropriate data points above threshold
4.4	Fluoride	Decay increased if fluoride <2	Decay increased if fluoride >2	Incorrect data scale direction

This process demonstrates the importance of calibrating a model, as well as the difficulties associated with this process. Had this step not been taken, the modelling could have been conducted using inaccurate parameters which bore no relation to the theory or the original data source. This could have nullified the validity of the results. While the agent-based models created for this research were of a manageable size to check in this way, manually calibrating parameter values can be an arduous task. Some researchers have employed optimisation methods to calibrate their models, such as genetic algorithms (GAs), which are able to undertake parallel searches through a large number of parameters (Ngo and See, 2012). This technique works with randomly generated samples of solutions, which are each evaluated for their fit against the model using pre-defined functions. This process loops through the solutions, and uses a ‘survival of the fittest’ approach, whereby the program is terminated once a solution has been found that satisfies the threshold condition, eliminating the other solutions. In this respect GAs are not that dissimilar, both conceptually and methodologically, to certain types of spatial microsimulation modelling such as simulated annealing (Morrissey et al, 2008).

However, GA techniques were not employed in this research for a number of reasons. Firstly, creating and running such an algorithm was beyond the scope of this research. Secondly, and perhaps most importantly, there seemed to be no guarantee that this method of checking parameters would adhere to the necessary rules that match the theory used in the research. In other words, the GA may come up with an ‘ideal’ set of parameters (in the form of threshold numbers for example), but these may bear no resemblance to the data values from the original ADHS file, or properly match the theoretical concepts being tested in the model. This risks invalidating the theoretical basis of the modelling, which again would defeat the point of it. Thus it could be argued that manual calibration, although more arduous, is potentially more appropriate for models based on theory.

Time lag events affecting agents

Once the theoretical pathways had been converted into the appropriate sections of code, the next step was to decide at which intervals each section of code would be implemented in the model. Table 31 shows the different pathways in the simulation, and the number of ticks between each application of these pathways. Some pathways were assigned a greater number of ticks between their uses in the simulations, which was done to reflect the differences between the immediate conditions people live in which may affect them on a more regular basis, and those which are likely to have longer term effects, or effects at less regular intervals. This is in line with previous research, including Mitchell et al. (2002), who stated that the negative effects of unemployment on health do not take place after one day, but over a longer time period. Other authors have suggested that changes in social life may affect health quickly, while the subtler changes of growing inequality may take longer to manifest (Adler and Newman, 2002). Macintyre et al. (2002) have also stated that consideration is needed for what appropriate time intervals between environmental exposure and health effects may be, as a ‘zero time lag between exposure and outcome is usually implausible in social epidemiology’ (p.135).

Table 31 – Pathways and ticks between implementation

Pathway	Main variables associated with pathway	Ticks/days	Time period
2.1	House price and cost delaying treatment	7	Weekly
2.2	House price and cost delaying treatment	7	Weekly
2.3	House price and cost delaying treatment	7	Weekly
3.1	NS-SEC and income estimates	14	Bi-weekly
3.2	Education	14	Bi-weekly
3.3	Fluoride intake and sweet consumption	7	Weekly
3.4	Sweet and sugar consumption	7	Weekly
3.5	Fluoride intake	7	Weekly
4.1	Years of lost life	14	Bi-weekly
4.2	Years of lost life	14	Bi-weekly
4.3	Years of lost life	14	Bi-weekly
4.4	Crime scores	7	Weekly
4.5	Crime scores	7	Weekly
4.6	Crime scores	7	Weekly

This research attempted to implement a similar approach, by having some of the model interactions take a more immediate effect than others (weekly vs bi-weekly). This was also in line with trying to mimic the accumulation model of the life course approach. As with defining the probabilities for visiting patterns to different locations (Tables 21 and

22), despite having an idea of the general trend associated with the variables, this was still difficult to quantify. The same was true of tooth decay, as there is no exact breakdown of how long it takes for each process to influence tooth decay in some way. This represents the artificial nature of these processes within the agent-based models, as tooth decay is not measured or recorded in the same way in the agent-based models as it is in clinical practice. People are not told that they have 2.34 decayed teeth for example, and tooth decay does not decrease once it is present.

The values in Table 31 were therefore chosen based on the time frame of 2 years which would be simulated in each model. Given that scores for each variable increased or decreased by 0.01, with interactions occurring every 7 or 14 ticks, this meant that an individual had the chance to cross a threshold for a variable within the simulation time of 730 ticks (i.e. going from 2 decayed teeth to 3). The selection of periods of either 7 or 14 days could be considered somewhat arbitrary, however there is no real empirical evidence to base such measures on, and this approach still allowed for the differentiation between shorter term and longer term processes. Indicators such as socio-economic position (NS-SEC), income, education, and long-term health related issues ('years of lost life') were assumed to have more of a lagged effect on the health of individuals, whereas housing conditions, not being able to pay for treatment or other material items, fluoride and sugar intake, and the threat of crime were all assumed to have more immediate effects. The issue of implementing the lagged effects to mimic the accumulation model of health was addressed using the 'mod' function in NetLogo. This allowed for sections of code to be run every 'x' ticks in the model. Thus the mod function acted as a way of implementing rules at certain time points, allowing for the lag effects of certain parts of the model to be incorporated in a more realistic fashion.

The verification of the 'mod' function was essentially carried out as part of the verification of the theoretical pathways in the 'updating agent characteristics' section (page 207). In this test the theoretical pathways were set to kick in after 50 ticks of the model, which occurred correctly on each occasion except one (pathway 3.3), where the pathway had been incorrectly coded. This demonstrated the reliability of the function for implementing the lagged effects into the final models.

5.3. Simulation experiments

5.3.1. Validation

Before covering the results of the two simulation experiments, it is important to review the validation process undertaken as part of this research, as this is not included as part of the ODD protocol. As with spatial microsimulation modelling, validation is a vital part of the agent-based modelling process. It could be argued that evaluating agent-based models is harder than evaluating the outputs of spatial microsimulation models, as the results of an IPF model generally give a set of numbers which are easy to analyse statistically. Agent-based models on the other hand do not always produce outputs in this way, meaning it is important to make sure the right sort of data is teased out of the models, or appropriate validation measures identified.

Xiang et al. (2005) point to the need for validation of agent-based models, stating that this involves checking whether the theoretical and conceptual models lead to a realistic representation of real world interactions, and that output is consistent with corresponding real data. The authors go on to list a variety of ways in which agent-based models can be validated. These are listed below:

- Face validity – this step involves a more subjective approach, where experts in a given field are asked to judge if a model behaves as it should, and whether it is accurate enough compared to findings in the real world. There are two types of face validity: animation, and graphical representation. Animation involves 'the graphical display of the behaviour of the model over time' (p.48), as some software (including NetLogo) have visual elements that allow properties to be tracked as the simulation runs. Graphical representation involves displaying the output data of the model (i.e. means, distributions, etc.) in a graph format. Thus, experts are asked to judge the final results, rather than the whole sequence of the simulation. Turing Tests are a similar type of expert analysis, where the expert is given output from real world data and the simulation, and asked to see if they can tell which is which.

- Internal validity – Xiang and colleagues (2005) state that this process involves comparing the results of several runs of stochastic simulations using different random seeds (similar to random number generators) in each run. The idea is that if the model is internally valid, similar results should be seen for each run, regardless of the effects of the random seed. This would not necessarily work with the agent data in this research as the IPF simulation technique is designed to produce consistent data outputs with regard to individual characteristics. Microsimulation methods such as combinatorial optimisation may be more suited to this random seed approach. In this research it may be more relevant to apply this approach to the setup locations of agents, or the effects of certain neighbourhood based factors within the model instead.

- Historical data validation – this technique (which essentially combines elements of verification, calibration and validation) is useful when historical data exist on a system. The data can be used both to build and parameterise the model, as well as to check if the model is behaving as would be expected.

- Parameter variability (sensitivity analysis) – this validation technique involves changing the values of initial parameters for the model, to see how this influences the model's subsequent behaviour and outputs. Changes to parameters should result in the same changes occurring in the model as would occur in the real world if such alterations were made. This is also a way to identify the most sensitive parameters, and make sure they are accurate enough to be reliable.

- Model to model comparison – also known as back to back testing, or docking, this technique involves comparing the results of numerous parts of the simulation to the results of other models. Models with similar outputs could be created using different conceptual models on different platforms for example, but may also make use of similar input data.

- Statistical tests – as with the validation of spatial microsimulation models, statistical analysis can play a large role in providing objective, quantitative evaluations of model output. Xiang and colleagues (2005) comment that the inclusion of such analysis 'can significantly increase the credibility of the model' (p.49). Usually statistical analysis is used to compare the output data with the corresponding system, or with data from other models that have used similar input data. Time series, means, variances and aggregations are all given as examples by the authors of graphical data for development, validation or Turing Tests. In their example validation of natural organic matter molecules, goodness of fit tests (Chi-Square, Kolmogorov-Smirnov) as well as factorial ANOVA are mentioned as potentially useful validation methods.

While agent-based models can be more difficult to validate compared to other types of models (namely spatial microsimulation models), the agent-based models in this research were affected by a similar problem to the microsimulation models, this being the lack of data on tooth decay in adults available at the small area level against which to validate model outputs. This does highlight a potential downside to such an approach, and is something that should be borne in mind for future research. Despite being an exploratory piece of research, this does not mean that these models could not be validated in some way. While not all of the methods suggested by Xiang et al. (2005) could be used in this research, certain of these could be made use of to test the data created through the simulations. Both models have already been verified and calibrated, with the tracing process suggested by Xiang et al. (2005) proving to be vital in verifying the models for example.

Further, model to model validation was also possible. As the two models were from two differing areas of Sheffield, and differ greatly with regard to deprivation, socio-economic status, and a range of other social indicators, it might be expected that they would have differing outputs with regards to tooth decay. Due to the different number of pathways involved in the simulations it was also possible to perform parameter variability tests using different combinations of these throughout the simulation runs. Finally, the outputs of the various different simulations were also compared statistically to see if any significant differences existed, and at which stage of the model these may have occurred. The next section will outline the process of running the final models, and the analysis of the results from these.

5.3.2. Simulation one – adding the pathways sequentially

The simulations were conducted using the University of Sheffield's supercomputer 'Iceberg'. Twenty-eight separate models were created for the first experiment (14 for Sheffield East, 14 for Sheffield West), with pathways being added sequentially with each new model. As stated previously, each simulation was run for 730 steps, which represented 1 day per tick, over the course of 2 years. Each pathway simulation was run for 10 iterations, and the average score from these iterations was taken as the score for that simulation, in order to allow for the probabilistic nature of the models. The results of the simulations for the Sheffield East study area are presented in Table 32, while Table 33 shows the results for Sheffield West. Within the tables, ‘model run’ refers to the number of pathways that were included (similar to Table 20). For example, model run 1 refers to the model containing only the first pathway (Pathway 2.1), while model run 2 contains the first two pathways (Pathways 2.1 and 2.2). This approach is repeated through to model run 14, which contains all of the pathways identified in Table 24.

Table 32 – Tooth decay scores from each of the simulation runs in the Sheffield East study area

Model run	Mean decayed teeth - original	Mean decayed teeth - final	Change (n)	Change (%)
1	46686.6	42828.86	-3857.74	-8.26
2	46686.6	46491.5	-195.1	-0.42
3	46686.6	53855.84	7169.24	15.36
4	46686.6	49338.12	2651.52	5.68
5	46686.6	49330.7	2644.1	5.66
6	46686.6	54459.4	7772.8	16.65
7	46686.6	23056017.53	23009330.93	49284.66
8	46686.6	23604131.47	23557444.87	50458.69
9	46686.6	23823930.36	23777243.76	50929.48
10	46686.6	22488133.03	22441446.43	48068.28
11	46686.6	20773675.92	20726989.32	44396.01
12	46686.6	17740303.38	17693616.78	37898.70
13	46686.6	22541373.66	22494687.06	48182.32
14	46686.6	23892514.47	23845827.87	51076.39

Table 33 – Tooth decay scores from each of the simulation runs in the Sheffield West study area

Model run	Mean decayed teeth - original	Mean decayed teeth - final	Change (n)	Change (%)
1	24379	19521.68	-4857.32	-19.92
2	24379	24706.26	327.26	1.34
3	24379	33068.94	8689.94	35.65
4	24379	28487.62	4108.62	16.85
5	24379	24733.1	354.1	1.45
6	24379	32063.6	7684.6	31.52
7	24379	50482041.9	50457662.9	206971.83
8	24379	50523178.77	50498799.77	207140.57
9	24379	50289110.81	50264731.81	206180.45
10	24379	50411316.67	50386937.67	206681.72
11	24379	50692145.71	50667766.71	207833.65
12	24379	50627428.37	50603049.37	207568.19
13	24379	51803747.65	51779368.65	212393.32
14	24379	50582640.46	50558261.46	207384.48

In order to assess whether there were any significant changes between the model runs, the results were analysed statistically. Originally a one-way repeated measures ANOVA was considered, however the results from the agent-based models were found to be non-normally distributed, and could not be transformed to a normal distribution using either log or square root functions. The Friedman Test was therefore also considered, this being the non-parametric equivalent to the repeated measures ANOVA. There is some debate over the use of non-normally distributed data in repeated measures tests however, with some literature pointing to the fact that ANOVA tests can still be applied to skewed datasets. This is summarised in Field et al. (2012), who state that ‘the evidence suggests that when group sizes are equal the F-statistic can be quite robust to violations of normality’ (p.413). Given this, and the fact that the post-hoc analysis options for ANOVA are far more extensive than those for the Friedman Test, it was decided that the data would first be tested using the Friedman Test then the ANOVA, and if the results were similar further analysis would be conducted using the ANOVA post-hoc (Bonferroni) tests. One of the assumptions of one-way repeated measures ANOVA is complying with Mauchly’s Test of Sphericity, however this only applies when more than two levels are included in the data, and as the results of the agent-based models were for two time points the data did not violate this assumption.

The descriptive statistics for the Friedman Test for the Sheffield East study area can be seen below in Table 34. The Friedman Test showed there to be a statistically significant difference between the original scores and the scores post simulation (X^2 = 68.600, P < 0.05), indicating that tooth decay scores increased as additional pathways were added.

Table 34 – Descriptive statistics for the Sheffield East study area (Friedman Test)

	N	Mean	Std. Deviation	Minimum	Maximum
Original score	140	46686.60	.000	46687	46687
New score	140	12729741.73	11752255.150	-21197920	28330281

This did not give much insight into the pattern associated with the data however, so these were further analysed using the ANOVA test. The descriptive statistics from this test are shown below in Table 35. The Sheffield East ANOVA showed a statistically significant difference between the time points in the data (F(1.000, 13.000) = 1410.471, P < 0.05), and as this matched the conclusions of the Friedman Test it was judged appropriate to use the Bonferroni post-hoc test results from the ANOVA model.

Table 35 – Descriptive statistics for the Sheffield East study area (ANOVA)

	Model	Mean	Std. Deviation	N
Original model	1	46686.60	.000	10
	2	46686.60	.000	10
	3	46686.60	.000	10
	4	46686.60	.000	10
	5	46686.60	.000	10
	6	46686.60	.000	10
	7	46686.60	.000	10
	8	46686.60	.000	10
	9	46686.60	.000	10
	10	46686.60	.000	10
	11	46686.60	.000	10
	12	46686.60	.000	10
	13	46686.60	.000	10
	14	46686.60	.000	10
	Total	46686.60	.000	140
New model	1	42828.86	.000	10
	2	46491.50	.000	10
	3	53855.84	.000	10
	4	49338.12	.000	10
	5	49330.70	.000	10
	6	54459.40	1.671	10

	7	23056017.53	1230455.505	10
	8	23604131.47	2853969.941	10
	9	23823930.36	1733739.780	10
	10	22488133.03	2694973.098	10
	11	20773675.92	1272643.452	10
	12	17740303.38	13844567.463	10
	13	22541373.66	1078614.480	10
	14	23892514.47	3024731.118	10
	Total	12729741.73	11752255.150	140

The post-hoc analysis demonstrated a significant interaction between time and model run ($p < 0.05$). Analysis of the pairwise comparisons showed that model runs 1-6 were not statistically significantly different from each other ($p = 1.000$), whereas these model runs were statistically significantly different to model runs 7-14 ($p < 0.05$). Model runs 7-14 were not statistically significantly different to each other, leading to the conclusion that the introduction of model run 7 (Pathway 3.4 - where agents interact with shops, which influences diet, and then health damaging behaviours) had the most significant impact on tooth decay scores in this area. These results can be seen in Appendix A.

The descriptive statistics for the Sheffield West study area for the Friedman Test can be seen below in Table 36. The overall model was statistically significant ($X^2 = 102.857$, $P < 0.05$), implying statistically significant differences between the number of decayed teeth in the population between the two time points.

Table 36 – Descriptive statistics for the Sheffield West study area (Friedman Test)

	N	Mean	Std. Deviation	Minimum	Maximum
Original score	140	24379.00	.000	24379	24379
New score	140	28969585.11	25180128.717	19522	63334820

The same data were then analysed using the one-way repeated measures ANOVA. The descriptive statistics for this can be seen in Table 37.

Table 37 – Descriptive statistics for the Sheffield West study area (ANOVA)

	Model	Mean	Std. Deviation	N
Original score	1	24379.00	.000	10
	2	24379.00	.000	10
	3	24379.00	.000	10
	4	24379.00	.000	10
	5	24379.00	.000	10
	6	24379.00	.000	10
	7	24379.00	.000	10
	8	24379.00	.000	10
	9	24379.00	.000	10
	10	24379.00	.000	10
	11	24379.00	.000	10
	12	24379.00	.000	10

	13	24379.00	.000	10
	14	24379.00	.000	10
	Total	24379.00	.000	140
New score	1	19521.68	.000	10
	2	24706.26	.000	10
	3	33068.94	.000	10
	4	28487.62	.000	10
	5	24733.10	.000	10
	6	32063.60	.000	10
	7	50482041.90	432210.453	10
	8	50523178.77	246795.721	10
	9	50289110.81	534639.868	10
	10	50411316.67	580296.236	10
	11	50692145.71	346456.806	10
	12	50627428.37	313984.183	10
	13	51803747.65	4067459.632	10
	14	50582640.46	467343.027	10
	Total	28969585.11	25180128.717	140

The ANOVA again demonstrated a statistically significant difference between the two time points ($F(1.000, 13.000) = 91988.461$, $P < 0.05$). Thus, as the Friedman and ANOVA tests returned similar results, it was again deemed appropriate to use the Bonferroni post-hoc analysis as part of the ANOVA test. Firstly, the post-hoc tests showed there to be a significant interaction between time and model run ($p < 0.05$). As can be seen from the data tables in Appendix B, model runs 1-6 showed no statistically significant differences to each other (p = 1.000). However, model runs 1-6 were statistically significantly different to model run 7 ($p < 0.05$), and all those that followed (8-14). Model runs 7-14 again showed no statistically significant differences to each other. Once again this pointed to model run 7 (Pathway 3.4) as a key point in the modelling, where tooth decay scores increased significantly but not thereafter.

The trend associated with these patterns are highlighted for Sheffield East (Figure 56) and Sheffield West (Figure 57) below. Time 1 represents the baseline mean tooth decay score, while time 2 represents the mean scores after each pathway was added.

Figure 56 – Trend of mean scores for each model run (Sheffield East)

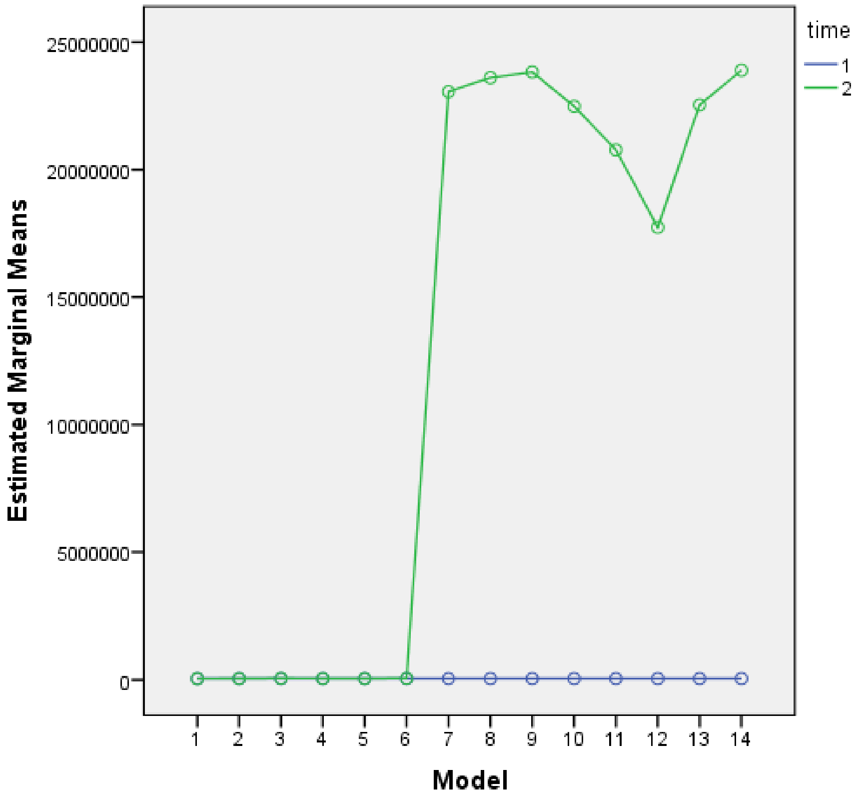
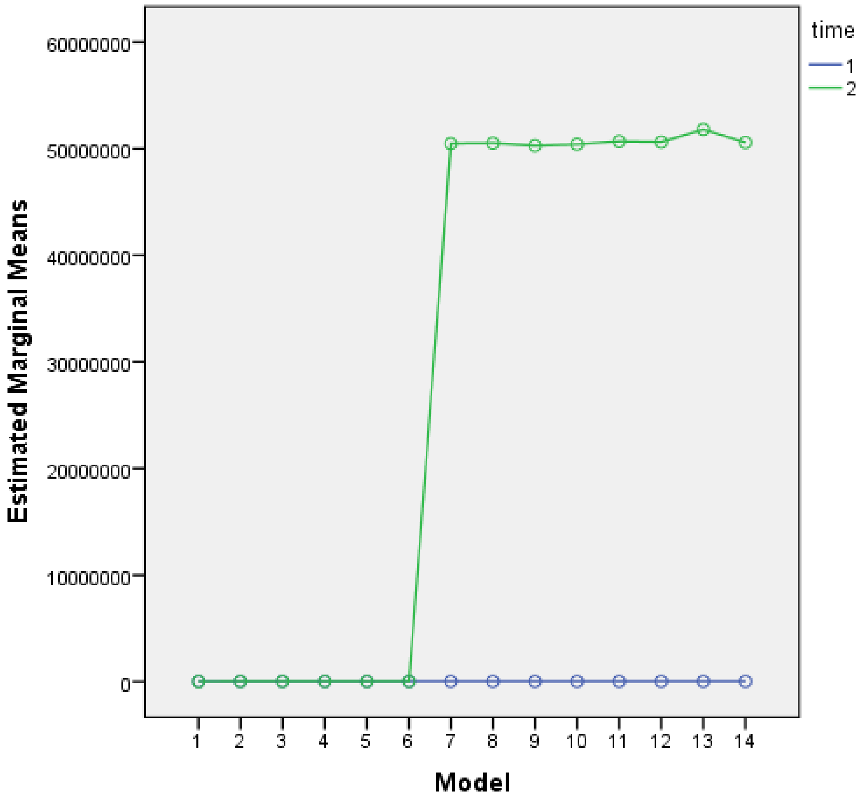


Figure 57 – Trend of mean scores for each model run (Sheffield West)



As can be seen from Figures 56 and 57, the trend in the data is more steady in the Sheffield West model, where the data showed little variation beyond the implementation of model run 7. Figure 56 shows that in the Sheffield East study area there was more variation after model run 7 (Pathway 3.4) had been implemented, although this was not statistically significant. It is interesting that Sheffield East saw a reduction in its mean tooth decay score upon the implementation of model run 10 (Pathway 4.2), 11 (Pathway 4.3) and 12 (Pathway 4.4). These pathways saw the introduction of tooth brushing habits (Pathway 4.2), attendance, fluoride and tooth brushing (Pathway 4.3), and fluoride intake via social capital (Pathway 4.4). Again while these pathways did not have a statistically significant impact on the overall trend, fluoride consumption, tooth brushing and dental attendance are all variables that would be expected to be beneficial to oral health, and it is interesting that these resulted in large reductions in Sheffield East, but not Sheffield West. What is clear from both figures however, is the large effect that the implementation of the seventh pathway had in both study areas.

These results show that there was a statistically significant difference between the original tooth decay scores for each of the study areas, and the new tooth decay scores after the pathways had been added. Further analysis of this revealed an interesting pattern, which involved no statistically significant differences after the first 6 model runs (Pathways 2.1, 2.2, 2.3, 3.1, 3.2 and 3.3), before a statistically significant increase in the number of decayed teeth upon the addition of the seventh model run. This involved the addition of Pathway 3.4, which allowed agents to visit shops and supermarkets, through which they consumed sugar. Beyond this, there were no statistically significant differences in model runs 8-14 (Pathways 3.5, 4.1-4.6). The ability to visit shops and supermarkets, and the consumption of sugar through these visits therefore had the most influence on the model outcome variable score, and stands out as the most important pathway statistically.

5.3.3. Simulation two – analysing the effects of a new shop

The next set of simulations were also conducted using the University of Sheffield's supercomputer. The scores from model run 14 were used as the baseline data, as this represented the agent-based model with all of the theoretical pathways included. The extra shop locations were then added to this (see Section 5.2.5), and tested separately. The baseline and simulation scores for Sheffield East can be seen in Table 38.

Table 38 – Baseline model compared to the new simulation (Sheffield East)

Simulation run	Decayed teeth – baseline	Decayed teeth – new simulation
1	23892514.47	24628229.57

The original idea had been to compare these two datasets using dependant t-tests, however the data for both the baseline data and the simulation involving the extra shop were non-normally distributed. The Wilcoxon Signed-Rank Test, the non-parametric equivalent of the dependant t-test, was therefore used instead. The descriptive statistics for this test for Sheffield East are shown below in Table 39.

Table 39 – Descriptive statistics for the Wilcoxon Signed-Rank Test (Sheffield East)

	N	Mean	Std. Deviation	Minimum	Maximum
Original score	10	23892514.4660	3024731.11816	18421135.12	26848098.84
New score	10	24628229.5720	2664690.91004	20328246.40	29180923.78

The test statistics showed there to be no statistically significant differences between the two scores ($Z = .000$, $p = 1.000$), implying that the tooth decay scores in this area remained relatively stable through this experiment. The Sheffield West data were then analysed in a similar manner, and the baseline and simulation scores can be seen below in Table 40.

Table 40 – Baseline model compared to the new simulation (Sheffield West)

Simulation run	Decayed teeth – baseline	Decayed teeth – new simulation
1	50582640.46	47456156.26

These data were also checked for the normality of their distribution. The data for the baseline model were normally distributed, while the data for the new simulation with the extra shop location were non-normally distributed, so again the non-parametric Wilcoxon Signed-Rank Test was used. The descriptive statistics for the Sheffield West study area are presented below in Table 41.

Table 41 – Descriptive statistics for the Wilcoxon Signed-Rank Test (Sheffield West)

	N	Mean	Std. Deviation	Minimum	Maximum
Original score	10	50582640.4620	467343.02726	49794767.24	51535015.52
New score	10	47456156.2640	785236.81633	45880398.88	48406913.84

This test showed there to be a statistically significant difference between the two scores ($Z = 10.000$, $p = .002$), indicating that the presence of an extra shop had a significant effect on tooth decay scores in the Sheffield West study area. Looking at the data in Table 41, this can be inferred as a statistically significant decrease in tooth decay scores.

Overall this demonstrates that adding an extra shop to each study area had differing effects. When added to the Sheffield East study area tooth decay scores remained similar to those in model run 14 (the final model in Table 20, containing all of the theoretical pathways), with no statistically significant differences between these two models. However, the addition of a shop to the Sheffield West study area resulted in a statistically significant reduction in tooth decay compared to model run 14. While additional shops may be expected to increase decay due to the increase in available produce, it may be that in the case of Sheffield West, the produce available aids in reducing tooth decay scores. This will be discussed in greater detail in Chapter 6 (Section 6.2.2).

5.4. Conclusions

This chapter has helped to achieve both the second and third research objectives outlined in Section 1.2.

- To build simulation models capable of representing these theoretical pathways
- To use these simulation models to find the most influential theoretical pathways within different neighbourhoods within Sheffield

Regarding the achievement of the first of the two objectives listed above, the agent-based models were combined with the base data from the spatial microsimulation model, to create fully functioning simulation models for two contrasting locations within Sheffield, which represents a rare example of the combination of these two techniques. This stage of the research allowed for the individuals in a representative population dataset to be combined with an interactive modelling environment. The flexibility and interactivity of this environment allowed the second objective to be achieved, as the individual pathways were added, then tested sequentially to assess their effect on the outcome variable score in the two study areas. By achieving this second objective (the third overall) all three of the main elements of this research were brought together, culminating in the achievement of the overall aim of this research, to investigate how neighbourhood effects may influence patterns of spatial inequalities in tooth decay, in an exploratory manner.

This chapter has demonstrated the flexible nature of agent-based models. Further, it has also demonstrated the complexity of such models, and the need for careful verification and calibration as the minimum that should be conducted when creating them. Standardised approaches such as the ODD protocol may provide a way to communicate the details of such complex models in an understandable and reproducible way. While the values from the models were not as accurate as would have been liked, the trends associated with these demonstrated that the pathway involving the introduction of shops/supermarkets, which influenced agents' diet and health damaging behaviours, had the most significant effect of all of the pathways. This was the case for both study areas. The addition of an extra shop to each area had a statistically significant impact in the Sheffield West study area, but not in the Sheffield East study area. A more detailed discussion of the potential mechanisms behind both of these simulations will be provided in Chapter 6.

As with previous chapters a number of limitations were noted throughout, including data availability, pathway simplification, the use of administrative data to define the spatial units used in this research, and the computational intensity of the agent-based models. These will be discussed in more detail in Chapter 6.

Chapter 6 – Discussion

‘We do not know a truth without knowing its cause’ – Aristotle (384–322 BC)

6.1. Introduction

The overall aim of this thesis was to investigate how neighbourhood environments influence patterns of spatial inequalities in tooth decay. To achieve this aim, the objectives of this research were to identify theoretical pathways by which neighbourhoods influence tooth decay, to build simulation models capable of representing these theoretical pathways, and to use these simulation models to find the most influential theoretical pathways within different neighbourhoods. This research was conducted in an exploratory manner, due to the lack of previous research to guide it, and the difficulties in producing accurate real world values. The trends associated with the output of the agent-based models were therefore the main outcome studied in this research, as opposed to the exact values produced through these simulations. The goal was to analyse the trends from the data from the agent-based models to determine which of the pathways had the most impact on tooth decay, from which potential key mechanisms could be identified.

This research was the first to take a geographical approach to studying neighbourhood determinants of tooth decay, as well as being the first in the field to incorporate multiple data sources and a theoretical framework into an agent-based model. The theoretical framework (Macintyre et al, 2002) was used to identify pathways by which neighbourhoods influence tooth decay, before using spatial microsimulation to create a population of individuals with associated characteristics, which the pathways were tested on using the agent-based models. In general, there has been a lack of simulation modelling within Dental Public Health, with only five other examples of agent-based models identified within the field (Metcalf et al, 2013; Roudsari et al, 2016; Wang et al, 2016; Sadeghipour et al, 2016; Sadeghipour et al, 2017). This study has highlighted the benefits of such an approach for Dental Public Health, and it is hoped that this work will be built on in future.

The discussion of this research will be split into three sections. Section 6.2 will discuss the key findings in relation to the main themes and determinants of tooth decay identified in Section 2.5 of the literature review, and whether the importance of these themes in determining tooth decay scores was supported or not by the agent-based modelling. The themes that were observed to have a significant impact on the outcome variable score within the agent-based models will be addressed first, before discussing those that did not. It is important to remember that just because a pathway was not judged to have a significant impact on the model, it does not mean that the themes or variables in that pathway are not significantly associated with tooth decay. This is because the statistical analysis looked for pathways that lead to a statistically significant increase or decrease in the total tooth decay score for each study area, rather than significant associations between variables and tooth decay individually. Section 6.3 will then provide a brief comparison of the results of key themes and determinants between the two study areas, with a discussion of these trends. Finally, Section 6.4 will consider the methodological and conceptual strengths and limitations of the overall approach taken within this thesis, before discussing assumptions that were made in designing the models.

6.2. Key findings

6.2.1. Key themes and determinants supported by the agent-based modelling

Area of residence and features of the built environment

A number of built neighbourhood level variables have been studied in relation to oral health. Tellez et al. (2006) included churches in their analysis and demonstrated associations between these and cavitated carious lesions, while Aida et al. (2008) found associations between community centres and tooth decay. Borenstein et al. (2013) included a variety of neighbourhood level features including supermarkets, fast food restaurants, number of social services, and number of parks per Census Tract. They found significant associations between social services and dental visits, as well as between park space and self-rated oral health. Other studies have included neighbourhood level features in the form of people’s opinions on access to them, including public libraries, medical services, parks and banks, however no significant relationships with self-reported oral health were found (Finlayson et al, 2010). Studies from Brazil have also included a number of other neighbourhood based variables such as overcrowding which was associated with dmft scores (Antunes et al, 2002), areas of substandard living (Moyses et al, 2008) which proved to have associations with dental trauma, and urban social facilities per health district (Carvalho et al, 2010), which were found not to be a predictor of dental trauma in children.

Beyond these variables, there were two that stood out in particular, due to their more common occurrence within the literature. The first of these was dental practices. Given the literature regarding the importance of dental attendance (Listl, 2012), as well as the potential uneven distribution of these services (Jones et al, 2001), dental practices were an obvious feature of the local environment that needed to be included. The second variable was the presence of shops and/or supermarkets. Given that the oral health literature has shown decay to be influenced by diet (Moynihan and Petersen, 2004), and that other authors have hypothesised the importance of shops (with regard to both access, and available foods) for oral health (Mobley et al, 2009; Fonseca, 2012) it was the logical decision to include shops as a second neighbourhood based feature. Both of these features were represented by point data, marking the geographical location of the relevant buildings that fell within the study area boundaries. Shops and dentists were both included as part of the third theme within the theoretical framework used to guide this research (Macintyre et al, 2002), namely 'services provided, publicly or privately to support people in their daily lives'.

The locations of dental practices were included in one pathway (Pathway 3.5) which was concerned with the pathway between dental attendance and dental knowledge. This pathway did not have a significant effect on the outcome variable score in either study area, which was extremely surprising given the emphasis in the literature on the importance of attendance at dentists (Donaldson et al, 2008), and the subsequent benefits of this (Listl, 2012). Practices were not differentiated based on whether they were NHS or private dentists however, and this may have influenced this result. However, given that there were only two practices in the Sheffield West model, and one in the Sheffield East study area, it could be argued this would have made little difference in the context of this research. Such differentiation in a study of the whole city may have had more of an impact. The locations of shops and supermarkets were included in one pathway (Pathway 3.4), which analysed the interactions between visiting a shop, diet, and health damaging behaviours. This pathway was shown to have a significant impact upon the outcome variable score in both the Sheffield East and Sheffield West study areas. The importance of shops and supermarkets has been stressed in the literature so it is not surprising to see them have a significant impact on tooth decay, and it may be that the dietary parts of this pathway also played a role in its significance. This will be discussed in more detail in Section 6.2.2.

Diet

Diet has been shown to be vitally important to oral health, and tooth decay in particular. Nutritional intake has been shown to differ between income groups (National Diet and Nutrition Survey, 2014), while women with lower educational attainment were shown to have less control over family food choices, with economic pressures limiting the amount of healthy food available (Barker et al, 2008). Studies have also shown that areas of low socio-economic status were more likely to be served by shops containing low-nutrient food (Mobley et al, 2009), while Fonseca (2012) identified low income areas as tending to be served by fewer large supermarkets that offer a variety of produce. Sheiham (2006) had commented on the susceptibility of undernourished children to decay, while decay leading to tooth loss has been associated with diets lacking fruit, vegetables and other vitamins (Moynihan and Petersen, 2004).

Conclusive links have also been made between sugar consumption (including frequency of consumption) and decay (Sheiham, 2001). High sugar intake through soft drinks has been linked to decay in low income adults (Burt et al, 2006), as well as in children from low socio-economic backgrounds (Warren et al, 2009). Dietary related diseases such as metabolic syndrome have also been associated with caries and decayed teeth (Timonen et al, 2010; Ojima et al, 2015). Diet featured in three of the themes of the theoretical framework used in this research, the first being theme two (‘healthy environments at home, work and play’), due to the effects material circumstances can have on diet through financial constraints (Barker et al, 2008; Fonseca, 2012). Diet also featured in the third theme (‘services provided, publicly or privately to support people in their daily lives’), through health damaging behaviours associated with dental knowledge (Moynihan and Petersen, 2004), and through produce purchased from supermarkets (Mobley et al, 2009). Finally, diet also featured in the fourth theme (‘socio-cultural features of a neighbourhood’), due to the effects of collective health behaviours affecting dietary outcomes (Moynihan and Petersen, 2004; Warren et al, 2009).

Diet was therefore featured in four separate pathways in the agent-based modelling; three times as 'diet', and once as 'health damaging behaviours'. The first of these (Pathway 2.2) was the pathway between material circumstances, financial constraints, purchasing power, diet and sugar intake. Within this pathway diet was represented by the number of cakes eaten per week, while sugar consumption was also included as an additional variable representing sugar intake. The second pathway to include diet was Pathway 3.3, the pathway between education, dental knowledge, and health damaging behaviours. In this pathway diet was represented by sweet consumption as a proxy for health damaging behaviours. The third use of diet in the agent-based models came in Pathway 3.4, between supermarket/shop visiting, diet and health damaging behaviours. In this pathway, diet was represented twice, first through sugar intake acting as a proxy for diet, with consumption of sweets as a proxy for health damaging behaviours. The fourth and final use of diet was in Pathway 4.1 which measured the interactions between health behaviours, diet and nutrition/sugar intake. In this pathway diet was represented by the number of cakes eaten per week, while nutrition/sugar consumption was represented by the 'high sugar intake' variable from the ADHS.

It is surprising that only one pathway containing dietary based pathways, measuring interactions between visiting a shop, diet and health damaging behaviours (Pathway 3.4), was found to have a significant impact on the outcome variable score. This significant effect occurred in both study areas. The addition of the shop locations to the model may also have played a role in these results, given the links between these locations and diet demonstrated in the literature (Mobley et al, 2009; Fonseca, 2012).

The pathway between material circumstances, financial constraints, diet and sugar intake (Pathway 2.2) included sugar consumption for the 'sugar intake' construct, which has been shown to be harmful to oral health (Sheiham et al, 2001; Burt et al, 2006). While the inclusion of sugar in this pathway is a logical choice based on the literature, it may be that using the number of cakes eaten per week for the 'diet' variable was not an accurate enough proxy. The pathway between education, dental knowledge and health damaging behaviours (Pathway 3.3) may also have been affected by the operationalisation of variables, as sweet consumption may not be an accurate enough variable to represent health damaging behaviours. Nevertheless, it is surprising that a variable concerning sugar consumption did not have a significant influence, particularly when it formed part of another dietary based pathway that did have a significant effect.

Pathway 4.1, the pathway between health behaviours, diet and nutrition, may have also suffered from the same problem as the previous two pathways, as the lack of impact may be due to the use of the ‘ncakes’ variable as a proxy for diet. This could potentially suggest that sugar consumed in other ways not specifically related to cake consumption is more damaging to oral health.

Given the significance of the pathway between shops, diet and health damaging behaviours, it may be that certain combinations of dietary behaviour and sugar consumption (i.e. high sugar intake and consumption of sweets) have more of an impact on tooth decay than others, or more influence than individual dietary variables, such as the case of sweet consumption in the pathway between education, dental knowledge and health damaging behaviours (Pathway 3.3). Given that the two themes that were shown to have a significant impact on the outcome variable score (shop locations and diet) were from the same pathway, these will be discussed together in Section 6.2.2 below.

6.2.2. Shops, sugar and tooth decay

Pathway 3.4 started by introducing shops as locations for agents to visit, however once they arrived the processes that occurred were based at the individual level. As Table 24 shows, the second and third sections of this pathway were concerned with high sugar intake, and the consumption of sweets by individuals. Section 6.2.1 outlined the adverse impacts on oral health, particularly tooth decay, associated with high consumption of sugars. While the sweet consumption construct referred to sugar intake through a particular type of produce, the high sugar intake variable (‘highsug’) included high intake (6 or more times a week) of cakes, biscuits, puddings, pastries, sweets, chocolate or fizzy drinks, so it is not clear which sugar related produce may be the most important in this construct. It may well be that sweet consumption was the most important part of this second construct, however there is no way to tell. There is also no way to gauge from the results which individual elements of the pathway influenced the significant outcome the most.

One explanation for the significance of this pathway could be that individual level behaviours which are known to be important to oral health (for example, diet and sugar intake) can be mediated by features of the physical environment, in this case shops. This could indicate that both neighbourhood and individual level factors have been identified that could play an important role in tooth decay scores, with one neighbourhood level variable (shops/supermarkets) and two individual level variables (diet and health damaging behaviours) associated with a significant effect on the model. This would be broadly in line with the findings of previous conventional studies of the effects of neighbourhoods on health, which have argued for the importance of contextual environments, including Curtis and Rees-Jones (1998) who demonstrated empirical evidence for the role of context in explaining health inequalities, independent of individual level attributes. This was despite individual level characteristics 'explain[ing] more of the statistical variability in health' (p.667). Pickett and Pearl (2001) and Riva et al. (2007) echoed these assertions, despite the latter commenting on the methodological weakness of much of the literature. Macintyre et al. (2002) summarised general trends within the health inequalities literature, stating that 'most investigators have tended to conclude that where you live matters for health, although probably not as much as who you are' (p.128). Pearson et al. (2013) have previously stated that there are three elements of neighbourhoods that influence health, and have generally been found throughout the literature – built, physical and social elements. This research has demonstrated that all three of these elements are potentially important for tooth decay, although only two were tested (built and social) due to the physical elements (water fluoridation) not being present in the study areas.

The findings of this research have therefore added to knowledge of theory surrounding neighbourhoods in relation to oral health which was previously a very underdeveloped section of the literature, by demonstrating that features such as shops and supermarkets are potentially the most important contextual level features for tooth decay through the dietary processes that occur through them. More specifically, the agent-based models have allowed for a better understanding of the effects of theoretically relevant processes by adding them cumulatively in an interactive and dynamic nature, allowing for the effects of individual pathways to be assessed.

While the connection between increased tooth decay and the intake of sugar is not surprising, fewer studies have investigated the importance of shops to oral health outcomes. Of those that have, most have argued for (Mobley et al, 2009; Fonseca et al, 2012) or found significant relationships between shops and decay (Tellez et al, 2006; Aida et al, 2008), while one study found no associations (Borenstein et al, 2013). These studies provide the main evidence in supporting the idea that shops and supermarkets may be important in determining levels of tooth decay. A review by Mobley et al. (2009) found that residential location influenced the ability of families and individuals to buy healthy food, this being a particular problem for low socio-economic groups and ethnic minority populations. Mobley and colleagues further stated that in rural and poorer urban areas stores selling ‘high energy, low nutrient dense foods’ (p.6) were more common than larger stores offering access to a range of healthy and nutrient rich foods. Fonseca’s (2012) review also revealed that low income groups have particular problems accessing such stores, due to the price of healthy produce, and the distance of the stores from their homes, representing a double burden. Tellez et al. (2006) demonstrated positive associations between cavitated carious lesions and number of grocery stores in a neighbourhood in their multilevel modelling of the association between neighbourhood characteristics and caries, in a low income African-American population. Aida et al. (2008) also found significant associations between the number of grocery stores per resident and higher levels of dmft in their multilevel analysis of the effect of social context on caries in 3-year-olds in Japan. Conversely, Borenstein et al. (2013) used multilevel modelling to test for associations between oral health and individual level, socio-economic and neighbourhood level factors in Toronto. While social services per Census Tract and park space showed associations with oral health outcomes, supermarkets showed no significant associations with self-rated oral health, dental visits, or dental insurance coverage.

While previous research has highlighted the importance of dentists to oral health (Donaldson et al, 2008; Listl, 2012), the finding here that shops were more important for tooth decay is conceptually logical. For example, in the majority of urban areas in the UK, shops and supermarkets are likely to outnumber both dentists and educational facilities, and are also likely to be visited on a more regular basis, even when taking into account individuals attending regular classes at educational facilities, or with quick repeat appointments at dentists. In the Sheffield East study area, there were six shops compared to one dentist and two educational facilities, while in the Sheffield West study area there were four shops and two dentists. Given that shops made up the majority of the physical locations in the models, and had the highest probability of being visited, it is perhaps hardly surprising that they had the largest effect. This in a way acts as a further argument for the validity of the models. The results are also consistent with those that place importance on food environments in the wider public health literature. For example, Inagami and colleagues (2006) found that the type of shop and distance travelled to these were associated with BMI levels. This research also found that supermarket presence was associated with lower obesity levels, with the converse true of convenience stores (Morland et al, 2006), as well as fast food establishments (Morland et al, 2009). Caution should be exercised in that all of these studies were based in North America, as studies of diet and food environments have been shown to differ between US and UK contexts (Macdonald et al, 2011).

The work of Mobley et al. (2009) and Fonseca (2012) regarding the difficulties that certain groups face in accessing healthy and nutritious food raises an important issue, as it may not simply be the presence of shops and supermarkets that are important, but rather the types of goods they sell. For example, there does not seem to be a guarantee that access to supermarkets and shops with a wide range of food will always lead to positive health habits. This is supported by the work of LeDoux and Vojnovic (2014). Their research investigated the link between the residential neighbourhood food environment and diet among low income populations in Detroit. Through the use of household questionnaire data on dietary habits, as well as GIS to measure distances to supermarkets, and a measure of 'cumulative opportunities' to capture individual's immediate food environments, the authors found that while residents with a greater number of supermarkets in their vicinity consumed healthier food, they were also more likely to consume a higher number of sugary and salted items as well. The cumulative opportunities of neighbourhood food environments were shown to be more important than proximity to supermarkets, as fast food establishments in the local environments reduced the intake of fruit and vegetables. Proximity also played a key role though, and LeDoux and Vojnovic (2014) comment on the complicated way the influence of supermarkets decreases with increased distances from them, and how some residents were 'forced by the composition of their food environment to expend additional resources to reach food outlets that provide affordable, quality and nutritious foods' (p.16). This is similar to conclusions drawn by other authors cited in this research (Mobley et al, 2009; Fonseca, 2012). However, the complexity of the findings of LeDoux and Vojnovic's (2014) work should be born in mind especially in the context of the findings of this research, as proximity or types of food sold in a supermarket may not be enough to explain shopping and/or consumption patterns on their own.

With this previous research in mind, it was interesting to note that the second simulation experiment (where an extra shop was added to each area to assess its effect on tooth decay scores) produced divergent results for the two study areas. In the Sheffield East study area, there was no significant difference between the baseline tooth decay score (a model including all of the theoretical pathways) and the new simulation (a model with all the pathways and an additional shop). However, in the Sheffield West study area there was a statistically significant reduction in the tooth decay score between the baseline and new simulation model. This indicates that the presence of an extra shop in Sheffield West helped to reduce tooth decay while, with the same scenario, levels of tooth decay remained similar in Sheffield East. While the statistical tests offer no details as to why this may have occurred, one explanation could be that the types of food available in Sheffield West may be more conducive to better diets and oral health outcomes, whereas the types of food available in Sheffield East could maintain tooth decay scores on the current trend. This could tentatively indicate that both the number of shops and the types of food sold are important for decay in Sheffield West, while only the types of food sold are important to decay in Sheffield East. The trends in Sheffield West would seem to fit with the conclusions of LeDoux and Vojnovic (2014), in that the cumulative opportunities of local food environments seem to be influential, as more shops, possibly selling certain types of produce, helped to reduce decay scores. The cumulative opportunities of local food environments in Sheffield West may therefore be more beneficial for oral health, as while decay scores decreased in this part of the city, they remained stable in Sheffield East, suggesting fewer opportunities in the latter local food environment to influence decay scores.

6.2.3. Themes not supported by the agent-based modelling

Income

Income has been shown to be a key determinant of inequalities in tooth decay, and there is a large amount of literature that has established this association, as reviewed in Section 2.5. Income inequalities measured using the Gini coefficient have been shown to have negative effects on the oral health of societies (Pattussi et al, 2001; Celeste et al, 2009), while other studies measuring average or personal income have also shown these measures have significant associations with tooth decay scores (Aida et al, 2008; Costa et al, 2012), particularly for those in the lowest income brackets (Geyer et al, 2010).

Income has been related to tooth decay via a number of potential mechanisms, including access to dental care, access to fluoridated water, as well as access to helpful information about oral health and products that help to maintain it (Costa et al, 2012). Disinvestment in public services, the erosion of social capital, and psychological stresses through social comparisons have also been highlighted as ways in which income might lead to inequalities in decay (Bhandari et al, 2015). Income levels have also been linked with diet, with regard to both the types of food that families can afford, as well as the types of foods available in low income areas (Fonseca, 2012). Through these mechanisms income is closely linked with the material circumstances in which people live. Income was included within the third theme in Macintyre and colleague's (2002) theoretical framework ('services provided, publicly or privately to support people in their daily lives') as a measure of earnings from employment opportunities.

Within the agent-based models, income was present in the pathway between employment and social hierarchy position (Pathway 3.1), represented by the 'employment' section of the pathway. The data used to represent this were model based income estimates for MSOAs, which was the estimated average weekly household income in each geographical area. This was the only pathway that directly assessed the impact of income. The results showed that this pathway did not have a significant effect on the outcome score of the model when added in either study area. This is somewhat surprising given the overwhelming support for the effects of income on tooth decay, as detailed above. While income seems to have been represented appropriately by its proxy variable, it may be that the binary nature of the NS-SEC variable representing 'social hierarchy position' caused the data to become too polarised in nature, possibly affecting the results for that pathway. The issues surrounding the NS-SEC variable will be discussed in further detail in the review of socio-economic position later in this chapter.

As outlined in Section 2.5, income is also an important marker of material circumstances, and is associated with the ownership of a number of important material possessions (Wilkinson, 1997; Galobardes et al, 2004). This has also been demonstrated with regard to oral health through measures such as the Townsend Deprivation Index (O'Hanlon et al, 1997), and through the presence or absence of material items and concepts such as inside toilets, housing construction material, car ownership, housing tenure, overcrowding, and piped water availability (Nicolau et al, 2005). Diet and oral health related resources could also be thought of as material circumstances impacted through income (Fonseca, 2012; Costa et al, 2012). Material circumstances was featured in the second domain of the theoretical framework, ‘healthy environments at home, work and play’, and primarily revolved around the material circumstances that individuals found themselves living in, and the associated financial constraints that occurred as a result. These in turn impacted on the ability of individuals and families to purchase certain types of food as well as access certain knowledge resources. Finding such variables to be important was hardly surprising, given the links between more general health inequalities and measures of social position (Wilkinson and Marmot, 2003), as well as the importance of diet for other health conditions and mortality (Kant, 2004).

Within the agent-based models, material circumstances were represented by house price statistics for MSOAs, taken from the Neighbourhood Statistics website. This represented the median house price value for each geographical area. This was seen as an appropriate proxy due its use as a marker of material standards in previous studies, which also identified this variable as an important indicator (Nkosi et al, 2011; Tunstall et al, 2013). Material circumstances was the starting point for three pathways including: a pathway between financial constraints, stress, and smoking as a coping mechanism (Pathway 2.1); through financial constraints which impacted on purchasing power and diet (Pathway 2.2); and finally through access to knowledge resources impacting on oral health habits (Pathway 2.3). None of the pathways involving material circumstances significantly changed the tooth decay scores in either study area within the statistical analysis when added. Again, this is surprising given the evidence for a link between tooth decay and material circumstances (and its associated variables) is fairly strong, and the operationalisation of the material circumstances construct was also appropriate. For example, there is a strong link between material circumstances and income (Costa et al, 2012), which has been shown to have strong associations with tooth decay within the literature (Geyer et al, 2010).

Education

Similar to income, education has also been shown to be a strong determinant of tooth decay. Education in early years has been shown to be a marker of inequalities in tooth decay scores (Muirhead and Marcenes, 2004), while the effects of education on inequalities in tooth decay have been shown to continue into adulthood (Brennan et al, 2007; Geyer et al, 2010; Mamai-Homata et al, 2012). A systematic review on the determinants of tooth decay in adults also found that 6 of the 9 studies related to education found similar associations (Costa et al, 2012).

Education has been hypothesised to be associated with differences in knowledge related to dental hygiene practices and the prevention of decay (Geyer et al, 2010). Education may also act as a mediator between tooth decay and socio-economic position, influencing income and access to preventative oral health measures, and has also been posited to influence health literacy and behaviours (Singh et al, 2013; Sabbah et al, 2015). Education was featured in the third theme within the theoretical framework used to guide this research ('Services provided, publicly or privately to support people in their daily lives'). The inclusion of this variable was a logical choice given its links to tooth decay as well as other health outcomes affected by the social determinants of health (Wilkinson and Marmot, 2003).

Education was present in two pathways within the agent-based models. The first of these was the pathway between education and social hierarchy positions (Pathway 3.2). In this pathway education was represented by whether an individual's highest qualification was above or below degree level. This pathway did not have a significant impact on the overall tooth decay scores in either study area, which was surprising given the importance placed on education within the oral health inequalities literature, as indicated in Section 2.5. Similar to the issue with the NS-SEC variable in the income related pathway, it may be that the binary nature of the education variable did not help in capturing the potential social gradient associated with this variable.

The second use of education within the agent-based models was in the pathway between education, dental knowledge and health damaging behaviours (Pathway 3.3). This time education was represented as point data within the models, representing the locations of further education facilities in the city, and those that offered apprenticeships. These locations were only present in the Sheffield East model, and were limited to use by those aged 16-24. As with the previous education based pathway, this was also shown not to have a significant influence on the outcome variable score in either study area. This is perhaps less surprising than the findings for the previous education based pathway, as education is often measured as an individual level variable rather than as location data, and only two education based facilities were present in the Sheffield East study area. This does, however, demonstrate the benefits of agent-based modelling in being able to test such scenarios. It may be that limiting the use of such facilities to one age band influenced the results. This decision was taken as this age group is the most likely to start apprenticeships (Department for Education, 2017), although there were also large numbers of starters from other, older age groups. However, given the results of the rest of the pathways, it seems unlikely that changing this would have had a significant impact on the outcome.

Employment

Costa et al. (2012) have shown that occupations of higher standing are linked with lower severity of tooth decay in adults, while a number of authors have demonstrated that unemployment is a risk factor for tooth decay (Tellez et al, 2006; Roberts-Thomson and Stewart, 2008). Parental occupation has also been shown to be a risk factor for tooth decay in children (Vanobberge et al, 2001; Gokhale and Nuvvula, 2016). Further, unemployment has been associated with oral health related behaviours, with lower employment status associated with less frequent use of dentists (Guiney et al, 2011), as well as xylitol use and the increase of behaviours such as smoking and drinking (Al-Sudani et al, 2016). Employment fell into the third theme of the framework of Macintyre and colleagues (2002) ('services provided, publicly or privately to support people in their daily lives'). Much like education, it is not surprising to see employment (or unemployment) emerging as an important variable within this theme, given its links to tooth decay development, as well as other health conditions through the social determinants of health (Wilkinson and Marmot, 2003).

Employment was included through the use of model based income estimates at the MSOA level, on the pathway between employment and social hierarchy position (Pathway 3.1). As detailed in the 'income' section above, this pathway did not have a significant effect on the outcome of the model for tooth decay scores in either study area. There seems to be enough evidence to suggest that employment is an important variable for tooth decay scores, despite some research finding this not to be the case (Julihn et al, 2006). Therefore, this variable would perhaps have been expected to have a significant impact on the outcome variable scores, particularly as it was represented by income, which also has strong associations with tooth decay (Celeste et al, 2009; Costa et al, 2012). As previously mentioned, the binary nature of the NS-SEC variable used to represent social hierarchy position may have caused the data to become too polarised, not allowing for the potential social gradient in this variable to be explored.

Socio-economic position

Watt and Sheiham (1999) have highlighted the importance of socio-economic status for inequalities in decay in both adulthood and childhood, while Schwendicke et al. (2015) have shown lower socio-economic positions to be associated with a higher risk of decay. Longitudinal research has also shown that differences in socio-economic inequalities in a number of oral health indicators (including tooth loss from caries) persist throughout the life course (Thomson et al, 2004). Hobdell et al. (2003) have also demonstrated socio-economic gradients in dental caries, implicating wider structural processes in such patterns. However, Costa et al. (2012) found sizeable variance in the results of studies assessing the impact of socio-economic status, with the inconsistency in variable classification mooted as a possible cause for these results. As with the previous two variables, socio-economic position featured in the third of Macintyre and Colleagues' (2002) framework themes ('services provided, publicly or privately to support people in their daily lives'), and was a logical addition to this research due to its links with decay and the social determinants of health (Wilkinson and Marmot, 2003).

Socio-economic status was featured in one of the theoretical pathways, between employment and social hierarchy position (Pathway 3.1). Within the agent-based models socio-economic position represented the 'social hierarchy position' construct, and was operationalised using the National Statistics Socio-Economic Classification (NS-SEC). This variable has been noted in the 'income' and 'employment' sections of this discussion chapter, as well as in previous chapters (See Section 5.2.7), due to the difficulty in using it in this analysis. This is primarily due to the non-hierarchical nature of the data, which would make it inappropriate to create arbitrary thresholds between groups in the middle of the data scale. Instead it was decided to use category 8 ('long term unemployed and never worked') as a reference group in contrast to the other seven categories (those that were employed in some way).

While this matches the way employment has been classified in a number of Dental Public Health studies (Tellez et al, 2006; Roberts-Thomson and Stewart, 2008; Costa et al, 2012), as well as in wider health inequalities literature (Wilkinson and Marmot, 2003), this also created a binary variable out of what is more than likely a far more complex indicator. As already reported, this pathway was not shown to have a significant influence on the overall outcome variable score in either study area. While it cannot be stated for certain that the classification of the NS-SEC data influenced this lack of statistically significant impact, Costa et al. (2012) noted that the different classifications for socio-economic positions, and the inconsistency associated with this, was a potential reason for the inconsistent findings of their review. The classification of the NS-SEC data may have had the same impact here, although it would have been difficult to avoid this approach without violating the structure of the data.

Psychological factors

Psychological factors and variables have been linked to tooth decay through a number of different mechanisms. Boyce et al. (2010) have demonstrated the biological effects of financial stresses and associated links to caries development in children. Other studies have also found links between psychological stressors, resources and self-reported oral health (Finlayson et al, 2010), as well as between self-reported oral health, personal constraints and chronic stress (Sanders and Spencer, 2005). Certain stresses have also been associated with the uptake of coping mechanisms such as smoking (Kassel et al 2003; Ng and Jeffrey, 2003; Siahpush and Carlin, 2005) which can be damaging to oral health. Psychological factors were featured in the second theme of Macintyre and colleagues' (2002) theoretical framework ('availability of healthy environments at home, work and play'), due to this variable's links with financial and material constraints, as well as in the fourth theme ('socio-cultural features of a neighbourhood') through its links with social capital.

Psychological factors formed part of two of the pathways included in the agent-based models. The first of these was the pathway between material circumstances, financial constraints, stress, and smoking (Pathway 2.1). The second pathway including psychological factors measured the interactions between social capital, stress and smoking (Pathway 4.6). Within both of these pathways the variable was represented by data on psychological discomfort ('psychological discomfort – self tense') from the ADHS. While it was not clear if this was a general variable on psychological discomfort or one related to oral health, it represented the only psychological stress based question in the ADHS. Again, these pathways were shown not to have a significant impact on the overall outcome variable scores when added to the agent-based models in either study area.

While the literature on psychological stress may not be as conclusive as the literature on variables such as income or education, it is still somewhat surprising to see this lack of statistical impact. The literature does tend to suggest an association between stress and tooth decay, however it may be that this is acting through a different mechanism, possibly biological (Boyce et al, 2010), which could not be captured in this analysis due to a lack of biology based responses in the ADHS. Ideally, although far from easy, future surveys or empirical work could consider the inclusion of such variables. It may be that the inclusion of smoking in the same pathway could explain the lack of significance. Smoking had a far more mixed literature than the other variables in Pathways 2.1 and 4.6, and thus could have contributed to the lack of overall significance of these pathways.

Coping mechanisms

Two coping mechanisms relevant to oral health were identified as potential risk factors within the literature. Excess alcohol consumption was shown to be a risk factor for tooth wear due to gastric acid regurgitation, as well as for dental trauma (Harris et al, 1997). Franceschi et al. (1999) also found that alcohol consumption increased the risk of oral cancers in each stratum of smokers. However, other research has shown that alcohol did not compromise the oral health of its participants (Harris et al, 1996). Due to the inconsistent findings, and the lack of links with dental decay, alcohol consumption was not included in the analysis.

Conversely, smoking was seen as a potentially more relevant risk factor for caries and tooth decay. Smoking can be a reaction to certain types of stresses in both adolescents and adults (Kassel et al, 2003; Ng and Jeffrey, 2003), and has also been associated with financial pressures (Siahpush and Carlin, 2005). There is debate in the literature as to whether smoking causes tooth decay, with some finding associations between the two (Hudson et al, 2007; Bernabe et al, 2014), while others reported no direct aetiological links (Reibel, 2003; Vellappally et al, 2007). There is still a suggestion that smoking may be a risk factor for decay however (Axelsson et al, 1998; Reibel, 2003), with chemical processes making teeth more susceptible to this (Vellappally et al, 2007).

Locker (1992) has also reported links with caries in older adults among other conditions such as periodontal disease. Winn (2001) has also posited that chewing tobacco may increase tooth decay through the high proportion of fermentable sugar found in these products. Smoking featured in the second theme of the theoretical framework ('healthy environments at home, work and play') as a coping mechanism for stress as a result of financial constraints, as well as the fourth theme ('socio-cultural features of a neighbourhood'), as a stress related response to social capital.

Smoking was present in two pathways within the agent-based modelling (2.1 and 4.6). The first of these (Pathway 2.1) was the pathway between material circumstances, financial constraints, stress and smoking. In this pathway, smoking was represented by data on whether an individual had been given advice on giving up smoking. This was preferred to another variable on smoking status, as the former was seen as a more appropriate indicator of having a problem with smoking. Pathway 2.1 was not found to have a significant impact on the outcome score of the model in either study area, although unlike some other themes tested in the model this was perhaps not a huge surprise. The evidence on the effects of smoking were hardly unanimous within the literature, and it may be more suited to investigations of periodontal disease (Pihlstrom et al, 2005). The limited data on smoking available in the ADHS meant that the construct may not have been operationalised as accurately as it could have been.

The second pathway (Pathway 4.6) involving smoking featured the pathway between social capital, stress and smoking. A similar pathway with stress and smoking was again present, with smoking again represented using data on whether an individual had been given advice about giving up smoking. Perhaps unsurprisingly, as with the first pathway including smoking, no significant change to the outcome variable score was found upon adding this pathway to the agent-based models, in either study area. Within this analysis it is hard to pinpoint which element of each pathway may have caused or not caused the observed results, however in this case it again seems likely that the inclusion of smoking and its less than ideal operationalisation may have been detrimental to this pathway.

Dental health habits

A number of dental health habits have been found to be important due to their effects on tooth decay. Dental behaviours of children have been linked to oral health knowledge of both parents and the child (Poutanen et al, 2006), while parental behaviours were also associated with the oral hygiene practices of children. Parental attitudes and efficacy towards oral health behaviours have also been shown to have a strong influence (Adair et al, 2004) on dental health practices in children. In the latter study deprived families were also shown to exhibit less confidence in being able to control such patterns. Links between education and income have also been found with regard to oral health related behaviours (Sabbah et al, 2009).

Socio-economic disparities in attitudes towards oral health have also been identified, with deprivation and education both associated with levels of dental knowledge and attitudes towards these (Williams et al, 2002). Singh et al. (2013) have also found educational gradients in the clustering of oral health behaviours, with Riley et al. (2006) reinforcing the importance of education for oral health related behaviours. Parents with higher dental knowledge have been shown to have children with lower levels of decay (Chu et al, 1999), and parental influence has been shown to be important in the development of decay in children (Hooley et al, 2012). Kumar et al. (2016) have also demonstrated the importance of regular brushing for incidence and increment of carious lesions. Despite this, Sanders et al. (2006) have disputed the idea that adults from less advantageous backgrounds care less about their oral health than those who are less deprived.

Dental health habits were featured in three of the themes as part of the theoretical framework used in this research. Firstly ‘healthy environments at home, work and play’, through financial constraints impacting upon oral health related knowledge (Williams et al, 2002; Singh et al, 2013). Secondly the ‘services provided, publicly or privately to support people in their daily lives’ theme, through dental knowledge acquired through education (Singh et al, 2013) or dental visits (Tickle et al, 2003). Finally, dental health habits were included in the ‘socio-cultural features of a neighbourhood’ theme, through collective oral health related attitudes and practices (Poutanen et al, 2006; Riley et al, 2006).

Dental health habits were present in six pathways within the agent-based modelling, by far the most of all of the constructs considered in this research. The first of these

(Pathway 2.3) was the pathway between material circumstances, financial constraints, purchasing power, diet and sugar intake. In this pathway dental health habits represented as ‘dental knowledge’ was operationalised by data on the fluoride intake of an individual. The second use of dental health habit related variables occurred in Pathway 3.3, between education, dental knowledge, health damaging behaviours and tooth decay. Dental health habits were again represented by the ‘dental knowledge’ construct, again in the form of data on fluoride intake. The third pathway using a habitual dental health variable was Pathway 3.5, in the pathway between dental service usage, dental knowledge, and tooth decay. Again dental knowledge was the construct used, with fluoride as its operationalised data source.

The fourth use of dental health habits came in Pathway 4.2, on the pathway between health behaviours, oral health habits and tooth decay. Oral health habits were represented by tooth brushing habits, which was operationalised with data from the ADHS on the frequency of tooth brushing of an individual each day. Both datasets for fluoride intake and tooth brushing levels were present for Pathway 4.3, a pathway running between health behaviours, dental attendance, dental knowledge, oral health habits, and tooth decay. Fluoride intake was used to operationalise dental knowledge, while tooth brushing frequency was again used to operationalise oral health habits. The final use of dental habit related variables was in Pathway 4.4, which was concerned with social capital and dental knowledge. As with previous pathways, fluoride intake was used to represent dental knowledge.

Given the emphasis on the importance of dental health habits, it was surprising to see that none of the pathways associated with variables related to this had a significant influence on the outcome variable score in either study area. Pathway 2.3 for example appeared to be entirely appropriate in its operationalisation (using fluoride level as a proxy for dental knowledge), as was also the case for Pathway 3.5. Pathway 3.3 was operationalised in the same way, however the use of sweet consumption as a proxy for health damaging behaviours in this pathway may not have been as appropriate, and could have affected the results of this pathway.

The oral health habits construct in Pathway 4.2 was represented using data on tooth brushing habits, which again would seem to provide a good fit between construct and variable, while the other variable in the pathway (‘health behaviours’, represented by ‘years of lost life’) was also deemed a good fit. The lack of significant impact was particularly surprising given how important tooth brushing is considered for good oral health (Kumar et al, 2016). Official guidance also suggests that teeth should be cleaned twice a day (NHS Choices, 2015), so creating a threshold in the agent-based models between those who cleaned their teeth twice a day, and those who cleaned their teeth less frequently than this also seemed logical. Similarly, the operationalisation of dental health habits in Pathways 4.3 (tooth brushing frequency for oral health habits, and fluoride level for dental knowledge) and 4.4 (fluoride level for dental knowledge) were also deemed a good fit.

Attendance at services

Dental attendance has been shown to vary by deprivation, with those from deprived areas being more likely to attend if symptomatic (Lang et al, 2008). Eckersley and Blinkhorn (2001) found similar trends regarding symptomatic visits, while also demonstrating a negative association between deprivation and attendance overall. Irregular attendance has also been related to adverse oral health outcomes, including dmft scores in 5-year-olds (Tickle et al, 1999) from deprived backgrounds, although more regular attenders sometimes have more fillings and treatment despite lower levels of decay (Sheiham et al, 1985). Other studies have suggested that service usage may increase with decreasing disease experience in children (Tickle et al, 2000). The importance of attendance is underlined by Tickle et al. (2003), who showed that previous treatment received was the most influential factor in parental healthcare preferences.

Further, Listl (2012) has demonstrated that a lack of dental attendance as a child can persist throughout the life course. Low income groups have been shown to experience considerable barriers to dental services (Wallace and Macentee, 2012), with some pre-existing socio-economic and demographic variables having higher associations with attendance (Guiney et al, 2011; Muirhead et al, 2009). Attendance has also been associated with oral health related behaviours (Hill et al, 2013). Donaldson et al. (2008) have also found gradients in oral health to be partially explained by dental attendance, which is also influenced by the effects of socio-economic status on barriers to regular visits.

Dental attendance fell under three separate themes from the theoretical framework used in this research, including: the second theme (‘availability of healthy environments at home, work and play’); the third theme (‘services provided, publicly or privately to support people in their daily lives’); and the fourth theme (‘socio-cultural features of a neighbourhood’). Dental attendance would also have formed part of the fifth theme (‘the reputation of an area’), due to the ability of dentists to set up where they liked before 2006 (Landes and Jardin, 2010), however this theme was not included in the final analysis, as it was essentially repeating a pathway from the third theme.

Attendance at services was included within four of the pathways tested in the agent-based models. Pathway 2.3 focused on the interactions between material circumstances, financial constraints, dental knowledge and healthy habits. In this pathway, attendance represented the healthy habits construct, and was operationalised by the ‘general dental attendance’ variable (which related to more general habits of attendance, rather than the exact frequency of visits). Pathway 3.5 was concerned with the interactions between dental attendance and acquired dental knowledge. This time attendance was represented by data on the frequency of attendance at a dentist. Pathway 4.3 involved the pathway between health behaviours, dental attendance, dental knowledge, and oral health habits, where dental attendance was again measured using the frequency of visits. Finally, Pathway 4.5 included the pathway between social capital, and healthy behavioural norms, with attendance (representing the healthy behavioural norms construct) represented by general habits of attendance, rather than the exact frequency.

None of the above pathways were shown to have a significant impact on the outcome variable score when added to the agent-based models in either study area. This was surprising given the importance of dental attendance stressed within the literature. All of the pathways appeared to be operationalised appropriately, as two slightly differing variables on dental attendance were used. As such it is difficult to see a methodological issue that may have affected the results of these. Conceptually it would however be wrong to assume that individuals would only attend dental services in their own, or surrounding neighbourhoods, or more specifically, the study areas used within this research. It could be that some would travel further from home to use other services for a variety of reasons; the Charles Clifford Dental Hospital, for example, if problems required particular attention. It may be that individual level attendance patterns may have more of an influence if a model of the whole city was conducted.

Further, previous research has concluded that the presence of dentists is not related to outcomes in decay (Aida et al, 2008). Conversely, there is plenty of evidence to suggest that dental attendance is beneficial to oral health (Montero et al, 2014), as well as evidence of less regular attendance being associated with higher levels of tooth decay (Tickle et al, 2000; Donaldson et al, 2008). Of course the pattern is more complicated than this, as more regular attendance can also be indicative of more decayed teeth, rather than good oral health habits (Sheiham, 1985). It may be that the variables taken from the ADHS do not offer a comprehensive enough coverage of the interactions that occur when visiting the dentist. For example, simply looking at frequency of attendance, or attendance related habits may not capture more complicated decisions such as choice of dentist and the location of this, which may be affected by other factors. Possible suggested reasons (other than proximity and location) include opening hours, treatments offered, available facilities, dentist reputation and manner, and treatment effectiveness, while factors such as patient age, charge paying status, dentist age and gender have all been shown to be associated with continued attendance at the same surgery (Lucarotti and Burke, 2015). Although one of the advantages of agent-based models is the ability to capture complex interactions, programming a list of interactions such as those listed above into a model remains a challenge.

Water fluoridation

Water fluoridation has been shown to have the potential to be an extremely effective Dental Public Health initiative. The World Health Organisation has recognised its potential to reduce caries in both adults and children (WHO, 2003). Systematic reviews on the topic have also shown its potential to reduce decay in children (McDonagh et al, 2000; Petersen and Lennon, 2004; Yeung, 2008), while also acknowledging mild side effects such as dental fluorosis. Other studies have shown the benefits of such initiatives for adults (Griffin et al, 2007). Fluoridation has also been shown to have the potential to reduce the social gradient in decay in UK cities (McGrady et al, 2012), with the work of Slade et al. (1996) on children with and without access to fluoridated water in Australia drawing similar conclusions. Within the UK the distribution of fluoridated water sources is far from universal, with around 6,000,000 people having access to fluoridated water of some kind, of which only 300,000 have access to naturally fluoridated water as of 2012 (British Fluoridation Society, 2012).

Within the theoretical framework of Macintyre and colleagues (2002) this variable would have been the sole relevant variable included within the ‘physical features of the environment shared by all residents in a locality’ theme, for which the authors used examples such as water quality and air pollution. However, the city of Sheffield is not fluoridated either naturally or artificially (bar small areas to the southeast of the city), and neither were the two study areas used in this research. Thus, neither the physical environment theme within the framework, or water fluoridation were featured in the agent-based models. This variable is potentially extremely important to Dental Public Health however, and should be the subject of future simulation modelling, where relevant.

Social capital

Social capital has been linked to children’s use of dental services through a variety of mechanisms, including diffusion of knowledge of health promotion, maintaining healthy behavioural norms, promoting access to services and amenities, and psychosocial support (Lida and Rozier, 2013). Contextual level social capital has been shown to be beneficial to the self-rated oral health of elderly citizens (Aida et al, 2011), while community centres have also been shown to be significantly related to dmft scores, potentially through chances to increase social cohesion and solidarity (Aida et al, 2008). Tellez et al. (2006) also found associations between social institutions and better oral health. Neighbourhood level caries variation has also been related to perceived levels of empowerment in adults, although not to individual level social capital (Santiago et al, 2014). Conversely, caries has also been shown to be higher amongst those living alone (Avlund et al, 2003).

While not all studies of social capital and tooth decay have found links between the two (Mathur et al, 2016), Rouxel et al. (2014) concluded that, despite the underdeveloped literature and mixed results of their review, social capital appeared to be a potential determinant of oral health. Social capital was featured in the fourth theme in the theoretical framework (‘socio-cultural features of a neighbourhood’), given its links to decay through social organisations (Tellez et al, 2006; Aida et al, 2011), as well as healthy behavioural norms and psychological effects (Lida and Rozier, 2013).

Within the agent-based models social capital featured in three pathways. As (contextual level) social capital formed the start of each of these pathways it was operationalised the same way each time, this coming in the form of crime data per LSOA from the 2015 Indices of Multiple Deprivation. This variable included information on violence, burglary, theft and criminal damage per 1000 residents. The first pathway to include social capital was Pathway 4.4, which concentrated on the interactions between social capital and dental knowledge, while Pathway 4.5 was concerned with social capital and healthy behavioural norms. Finally, Pathway 4.6 focused on the relationship between social capital, stress and smoking. None of these pathways proved to have a significant impact on the overall score in the model in either Sheffield East or Sheffield West, which was somewhat surprising given the evidence for a link between various indicators of social capital and tooth decay within the literature.

This brings into question the operationalisation of the data. Crime data was seen as an appropriate source for operationalising social capital, given the use of similar data in a previous oral health related study (Pattussi et al, 2001). A lack of other appropriate neighbourhood level data sources meant that this was also the only realistic option for this variable. Pathway 4.6 may have been affected by the issues associated with the 'stress -> smoking' pathway which were discussed earlier. The use of fluoride level and dental attendance in Pathways 4.4 and 4.5 respectively would seem to be appropriate however, so it is difficult to see why these pathways did not have a significant impact on the outcome variable score. This is an emerging theme within this analysis however, where certain combinations of behaviours were shown to be significant, while other behaviours were not.

It is worth remembering however that social capital is a difficult concept to include in quantitative analysis. There are numerous definitions of social capital (Celeste et al, 2009), with no agreed upon preference among these. Crime was used as a proxy for social capital in this case, representing the idea of community cohesiveness, which fits with a previous study of oral health and social capital that used homicide rate to act as a proxy (Pattussi et al, 2001). However, it may be that this is not accurate enough to capture the nuanced nature of social capital. Crime scores may not reflect the complicated processes that occur as part of this measure, and the ADHS did not have enough relevant variables to supplement this.

Access to services

Some authors have hypothesised that attendance rates may not be helped by some areas being underserved by dentists (Lang et al, 2008). Historically there has been an uneven distribution of dentists in the UK (Cook and Walker, 1967), the evidence for which has been added to in more recent studies (Jones et al, 2001), which may not have been helped by practices being free to set up wherever they wished before 2006 (Landes and Jardin, 2010). However, Macintyre et al. (2008) found no evidence to suggest an uneven distribution of dentists in Glasgow, UK. Some studies have also suggested that certain regions of the UK, such as the south of England, have difficulty accessing NHS dentistry in particular due to the rise in demand for private dentistry in these areas (Hancock et al, 1999; McGrath et al, 2001).

Within the agent-based models the geographical distribution of dental services was covered by the point data used to model the locations of the dental services. The pathway associated with access to dental services (See Figure 7) was deemed to be too similar to Pathway 3.5 (dental service usage -> associated benefits/knowledge), and that running this would give an undue additional weighting to this effect, which might have adversely impacted the results of the model. Whilst this seems unlikely now given the nature of the results, logic at the time dictated that this pathway should not be included.

6.4. Comparison of the two study areas

The decision was made to test the theoretical pathways in two study areas within Sheffield, to see if different pathways had different impacts in the two geographical locations. The two study areas were chosen based on clusters of LSOAs with some of the highest and lowest tooth decay scores. Thus the study areas differed with respect to mean tooth decay scores, which also happened to coincide with considerable differences in a number of demographic and socio-economic variables.

One of the most interesting trends from the analysis was the consistency of the results discussed above across the two study areas. Both study areas saw their only significant change in the outcome variable score occur upon the introduction of the pathway between shops/supermarkets, diet, and health damaging behaviours (Pathway 3.4). One explanation could be that the locations of shops and supermarkets and the interactions that occur with sugar as a result of this are not mediated by common socio-demographic characteristics. This is, of course, a potential explanation for the two study areas in this research, and caution is urged regarding their generalisability to the UK as a whole. This explanation would not negate the importance of socio-economic standing or better oral health behaviours and attitudes, but rather could show that problems related to tooth decay come from similar sources in both of this study's populations. Further, and as previously mentioned earlier in this discussion chapter (see Section 6.2.2), despite the theoretical pathways having very similar effects in both study areas, the second simulation experiment run as part of this research (adding an extra shop to each area) showed interesting differences between the two study areas. An extra shop was shown to lead to a statistically significant reduction in decay in the Sheffield West study area, while the pattern in Sheffield East remained relatively unchanged. This could indicate that opportunities within the local food environment in Sheffield West to reduce decay are influenced by both food availability and the number of shops, and in general are greater than in Sheffield East, where it may be that the types of food sold are more important to oral health.

Given the differences in demographic and socio-economic variables between the areas (presented in Table 16) that are also known to be associated with the social determinants of health (Wilkinson and Marmot, 2003), it perhaps would have been expected that pathways might have different influences and driving mechanisms between the two areas. Clear social gradients have been demonstrated with regard to oral health outcomes, with Watt (2012) stating that 'oral diseases are directly related to socioeconomic position in a step-wise graded fashion' (p.44). Similar patterns have also emerged for tooth decay (Hobdell et al, 2003). Watt and Sheiham (1999) reported social class to be a determinant of tooth decay, and the closest equivalent to this variable in this research, the NS-SEC classification, showed marked differences across the two study areas. Other indicators recognised as social determinants (Wilkinson and Marmot, 2003) including unemployment (Roberts-Thomson and Stewart, 2008), education (Brennan et al, 2007), and material circumstances (Nicolau et al, 2005) have also been linked with decay. All three of these variables (car ownership in the case of material circumstances, and degree ownership in the case of education) showed considerable differences between the two study areas. Given the trends from previous literature that has shown that while context is important, individual characteristics may be more so (Curtis and Rees-Jones, 1998; Macintyre et al, 2002), it is even more puzzling that more differences were not seen between two samples with very different socio-economic characteristics with regard to the pathways.

A potential explanation for the lack of expected differences could be seen in the results of Table 17. Watt and Sheiham (1999) have previously cited sugar consumption as playing a key role in causing decay, with the use of fluoridated dentifrice having the opposite effect. As can be seen in Table 17, the differences between the two study areas for variables such as using toothbrushes and paste, fluoride intake, high sugar intake, and both sweet and cake consumption are minimal. Even the dental attendance variables are not as markedly different as would be expected from reading the literature. With minimal differences in these variables, it may be that diet mediated by shops can impact on tooth decay regardless of socio-economic position, however those from more advantageous backgrounds may reap the protective effects of their prior socio-economic position, including lower numbers of decayed teeth (although intriguingly the final outcome variable scores for Sheffield West were higher than for Sheffield East).

Despite the overall mechanisms from the pathways appearing to be similar between the two study areas, there were some interesting, if not statistically significant differences between them. Firstly, as can be seen in Figure 57 (Section 5.3.2), the trend in Sheffield West was steady until the introduction of Pathway 3.4, then was fairly steady again after the increase that accompanied the introduction of this pathway. Sheffield East on the other hand displayed a slightly different trend (Figure 56, Section 5.3.2). Here it can be seen that after the increase brought about by the introduction of Pathway 3.4, there was a subsequent decrease in the tooth decay score upon the implementation of three further pathways: Pathway 4.2, the pathway between health behaviours and oral health habits; Pathway 4.3, the pathway between health behaviours, dental attendance, dental knowledge, and dental health habits; and finally Pathway 4.4, the pathway between social capital and dental knowledge. While these pathways did not have a significant impact on the outcome variable score, they represented the only other notable change in the trend line beyond the introduction of the pathway between shops, diet/sugar intake and health damaging behaviours (Pathway 3.4).

The themes associated with these pathways were health behaviours (operationalised as ‘years of lost life’ from the IMD, 2015), oral health habits (operationalised as tooth brushing frequency), dental attendance, dental knowledge (operationalised as fluoride consumption), and social capital. It would seem unlikely that dental health habits (tooth brushing, dental attendance, and fluoride intake) would benefit those from less socio-economically advantaged populations, given the extensive literature around theories such as fundamental causes (Link and Phelan, 1995), as well as neo-materialist theories (Bissell and Peacock, 2011) that posit that those in more advantageous social positions tend to benefit most from health related resources and interventions. It has also been shown that those in more advantageous positions are more likely to brush teeth more frequently and use other beneficial oral hygiene products (Watt and Sheiham, 1999). In addition, while levels of decay have decreased inequalities in tooth decay persist (Watt, 2007), suggested that these behaviours may not hold the key to closing the gap in tooth decay.

It may be that social capital and the idea of community cohesion has more of an influence on health in areas with lower incomes. This idea is supported by previous work that has found high levels of social cohesion modified psychological distress in lower socio-economic communities (Erdem et al, 2015), as well as mental health outcomes in low income groups, with health-related behaviours and improvements in these being one hypothesised pathway in the case of mental health outcomes (Fone et al, 2007).

6.5. Strengths and limitations of the research

6.5.1 Conceptual strengths

A major conceptual strength of the approach taken in this research was the flexible nature of the theoretical framework, which allowed for a variety of different interactions and pathways to be included. This was after all the original intention of Macintyre and colleagues (2002). The carefully thought out themes from the original framework meant that all relevant aspects of neighbourhoods appeared to be covered, or could be accommodated within the custom pathways. The pathways also allowed for the focus on both neighbourhood based features (both built and physical) as well as the collective social actions of individuals within groups in the study areas. This gave the theoretical side of the work balance, as it was possible to look at the interactions between individual and neighbourhood level variables. This fits with previous literature on neighbourhoods and their effects on health, which suggests there is reason to believe neighbourhood level variables have some influence, but that individual level characteristics are as, if not more, influential (Curtis and Rees-Jones, 1998). The approach of the framework allowed for this to be taken into consideration.

The methods used in this research also managed to strike a conceptual balance between the conventional and relational views of neighbourhoods described in Section 2.2. Conventional administrative boundaries were used to define the study areas used for the agent-based modelling, and were kept to the small area level (LSOAs) in line with people feeling that their neighbourhoods were represented by the more immediate areas surrounding their house (Haynes et al, 2007). However, LSOAs are still big enough to include local facilities and features of the built environment that were important to this research, such as dentists and shops. The clustering of the LSOAs also allowed for a variety of these features to be included over a larger geographical area, in an attempt to include a more relational approach to the research (i.e. people not being restricted to features in their immediate surroundings). While a difficult balance to manage, this approach demonstrates how a certain level of compromise between these two theoretical approaches can be achieved.

6.5.2. Methodological strengths

The flexibility of the spatial microsimulation modelling in creating a representative population with operationalised data from the theoretical pathways represents one of the major strengths of this work. The ability to combine Census variables with those of a behavioural or attitudinal nature from survey data is a novel approach, and one that is still underused in health inequalities research, having never been used before within Dental Public Health. It allows researchers to create custom datasets to fit their research, using a reliable population synthesis technique, with the creation of new target variables that would not have been available in such detail previously. This is particularly beneficial when dealing with theoretical frameworks that need operationalising and populating with variables. This technique could be particularly useful in a discipline such as Dental Public Health that often uses surveys and questionnaires to collect data, offering a potentially cheaper and quicker alternative to gathering population data. The creation of the microdata also allowed for a more detailed picture to be painted of the individuals and households involved in the analysis (Ballas et al, 2005a) through both statistical and spatial analysis, while the consistency of the deterministic method (IPF) resulted in consistent results each time the model was run. While the statistical nature of spatial microsimulation models has been seen as a negative at times when compared to the interactivity of other methods, Wu et al. (2008) demonstrate that it is also a positive, due to the presence of 'important statistical mechanisms that ensure the similarity of what it predicts and what is actually observed in the gathered data' (p.446).

The use of open source software such as R to conduct the modelling was an additional benefit, as the packages associated with this software allow for the management, manipulation and analysis of data in the same programming environment, as well as offering links to other open source software packages. Although not used in this research, the 'RNetlogo' extension (Thiele and Grimm, 2010), a way of linking the simulation capabilities of NetLogo with the statistical controls of R, is a good example of this. The output of these models are also highly flexible, and could be imported into traditional databases such as Excel of SPSS, as well as geographical software to be mapped, allowing for a spatial consideration of disease in a population at a spatial scale previously not available. Nowadays spatial microsimulation is more readily accessible than ever, due to attempts to make the process more efficient for new and existing users. For example, Jones et al. (2016b) have scripted a package for easy spatial microsimulation modelling in R ('rakeR'), while Tomintz et al. (2016) have created an interactive website (www.simsalud.org) which is designed to be easy to use, and is specifically aimed at non-programmers.

The data from the spatial microsimulation model formed the backbone of the research, as it provided a representative synthetic population that could be used as the base population in the agent-based models. The use of the agent-based models allowed for the inclusion of dynamic and interactive simulation modelling, which added a new dimension to quantitative analyses of oral health inequalities. In particular, these features give the models an advantage over traditional statistical techniques, which are less well suited to testing the types of interactions and mechanisms representing behaviours (Baker and Gibson, 2014). The 'low threshold, no ceiling' (Tisue and Wilensky, 2004) nature of NetLogo allowed the theoretical pathways to be programmatically added to help set the modelling on a theoretically relevant course. A further advantage of the use of agent-based models was the possibility to test the significance of a variety of potential relevant pathways where the literature was less conclusive, which could be discarded if not relevant for further research. The flexibility of such a novel method makes it a highly useful and adaptable tool.

6.5.3. Conceptual limitations

As with any research, there are also a number of limitations that need to be addressed. Concerning the conceptual basis of this research, it is important to acknowledge that while the work is based on theory from relevant literature, the lack of neighbourhood based analysis within the Dental Public Health field may have hampered this to some degree. While theoretical concepts were also taken from the more general health inequalities literature, an underdeveloped dental literature on the subject of neighbourhood effects may have led to a slightly less developed framework. Curtis and Rees-Jones (1998) stated that the way places interact with health was under theorised when they published their work, with different elements of the body of theory either not integrated well, or still relatively underdeveloped. This is still true of the Dental Public Health field today, and it is hoped that this research can contribute to this area, aiding similar studies in the future.

The way some pathways were constructed may also have affected their outcomes, if say, the effect of a more influential variable within a certain pathway was negated by one that was less influential. The marrying up of the theoretical concepts with data to represent them was also at times not as accurate as would have been liked, and this will be discussed in more detail in Section 6.5.4. Additionally, despite the deconstruction of the pathways being carried out in a thorough manner, and all of the singular pathways being present in the simulation at some point, it may be that this deconstruction from their original format (Figures 3-7, Section 3.5) caused some interactions between variables in nearby pathways to be left out.

In line with this, there were also a number of variables that were not included in the final analysis for a variety of reasons. Race and ethnicity were shown to be an important predictor of tooth decay in the literature review, however they did not form part of a pathway. This was due to both conceptual and practical issues. Firstly, from a conceptual point of view, no studies have laid out the exact mechanisms by which a person's race and/or ethnicity leads to tooth decay, which given the number of different races and ethnicities that would need to be modelled, is a limitation. It would not have been appropriate to start grouping ethnicity or race as 'white' versus 'non-white', as the situation is likely more complicated than that (Delgado-Angulo et al, 2015). From a practical point of view, this variable was not found to be statistically significant when judging potential constraint variables for the spatial microsimulation, leading to its exclusion. Additionally, trying to include all ethnicities and/or races would be difficult within the simulation modelling given the number of different categories, and would likely make the models slower due to the extra calculations and interactions. This is in no way meant to diminish the importance of these variables, and future research should look to include them where possible.

Similarly, variables such as age and gender, while again proving to be important predictors and being included in the spatial microsimulation modelling, did not fit explicitly within the remit of any of the pathways in Macintyre and colleagues' (2002) framework, and were thus not included in the pathways. Further, an increased number of age brackets would have added additional processing time to the models. This represents the types of trade-offs that can occur in such research, involving a balancing act between accuracy and speed (Shendarkar et al, 2008). Locations such as community centres and churches should also be considered in future research, as these came to light as potentially important locations after the initial models had been designed. A great amount of detail was covered in the pathways, however not everything could be included, as they either did not fit explicitly within the framework, or could not be operationalised as easily.

6.5.4. Methodological limitations

With regard to operationalising the theory in this research, the ADHS contained barely enough variables to cover the theoretical framework that had been constructed. Suitable proxies were found for each of the constructs in the pathways, however it could still be argued that more suitable variables exist, and may be found in other surveys. Due to the methodological difficulties surrounding the inclusion of more than one survey data source in the microsimulation, this issue was not addressed in this research. This also led to some of the theoretical pathways being simplified with steps taken out of a number of these. This iss a great shame as there are many excellent data sets available, yet there is also often not enough overlapping of variables between them, or themes covered in enough detail to form one comprehensive dataset. For example, the ADHS is excellent in its coverage of clinical data and a number of behavioural variables, but lacks data on types of food and nutrients that are consumed beyond sugars. The Understanding Society dataset has a much more extensive range of food consumption variables, yet lacks the number of responses related to oral health that the ADHS has.

While a shame, it is probably not surprising given that each survey will have a particular focus in mind, and even if a survey was created for a more general purpose, it would be near impossible to cover every socio-economic, demographic and health related variable in enough ways to satisfy an entire research field. This leads to a familiar position where, despite theoretical work attempting to overcome limitations seen in quantitative work, theoretical work can to some degree be limited by data availability. This is however still a preferable approach compared to designing a study based firstly on data choices, with theory as an afterthought.

This issue of data availability also affected the spatial microsimulation modelling in much the same way. While the models offer a flexible approach to combining surveys and Census data, they are still limited by the data that is available to them. Additionally, the size of the ADHS sample may also have affected the results of the spatial microsimulation modelling. Due to deleting missing data or NAs (which spatial microsimulation cannot process) the sample size fell from 11,380 individuals to 4,840. While a sizeable cohort, this was a large reduction from the original sample, and potentially much smaller than remaining sample sizes that may be left from other surveys. Conceptually, with a smaller pool of individuals and a potential reduction in the variety of characteristics in this pool, it may be harder for the method to recreate populations, and associated target variables as accurately. This theory is supported by the work of Ryan et al. (2009), who found that 'as input sample size increases, resulting populations experience gains in accuracy' (p.201). It must also be remembered that 'spatial microsimulation outputs are model estimates (the reliability of which depends on a wide range of factors) and not findings' (Ballas et al, 2005a - p. 14), and are not suitable for predicting variables influenced by local or external factors (i.e. transport, public transport services, or the presence of large universities or single employers).

One method that could be useful in both combining survey data and expanding sample sizes is data linkage, which Slack-Smith (2012) describes as 'using probabilistic matching of identifiers such as name, address, date of birth, sex and medical record number' (p.91). Slack-Smith (2012) has stated that data linkage could be useful for studying population trends in oral health outcomes over time, as well as potentially investigating causal pathways and relationships between different factors. Such work is still relatively rare within the field of Dental Public Health, yet has considerable potential and clearly shares some common ground with spatial microsimulation modelling.

On a similar note, techniques such as statistical matching could offer an alternative approach. This method works by matching records from one source to those in a second source using common identifiers, which allows information on the same units to be identified in both sources. For example, if one dataset contains variables X and Y, and a second source contains variables Y and Z, statistical matching aims to combine these sources, leading to files containing X, Y and Z (Moriarity and Scheuren, 2001). However the cases in the two files have little overlap, or are not available due to the potential to identify individuals, so generally records are matched with those from a second source that does not represent the same unit, but rather a similar one (Rodgers, 1984). Previous spatial microsimulation models have successfully combined multiple survey data sources (Edwards and Clarke, 2009), however research papers often include little detail and practical advice regarding how to conduct such methods.

A more practical concern with the research was the computational intensity of the agent-based models. Despite the use of the University of Sheffield's supercomputer, the models had to be scaled down to a certain degree in order for multiple iterations of these to be run in time. In line with this, issues with computing power, time, and technical difficulties with code meant that a number of intended features for the agent-based models had to be left out. These included: social interactions between agents; family structures within the spatial microsimulation dataset (as opposed to a population of individuals); differentiation of agent movement with regard to speed, as well as potentially including different transportation options; and finally adding demographic shifts and life cycles to the models. The latter was included in Crooks's (2008) exploratory update of Schelling's (1971) segregation model, where agents were removed when reaching certain ages, and new agents introduced and given a random social class and age. This would be a valuable addition in future work.

Further, while the manual calibration of the agent-based models worked in this research, for future research or potentially larger projects, more efficient parameter setting could be investigated through the use of genetic algorithms, and the inclusion of differential equations to determine agent rules. It is worth repeating earlier caution in the use of such methods however, as if the parameters created do not match or adhere to the data values of the survey data there is a risk that the theoretical side of the research could be compromised, negating the point of doing it. Such concerns could be alleviated through using approaches similar to Heppenstall et al. (2007), whose work on agent-based models of petrol markets revealed that GA parameters were a close match to the original parameters chosen based on analysis of the market, and that the GA parameters produced better estimates when considering the spatial trends in pricing. Additionally, while this exploratory research was more concerned with the trends produced by the outputs of the agent-based models, the tooth decay scores produced as part of this were not as accurate as would have been desired. Future research should look to fine tune these, so that more accurate predictions of specific scores may be obtained, rather than just the analysis of trends.

Finally, despite efforts to include relational geographies in this research, the use of geographical boundaries from the Census for the spatial units may still limit the inferences that can be made. For example, while multiple facilities were included within each study area, there is no guarantee that the study sample would not use facilities in other parts of the city. This could have been accounted for by modelling the whole city; however, this would have been far too computationally intensive. Even then there is no guarantee that individuals would not use services in other nearby towns such as Rotherham or Chesterfield. Perhaps most important to remember is that the models were only tested on two relatively small areas within one city, on a total of just over 17,000 individuals. It is highly unlikely that these populations are representative of the overall population of the UK, given the socio-demographic statistics presented for each study area in Table 16. Therefore, the generalisability of the findings in this study to other urban areas within the UK should be interpreted with caution.

6.5.5. Modelling assumptions

There are a number of other considerations that should be taken into account when discussing the results of the model, including the assumptions made in building the agent-based models. These are not weaknesses as such, however these decisions influenced the structure of the models themselves, and so are worthy of further reflection. For example, Johnson and Groff (2014) have commented that ‘similar to other types of modeling, findings are constrained by the assumptions and rules that underpin the model’ (p.514). There are six assumptions that will be discussed, including: the use of mean scores as thresholds; the changes in agent characteristics over the course of the model; setting the models to mirror the accumulation model of health inequalities; the changes to tooth decay scores over the course of the model; the timescale over which the model was run; and finally, the probabilities used to drive the models.

Firstly, throughout the agent-based modelling thresholds were used to differentiate between groups, and the subsequent rules that were applied to them. On occasions the mean score of a variable was used as this threshold. This was usually in the case of the data representing the neighbourhood based variables, such as house price data (material circumstances), crime scores (social capital) and years of lost life (health behaviours). Education was also used to differentiate certain agent actions. These cut-off points may have been too crude at times, and perhaps breaking down variables into more distinct classifications would have helped capture more of the nuances of a potential social gradient. This decision was partly taken for practical reasons however, as breaking data down into quintiles, or even smaller categories, would have increased the computing time and complexity of the models. Conversely, data from the ADHS was split based on associated data scales and use of literature, to keep values in line with real world data.

The second point regards the changing of agent characteristics as the agent-based models progressed. This was achieved through a score of 0.01 being either added to or subtracted from the relevant variable in question (i.e. tooth brushing frequency) depending on the rules for that agent. For example, if an agent was presumed to have higher sugar consumption, their sugar score would have 0.01 subtracted from it, in order to bring it in line with the real world value for high sugar consumption ('1'). Whether this was an accurate reflection of the importance of each of the variables and mechanisms is open to debate. It may be that some variables require higher 'weightings' than others to emphasise their importance, which may affect the outcomes of certain pathways and the model in general. An approach similar to that taken by Dibben et al. (2007) could be used to address this issue. Their research used survey data and regression analysis to assign alternative weightings to the domains of the Indices of Multiple Deprivation, and is an approach that could easily be replicated for the data in the ADHS. This may help to confirm whether some variables should be given a weighting higher than others in such modelling, and is possibly an area for future research.

A third assumption concerns the way that the agent-based models were set up to mimic the accumulation model of health inequalities (Poulton et al, 2002). The premise behind this was that adverse exposures increase risk and have an additive effect over time, which leads to longer term damage later in life. These exposures could be due to a mixture of various individual, environmental and structural elements (Sisson, 2007). Some processes and the subsequent occurrence of diseases or conditions will take longer than others, and this was allowed for within the model (see Table 31, Section 5.2.7). For example, the effects of lower education were presumed to have a longer term, and less immediate every day impact than the consumption of sugar, or not being able to afford certain food types. This was in line with previous research which found that the negative effects of unemployment on the health of an individual did not occur after one day, but rather over a longer term period of being unemployed (Mitchell et al, 2002). Previous research has also suggested that social changes may have more immediate impacts on health, compared to the longer term impacts and changes to health associated with growing inequality (Adler and Newman, 2002). Additionally, Macintyre et al. (2002) have posited that plausible time intervals during which neighbourhood environments may influence health should be included in research hypotheses, as a lack of time lag between exposure and effect is unlikely to be accurate.

The use of the accumulation model was a good fit for the approach taken within the analysis, as the addition or subtraction of 0.01 from individual characteristics matched the idea of additive exposures (or lack thereof) over time. This is in addition to the fact that the critical periods model of health inequalities would have been harder to implement programmatically. Thresholds were added to the simulations so that once an agent reached a certain critical point, their behaviour may change, or they would feel a different effect as a result of their behaviour. Conceptually this can be thought of as agents building an accumulation of certain conditions over time, before the effects take place once a critical threshold has been crossed at a certain point in time. This approach governed the way agent characteristics interacted and changed over the course of the model, so while it is highly unlikely that this influenced the overall trend seen in the analysis, it will have influenced the workings of the models that lead to them. Such approaches have been used more commonly in agent-based models of migration, with thresholds in place for events such as population or resources in a community declining sharply, or income falling below a certain level (Klabunde and Willekens, 2016). Upon reaching these thresholds an agent would migrate, possibly based on 'pull factors' such as socio-economic or environmental conditions (Hassani-Mahmooei and Parris, 2012).

The fourth assumption concerned the change in tooth decay scores in the model, which will now be discussed. Tooth decay scores were subject to the same addition or subtraction of 0.01 as the other agent characteristics. Of course it is not possible for people to have their tooth decay (or other variable scores) change by +/0.01, as this is not the way tooth decay is measured in the real world, and also because it is not possible for tooth decay scores to decrease once decay is present. This approach was taken to keep certain variables from becoming too large. As there were no births, deaths and more general demographic interactions within the model, the subtraction of values from certain characteristics was used to compensate for the lack of these demographic processes. This is because over a two-year period it would be expected that a certain amount of population turnover would take place, including births, deaths, individuals becoming 16 years old and hence becoming part of the study population, and individuals moving in and out of the study areas. As the inclusion of these dynamics were beyond the scope of the research, the decrease in tooth decay scores for some was an attempt to represent population change over time for those with better teeth. Conceptually, tooth decay scores are unlikely to increase in every member of a population over the two-year period, thus the subtraction acted as a way of keeping these scores in check, in lieu of the demographic processes that would normally occur. Previous agent-based models have included population turnover mechanisms, including Crooks (2008), who added deaths and births (with random locations and social classes) in his research updating Schelling's (1971) residential segregation models. Geard et al. (2013) have also showed how population turnover can be replicated using probabilities based on the base characteristics of the individuals in the simulation, while Kim et al. (2014) have described how national level survey data can be used to help define demographic transition rates.

The fifth assumption to be discussed regards the timescale over which the agent-based models were run. Based on clinical advice a time period of 2-3 years was identified as being appropriate for studying the process of tooth decay (from the inception of the disease to it showing visible signs). This was shortened to a 2-year period, based on the computational intensity of running the models. Each 'tick' in the model represented one day, which again was mainly due to the intensity of running the models, and the number of different iterations of each model that needed running. This prevented the use of shorter time spans such as hours or minutes. While the results of the models seemed to move very firmly in one direction, and it seems unlikely that a different timescale would have affected the overall trends seen in this research, the choice to use days over smaller time units will still have played a role in shaping the running of the model.

Finally, the last assumption to be addressed is that of the probabilities used the drive the agent movement in the simulations. Due to the lack of agent-based modelling conducted within the Dental Public Health field, combined with a lack of official data sources from which to make estimates, there was little on which to base the probabilities of visiting certain locations within the simulations, and defining these was a difficult task. Trends were therefore taken from the literature to guide the assignment of probabilities to make them theoretically relevant, however this still did not help with regard to assigning exact numbers to these probabilities. Therefore, the quantification of these values was to some degree subjective, although always with the literature in mind. Thus while the values themselves may not be exact in their nature, they are based on theoretically relevant trends and logic.

For example, as mentioned in Section 6.2.2, shops are likely to be the most visited of the three locations, so in both study areas these locations had the highest probabilities of being visited. It was then difficult to judge the frequency of visiting further education facilities and dentists, so both were given equal weightings in each area, rather than picking one over another without reasoning to do so. This approach was as logical as it could be given the uncertainty surrounding these parameters. The higher weightings given to shops likely did influence the simulations in some way, but it is hard to argue with the logic behind this decision. Even tweaking the probabilities in each study area would seem unlikely to affect the overall trends seen in this research. The use of education as a differentiator between the probabilities of groups in each area may have been too crude (being binary in nature), but was preferable to using the troublesome NS-SEC variable. Education was also a theoretically relevant variable to use as a differentiator, given its links to oral health related behaviours (Chu et al, 1999; Williams et al, 2002; Singh et al, 2013) and dental attendance (Riley et al, 2006). The binary nature of the variable may have influenced the workings of the model however, and likely would not have reflected any potential social gradient in oral health related behaviours.

6.6. Summary

This chapter has highlighted some surprising results from the agent-based modelling, namely that of all the potential determinants of tooth decay identified in Section 2.5, only one pathway containing diet and the locations of supermarkets had a significant impact on the outcome variable scores. This impact was the same across the two study areas, however this differed when a new shop was introduced in a separate simulation, suggesting different mechanisms through which diet and supermarket location may influence tooth decay in each area. Despite the advantages of studying tooth decay through combining theoretical work with two separate simulation methods, the findings of this research should be interpreted with caution due to the limitations listed in Sections 6.5.3 and 6.5.4, as well as the potential lack of generalisability of the results to other areas of the UK.

Chapter 7 - Conclusions and recommendations

7.1. Introduction

The following chapter will briefly discuss the overall findings of this thesis in relation to the original aim and objectives of the research, set out in Section 1.2. The main conclusions of the research will then be presented, before the implications of this work for future research and policy is discussed.

7.2. Summary of research findings

The overall aim of this research was to conduct an exploratory analysis to investigate how neighbourhood effects may influence patterns of spatial inequalities in tooth decay. In order to achieve this aim, three research objectives were set. These objectives will now be discussed in relation to the main findings from this research.

- To identify theoretical pathways by which neighbourhoods influence tooth decay

The above objective was achieved through a thorough search of the Dental Public Health literature on the subject of tooth decay. Along with literature from the Dental Public Health field, theory from the wider health inequalities field was also included in order to 'fill gaps' in the dentistry based theory, and to give guidance on wider patterns associated with certain themes. This literature was mapped onto a theoretical framework, conceptualised by Macintyre and colleagues (2002), which was designed to represent five conceptual themes by which neighbourhoods influence health inequalities. The framework was designed for the customisation of pathways for any health related subject. Through applying the literature from Dental Public Health, and the wider health inequalities field, a series of theoretical pathways by which various aspects of neighbourhood environments influenced tooth decay were identified. This stage of the research also helped to operationalise these theoretical pathways, as data for each of the individual elements of the pathways were taken from the Adult Dental Health Survey (2009), with data for the neighbourhood level elements collected from a variety of other sources, including the 2011 Census, the Indices of Multiple Deprivation (2015), house price data (ONS, 2015), and geographical locations represented through point data.

- To build simulation models capable of representing these theoretical pathways

The second objective was achieved through the combination of two simulation models. Firstly, spatial microsimulation modelling was used to create a representative synthetic dataset of individuals for two study areas within Sheffield, by combining data from the 2011 Census and the Adult Dental Health Survey (2009). The data that was operationalised as part of the previous objective was included in this process, so that the individuals created in the dataset were allocated appropriate values for all of these variables. The individuals in this dataset formed the basis of the agent-based models that were built to test the theoretical pathways in an interactive and dynamic way. Through the use of GIS, shapefiles and point data, a model was created for two study areas in Sheffield, complete with representative populations of individuals. The processes and mechanisms of the theoretical pathways were then added to the agent-based models programmatically using code in order to provide the theoretical basis for the simulations.

- To use these simulation models to find the most influential theoretical pathways within different neighbourhoods within Sheffield

The theoretical pathways were tested in the agent-based models in a cumulative fashion, with an additional pathway being added to the models after each run. Through this process it was possible to analyse which pathways had a statistically significant impact on the overall tooth decay score for the models (either an increase or decrease). Any pathway that had a statistically significant impact on the outcome scores was given further consideration in an attempt to understand the nature of the mechanisms that may have led to the change in scores. A second simulation was then run with the addition of an extra shop in each study area, to test whether this had any effect on tooth decay scores. Through this series of simulations, the third objective of this research was achieved.

7.3. Conclusions

- This was the first study to apply a neighbourhood based theoretical framework to an oral health related outcome, in order to create theoretical pathways by which neighbourhoods may influence tooth decay.
- This was also the first piece of research to combine the population synthesis methods of spatial microsimulation modelling with the dynamic interactivity of agent-based modelling, in an attempt to simulate theoretical pathways.
- Statistical analysis of the output of the agent-based models suggests that a pathway representing the interactions between shops, diet and sugar intake had the most influence on tooth decay scores, in both study areas.
- These findings point to both individual level and neighbourhood level variables being potentially important for tooth decay scores.
- The addition of an extra shop to each study area produced divergent results however, with decay decreasing in the more affluent Sheffield West, while remaining similar in Sheffield East. This could suggest that the food environment in Sheffield West may be more beneficial for tooth decay.

7.4. Recommendations for future research

7.4.1. Theory

The theoretical work in this research has shown how the Dental Public Health literature can be applied to a conceptual framework on the subject of neighbourhood effects. There is great potential for other oral health related outcomes, and the literature associated with this, to be applied to such frameworks, allowing for an interdisciplinary approach to the study of oral diseases. Creating the theoretical pathways also helped to build a picture of research areas that have yet to be explored in depth. In this research this included themes such as the effects of housing environments, work, and particularly play environments (the main themes of Pathway 2), which remain underdeveloped. This also applies to the literature on the theme of area reputations and their effects on oral health outcomes, as well as the general lack of neighbourhood based features that have been included in studies. It is perhaps not surprising that certain areas of neighbourhood analysis remain underdeveloped within the dental literature, particularly given the focus on individual behaviours in relation to clinical outcomes, and social determinants of health that have been more prominent themes. Further research to address these gaps would help to develop the neighbourhood based literature however.

7.4.2. Spatial microsimulation

Many studies within Dental Public Health rely on the use of patient questionnaires and large scale sampling methods to obtain data for further analysis. Spatial microsimulation could offer a quicker, and most likely more cost effective alternative to these methods, while also offering the flexible approach of being able to add variables at any point once the original model is completed. Beyond its use a data creation tool, the method would also be particularly useful for future research into small areas and neighbourhoods, as this remains an under researched area within Dental Public Health. Dynamic spatial microsimulation also offers the opportunity to project populations into the future, and assess the effects of certain policy changes over time on subsections of the population (Ballas et al, 2006). Dynamic spatial microsimulation can also project the attributes of individual microunits into a future state based on predefined transition rates (or rules), in order to assess future individual attributes (Kavroudakis et al, 2013), further adding to its potential for use within Dental Public Health.

Future research should also aim to improve the fit of the external validation for these models. While sample size is likely to play a role (Ryan et al, 2009), it is worth considering other strategies to improve this. More stringent selection of constraint and target variables could aid in this, despite the risk of the model not constraining as well if less constraint variables are applied. In line with this, methods such as data linkage (Slack-Smith, 2012) and statistical matching (Rodgers, 1984; Moriarity and Scheuren, 2001) should be investigated alongside spatial microsimulation methods, to see if these methods could benefit microsimulation models with regard to sample sizes and variable breadth.

7.4.3. Agent-based models

In relation to the findings of this research, the interactive dynamics of sugar purchasing and consumption by individuals and families would be a suitable future research subject for the use of agent-based modelling, and would have very relevant public health implications. Future research should aim to build on the models in this thesis by fine tuning these simulations so that they are capable of making more precise predictions for health related outcomes. In line with this, Chapter 5 made reference to genetic algorithms which could be used to help parameterise such models in a potentially more efficient manner. Caution regarding accuracy in relation to real world data remains, and future research would benefit from attempts to measure the differences in accuracy between manual calibration, and the results of GAs, such as in the work of Heppenstall et al. (2007). A number of additional interactions would also add to the realism of the models. Currently the models lack some form of social interactions between agents, as well as family structures, and demographic turnover over the course of the model. An example of the latter can be seen in the work of Crooks (2008). Public transport systems, as well as differentiated agent movement speed would add further realism. Experimenting with weightings for variables (Dibben et al, 2007) may also provide more accuracy with regard to identifying more influential predictors of oral health.

One topic that could benefit from the use of agent-based models is water fluoridation, given this method's ability to model and simulate the effects of environmental factors on the health of populations. Other structural concepts, such as media and advertising would also make suitable future research topics for agent-based modelling. One idea could be to test the effects of local advertising campaigns on behaviours such as tooth brushing for example. The inclusion of wider regional and national level policies within neighbourhood level analysis would also be possible, as work by Heppenstall et al. (2016) asserts that 'subsystems do not operate in isolation. In the short term they might appear to be independent from the rest of the system, but in the long run they are dependent on the aggregate system behavior' (p.4). Creating models with influences over a number of different aggregate scales would allow for a comprehensive analysis of oral health related outcomes.

Finally, given the success of incorporating a theoretical framework into an agent-based model in this research, future analyses should look to expand this approach to other relevant frameworks and theories. Cerda et al. (2014) provide an example of this in relation to the theory of fundamental causes (Link and Phelan, 1995), while the social determinants of health conceptual framework (WHO, 2010) would also benefit from this approach.

7.4.4. Oral health

The finding that shops and supermarkets are potentially important neighbourhood based features warrants further investigation, particularly into dietary patterns and the types of sugary foods that are sold at these locations that may be most harmful to oral health. This would have important policy implications given the political landscape surrounding sugar in the UK currently. The introduction of the sugar tax in the UK in April 2018 means this variable will likely take on extra significance within the Dental Public Health landscape. Future research into the behavioural and social responses to this tax, as well as the effects on sales of soft drinks, would provide valuable insight into a key determinant of tooth decay.

Regarding the results of this research, future work of a similar nature should also be conducted in different, or across a range of geographical areas to assess the consistency of such findings. The research in this study was conducted across two contrasting study areas in Sheffield, and as such may not be generalisable to the population of the UK. The methods used in this research could be used to replicate such an approach.

Biological factors are another theme that should ideally be incorporated into future work where possible. Boyce et al. (2010) have shown that biological factors can be important in predicting tooth decay in children, however little work within Dental Public Health has incorporated such measures. This may primarily be a data issue, as operationalising biological data from public health datasets represents a significant challenge. It is worth considering the inclusion of such variables in future research however, as they likely play important roles, and would expand the already interdisciplinary nature of this type of work. Agent-based models have previously been used to simulate cariogenic activity in the oral biofilm (Head et al, 2014), so adding a societal element to such research could lend valuable insight into the relationship between biology and socio-economic variables.

7.5 Recommendations for policy

The main findings of this research are very topical and policy relevant, given the government's intention to introduce the sugar tax in April 2018 (HM Revenue and Customs, 2016). This legislation ties into both of the main findings of this research, as sugary products are likely sold in greatest quantities through supermarkets and shops. Being able to assess the potential effects of this tax, and gain an insight into sugar consumption in general would be a valuable policy tool. The methods used in this research would be highly relevant in aiding in such analysis, as well as studying future trends. Agent-based models have been shown to be effective in analysing the effects of policy implementation, in scenarios such as urban planning and land use (Ligmann-Zielinska and Jonkowski, 2007), as well as social and behavioural responses to tax based scenarios (Hashimzade et al, 2015).

Away from the findings of this research, the study of other dental based policies including themes such as new dental contracts, and the effects of the implementation of these, would also be beneficial to the field, and again would be suited to the methods used in this research. As mentioned previously, the effects of water fluoridation is another policy relevant topic that would benefit from such an approach.

7.6. Summary of implications

In summary, this research has demonstrated the benefits of a multi-disciplinary approach to investigating Dental Public Health related problems, by way of incorporation and testing of theoretical concepts through the use of simulation modelling. This approach has implications both conceptually and methodologically for Dental Public Health, as well as for policymakers given the relevance of the findings. Future analyses would benefit from using a combination of these methods in both a research and a policy context. The future use of theory to guide such a research approach is vital, and the design of the methodological side of the analysis allows for the addition of a new dimension to the study of oral health related outcomes.

References

Adair, P.M., Pine, C.M., Burnside, G., Nicoll, A.D., Gillett, A., Anwar, S., Broukal, Z., Chestnutt, I.G., Declerck, D., Ping, F.X., Ferro, R., Freeman, R., Grant-Mills, D., Gugeshe, T., Hunsrisakhun, J., Irigoyen-Camacho, M., Lo, E.C.M., Moola, M.H., Naidoo, S., Nyandindi, U., Poulsen, V.J., Ramos-Gomez, F., Razanamihaja, N., Shahid, S., Skeie, M.S., Skur, P., Splieth, C., Soo, T.C., Whelton, H and Young, D.W. (2003) Familial and cultural perceptions and beliefs of oral hygiene and dietary practices among ethnically and socio-economically diverse groups. Journal of Dental Research 82(B66).

Adegoke, O. (2016) A theoretical and practical approach to a persuasive agent model for change behaviour in oral care and hygiene. Universiti Utara Malaysian Electronic Theses and Dissertation - http://etd.uum.edu.my/6029/

Adler, N.E and Newman, K. (2002) Socioeconomic disparities in health: pathways and policies. Health Affairs 21(2). P.60-76.

Aida, J., Ando, Y., Oosaka, M., Niimi, K and Morita, M. (2008) Contributions of Social Context to Inequality in Dental Caries: A Multilevel Analysis of Japanese 3-year-old Children. Community Dentistry & Oral Epidemiology 36. P.149–156.

Aida, J., Kondo, K., Kondo, N., Watt, R.G., Sheiham, A., and Tsakos, G. (2011) Income inequality, social capital and self-rated health and dental status in older Japanese. Social Science and Medicine 73(10). P. 1561-1568.

Alameda County Oral Health Needs Assessment (2006) More than a toothache: untreated dental disease in our school children - http://www.acphd.org/media/53546/toothache.pdf

Al-Sudani, F.Y.H., Vehkalahti, M.M and Suominen, A.L. (2016) Association of current employment status with oral health-related behaviors: findings from the Finish Health 2000 Survey. European Journal of Oral Sciences 124. P.368-376.

Anderson, B. (2007) Creating small area income estimates: spatial microsimulation modelling. Department for Communities and Local Government - http://opendepot.org/166/1/CWP-2007-07-Income-Deprivation-England.pdf

Anderson, B. (2013) Estimating Small-Area Income Deprivation: An Iterative Proportional Fitting Approach. P. 49-67. In Tanton, R and Edwards, K. L. (2013) Spatial Microsimulation: A reference Guide for Users. Springer.

Axelsson, P., Paulander, J and Lindhe, J. (1998) Relationship between smoking and dental status in 35-, 50-, 65-, and 75-year-old individuals. Journal of Clinical Periodontology 25(4). P.297-305.

Antunes, J. L. F., Biazevic, M.G., de Araujo, M. E., Tomita, N. E., Chinellato, L. E and Narvia, P. C. (2001) Trends and spatial distribution of oral cancer mortality in Sao Paulo, Brazil, 1980-1998. Oral Oncology 37(4). P.345-350.

Antunes, J. L. F., Frazao, P., Narvai, P.C., Bispo, C.M and Pegoretti, T. (2002). Spatial analysis to identify differentials in dental needs by area-based measures. Community Dentistry and Oral Epidemiology 30(2). P.133-142.

Antunes, J.L.F., Junqueira, S.R., Frazao, P., Bispo, C.M., Pegoretti, T and Narvai, P.C. (2003a) City-level gender differentials in the prevalence of dental caries and restorative dental treatment. Health and Place 9(3). P.231-239.

Antunes, J. L. F., Pegoretti, T., de Andrade, F. P., Jungueira, S. R., Frazao, P and Narvai, P, C. (2003b) Ethnic disparities in the prevalence of dental caries and restorative dental treatment in Brazilian children. International Dental Journal 53(1). P. 7-12.

Antunes, J. L. F., Narvai, P.C and Nugent, Z.J. (2004) Measuring inequalities in the distribution of dental caries. Community Dentistry and Oral Epidemiology 32(1). P. 41-48.

Antunes, J. L. F., Peres, M. A., de Campos Mello, T. R and Waldman, E. A. (2006) Multilevel assessment of determinants of dental caries experience in Brazil. Community Dentistry and Oral Epidemiology 34(2): 146-152.

Antunes, J. L. F., Borrell, C., Perez, G., Boing, A. F and Wunsch-Filho, V. (2008) Inequalities in mortality of men by oral and pharyngeal cancer in Barcelona, Spain and Sao Paulo, Brazil, 1995-2003. International Journal for Equity in Health 7(14).

Armfield, J.M., Mejia, G.C and Jamieson, L.M (2013) Socioeconomic and psychosocial correlates of oral health. International Dental Journal 63(4). P.202-209.

Auchincloss, A. H and Diez Roux, A. V. (2008) A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health. American Journal of Epidemiology 168(1). P. 1-8.

Auchincloss, A.H., Riolo, R. L., Brown, D. G., Cook, J and Diez Roux, A. V. (2011) An agent-based model of income inequalities in diet in the context of residential segregation. American Journal of Preventive Medicine 40(3). P.303-311.

Auchincloss, A. H., Gebreab, S. Y., Mair, C and Diez Roux, A, V. (2012) A Review of Spatial Methods in Epidemiology, 2000-2010. Annual Review of Public Health, Vol 33. Fielding, J.E. Palo Alto, Annual Reviews. 33. P. 107-22.

Avlund, K., Holm-Pedersen, P., Morse, D.E., Viitanen, M and Winblad, B (2003) Social relations as determinants of oral health among persons over the age of 80 years. Community Dentistry and Oral Epidemiology 31(6). P.454-462.

Axelrod, R. (1997) The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton University Press.

Bailey, R.L., Ledikwe, J.H., Smiciklas-Wright, H., Mitchell, D.C and Jensen, G.L. (2004) Persistent oral health problems associated with comorbidity and impaired diet quality in older adults. Journal of the American Dietetic Association 104(8). P.1273-1276.

Baker, S.R and Gibson, B.G (2014) Social oral epidemic(olog)2y where next: one small step or one giant leap? Community Dentistry and Oral Epidemiology 42(6). P.481-494.

Ballas, D., Rossiter, D., Thomas, B., Clarke, G and Dorling, D. (2005a) Geography Matters. Joseph Rowntree Foundation - http://www.jrf.org.uk/publications/geography-matters-simulating-local-impacts-national-social-policies

Ballas, D., Clarke, G.P and Wiemers, E. (2005b) Building a Dynamic Spatial Microsimulation Model for Ireland. Population, Space and Place 11(3). P.157-172.

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B and Rossiter, D. (2005c) SimBritain: A spatial microsimulation approach to population dynamics. Population, Space and Place 11. P.13–34.

Ballas, D., Clarke, G.P., Dorling, D., Rigby, J and Wheeler, B. (2006) Using geographical information systems and spatial microsimulation for the analysis of health inequalities. Health Informatics Journal 12(1). P.65-79.

Bambra, C. (2011) Health inequalities and welfare state regimes: theoretical insights on a public health ‘puzzle’. Journal of Epidemiology and Community Health 65. P. 740-745.

Barker, M., Lawrence, W.T., Skinner, T.C., Haslam, C.O., Robinson, S.M., Inskip, H. M., Margetts, B.M., Jackson, A.A., Barker, D.J and Cooper, C. (2008) Constraints on food choices of women in the UK with lower educational attainment. Public Health Nutrition 11(12). P.1229-1237.

Batty, M., Desyllas, J and Duxbury, E. (2003) Safety in numbers? Modelling crowds and designing control for the Notting Hill Carnival. Urban Studies 40(8). P.1573-1590.

Batty, M., Barros, J and Alves, S. Jr. (2004) Cities: Continuity, transformation, and emergence. CASA Working Paper – Paper 72 - http://discovery.ucl.ac.uk/214/1/paper72.pdf

Bedos, C., Brodeur, J.M., Arpin, S and Nicolau, B. (2005) Dental caries experience: a two-generation study. Journal of Dental Research 84(10). P.931-936.

Belstrom, D., Fiehn, N.E., Nielsen, C.H., Holmstrup, P., Kirkby, N., Klepac-Ceraj, V., Paster, B.J and Twetman, S. (2014) Altered bacterial profiles in saliva from adults with caries lesions: a case-cohort study. Caries Research 48(5). P. 368-375.

Ben-Shlomo, Y and Kuh, D. (2002) A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. International Journal of Epidemiology 31(2). P. 285-293.

Bernabe, E., Delgado-Angulo, E.K., Vehkalahti, M.M., Aromaa, A and Suominen, A. L. (2014) Daily smoking and 4-year caries increment in Finnish adults. Community Dentistry and Oral Epidemiology 42(5). P.428-434.

Bernard, P., Charafeddine, R., Frohlich, K.L., Daniel, M., Kestens, Y and Potvin, L. (2007) Health inequalities and place: A theoretical conception of neighbourhood. Social Science and Medicine 65 (9). P. 1839-1852.

Berry, B.J.L and Garrison, W.L. (1958) The functional bases of the central place hierarchy. Economic Geography 34. P.145-54.

Bertea, P.C., Staehelin, K., Dratva, J and Stutz, E.Z. (2007) Female gender is associated with dental care and dental hygiene, but not with complete dentition in the Swiss adult population. Journal of Public Health 15. P.361-367.

Bhandari, B., Newton, J.T and Bernabe, E. (2015) Income inequality, disinvestment in health care and use of dental services. Journal of Public Health Dentistry 75. P.58-63.

Binnie, J. (1997) Coming out of geography: towards a queer epistemology? Environment and Planning D: Society and Space 15. P.223-237.

Birkin, M and Clarke, M. (1988) SYNTHESIS—a synthetic spatial information system for urban and regional analysis: methods and examples. Environment and Planning A 20(12). P.1645-1671.

Birkin, M and Clarke, M. (1989) The Generation of Individual and Household Incomes at the Small Area Level using Synthesis. Regional Studies 23(6). P. 535-548.

Black, D. (1980) The Black Report – Inequalities in health. Department of Health and Social Security

Blaxter, M. (1990) Health and lifestyles. Routledge, London.

Blocker, A. (2015) Package 'ipfp'. The Comprehensive R Archive Network - https://cran.r-project.org/web/packages/ipfp/index.html

Borenstein, H., Renahy, E., Quinonez, C and O'Campo, P. (2013) Oral Health, Oral Pain, and Visits to the Dentist: Neighborhood Influences among a Large Diverse Urban Sample of Adults. Journal of Urban Health 90(6). P. 1-15.

Borrell, L.N., Northridge, M.E., Miller, D.B., Golembeski, C.A., Spielman, S.E., Sclar, E.D and Lamster, I.B. (2006) Oral health and health care for older adults: a spatial approach for addressing disparities and planning services. Special care in dentistry: official publication of the American Association of Hospital Dentists, the Academy of Dentistry for the Handicapped, and the American Society for Geriatric Dentistry 26(6). P.252-256.

Boren, D., Sullivan, P.S., Beyrer, C., Baral, S.D., Bekker, L.G and Brookmeyer, R. (2014) Stochastic variation in network epidemic models: implications for the design of community level HIV prevention trials. Statistics in Medicine 33(22). P.3894-3904.

Bourdieu, P. (1984) Distinction. A social critique of the judgement of taste. Cambridge (Mass). Harvard University Press.

Bower, E., Gulliford M., Steele, J and Newton, T. (2007) Area deprivation and oral health in Scottish adults: a multilevel study. Community Dentistry and Oral Epidemiology 35(2). P.118-129.

Boyce, W.T., Den Besten, D.K., Stamperdahl, J., Zhan, L., Jiang, Y., Alder, N.E and Featherstone, J. D. (2010) Social disadvantage in childhood dental caries: The convergent roles of stress, bacteria and disadvantage. Social Science and Medicine 71(9). P.1644-1652.

Boyle, M.H and Willms, J.D. (1999) Place Effects for Areas Defined by Administrative Boundaries. American Journal of Epidemiology 149(6). P.577-585.

Bradley, J.E., Kirby, A.M and Taylor, P.J. (1978) Distance decay and dental decay: A study of dental health among primary school children in Newcastle upon Tyne. Regional Studies 12(5). P. 529-540.

Brennan, D.S., Spencer, A.J and Roberts-Thomson, K.F. (2007) Caries experience among 45-54 year olds in Adelaide, South Australia. Australian Dental Journal 52(2). P.122-127

Broomhead, T., Baker, S.R., Jones, K., Richardson, A and Marshman, Z. (2014) What are the most accurate predictors of caries in children aged 5 years in the UK? Community Dental Health 31(2). P.111-116.

Brouwers, L., Boman, M., Camitz, M., Makila, K and Tegnell, A. (2010) Micro-simulation of a smallpox outbreak using official register data. Eurosurveillance 15(35).

Brown, L.J., Caldwell, S.B and Eklund, S.A. (1995) How fee and insurance changes could affect dentistry: Results from a microsimulation model. Journal of the American Dental Association 126(4). P. 449-459.

Brown, D.G., Riolo, R., Robinson, D.T., North, M and Rand, W. (2005) Spatial process and data models: Toward integration of agent-based models and GIS. Journal of Geographic Systems 7. P.25–47.

Burden, S and Steel, D. (2016) Constraint choice for spatial microsimulation. Population, Space and Place 22. P.568-583.

Burke, D.S., Epstein, J.M., Cummings, D.A.T., Parker, J.I., Cline, K.C., Singa, R.M and Chakravarty, S. (2006) Individual-based computational modeling of smallpox epidemic control strategies. Academic Emergency Medicine 13(11). P.1142-1149.

Burt, B.A., Kolker, J.L., Sandretto, A.M., Yuan, Y., Sohn, W and Ismali, A. I. (2006) Dietary patterns related to caries in a low-income adult population. Caries Research 40(6). P.473-480.

Butani, Y., Weintraub, J.A and Barker, J.C. (2008) Oral health-related cultural beliefs for four racial/ethnic groups: assessment of the literature. BMC Oral Health 8(26).

Cajka, J.C., Cooley, P.C and Wheaton, W.D. (2010) Attribute Assignment to a Synthetic Population in Support of Agent-Based Disease Modeling. Methods Report RTI Press 19(1009). P.1-14.

Campbell, M.H. (2011) Exploring the social and spatial inequalities of ill-health in Scotland: A spatial microsimulation approach. White Rose eTheses online - http://etheses.whiterose.ac.uk/1942/

Campbell, M.H and Ballas, D. (2013) A spatial micro simulation approach to economic policy analysis in Scotland. Regional Science Policy and Practice 5(3). P. 263-288.

Campbell, M and Ballas, D. (2016) SimAlba: A spatial microsimulation approach to the analysis of health inequalities. Frontiers in Public Health 4(230).

Cannon, W.B. (1929) Bodily changes in pain, hunger, fear, and rage. Appleton, Oxford, UK.

Carvalho, M.L., Moyses, S.J., Bueno, R.E., Shimakura, S and Moyses, S.T. (2010) A geographical population analysis of dental trauma in school-children aged 12 and 15 in the city of Curitiba-Brazil. BMC Health Services Research 10(203).

Casamassimo, P.S., Thikkurissy, S., Edelstein, B.L and Maiorini, E. (2009) Beyond the dmft - The human and economic cost of early childhood caries. Journal of the American Dental Association 140(6). P.650 – 657.

Castle, C. J. E and Crook, A. T. (2006) Principles and concepts of agent-based modelling for developing geospatial simulation. UCL Working Paper Series, Paper 110, Centre for Advanced Spatial Analysis, University College London - http://discovery.ucl.ac.uk/3342/1/3342.pdf

Cataife, G. (2014) Small area estimation of obesity prevalence and dietary patterns: A model applied to Rio de Janeiro city, Brazil. Health and Place 26. P.47-52.

Celeste, R. K., Nadanovsky, P., Ponce de Leon, A and Fritzell, J. (2009) The individual and contextual pathways between oral health and income inequality in Brazilian adolescents and adults. Social Science & Medicine 69(10). P.1468-1475.

Celeste, R.K., Fritzell, J and Nadanovsky, P. (2011) The relationship between levels of income inequality and dental caries and periodontal diseases. Cadernos De Saude Publica 27(6).P.1111 – 1120.

Cerda, M., Tracy, M., Ahern, J and Galea, S. (2014) Addressing Population Health and Health Inequalities: The Role of Fundamental Causes. American Journal of Public Health 104(4). P.609-619.

Chaffee, B.W., Gansky, S.A., Weintraub, J.A., Featherstone, J.D.B and Romas-Gomez, F.J. (2014) Maternal oral bacterial levels predict early childhood caries development. Journal of Dental Research 93(3). P.238-244.

Chen, B and Cheng, H.H. (2010) A review of the applications of agent technology in traffic and transportation systems. IEEE Transactions on Intelligent Transportation System 11(2). P.485-497.

Chroinin, D.N., Montalta, A., Jahromi, S., Ingham, N., Beveridge, A and Foltyn, P. (2016) Oral health status is associated with common medical comorbidities in older hospital inpatients. Journal of the American Geriatrics Society 64(8). P.1696-1700.

Chu, C.H., Fung, D. S. H and Lo, E.C.M. (1999) Dental caries status of preschool children in Hong Kong. British Dental Journal. Volume 187(11). P.616-620.

Clarke, M., Forte, P., Spowage, M and Wilson, A.G. (1984) A strategic planning simulation model of a district health service system: The in-patient component and results. In van Elmeren, W., Engelbrecht, R and Flagle, C.D (eds.) Systems Science in Health Care. Berlin: Springer.

Clark, S.D., Birkin, M and Heppenstall, A. (2014) Sub regional estimates of morbidities in the English elderly population. Health and Place 27. P.176 – 185.

Coates, B. E and Rawstron, E. M. (1971) Regional Variations in Britain. B.T Batsford Ltd. London.

Cook, P.J and Walker, R.O. (1967) The geographical distribution of dental in the United Kingdom. British Dental Journal 122. P.494-499 & 551-558.

Costa, S.M., Martins, C.C., Bonfim, M.L.C., Zina, L.G., Paiva, S.M., Pordeus, I.A and Abreu, M.H.N.G. (2012) A systematic review of socioeconomic indicators and dental caries in adults. International Journal of Environmental Research and Public Health 9(10). P.3540-3574.

Cresswell, T. (2009) Place. In Thrift, N and Kitchen, R. (eds.) International Encyclopedia of Human Geography. Oxford: Elsevier, Vol. 8. P. 169-177.

Crooks, A.T. (2006) Exploring cities using agent-based models and GIS. UCL Working Paper Series, Paper 109 - http://discovery.ucl.ac.uk/3341/1/3341.pdf

Crooks, A.T. (2008) Constructing and implementing an agent-based model of residential segregation through vector GIS. UCL Working Paper Series, Paper 133 - http://discovery.ucl.ac.uk/15185/1/15185.pdf

Crooks, A., Castle, C and Batty, M. (2008) Key challenges in agent-based modelling for geo-spatial simulation. Computers, Environment and Urban Systems 32(6). P.417-430.

Crooks, A. T and Heppenstall, A. J. (2012) Introduction to Agent-Based Modelling. In Heppenstall, A.J., Crooks, A.T., See, L.M and Batty, M. (2012) Agent-Based Models of Geographical Systems. Springer. London.

Cummins, S., Macintyre, S., Davidson, S and Ellaway, A. (2005) Measuring neighbourhood social and material context: generation and interpretation of ecological data from routine and non-routine sources. Health and Place 11(3). P.249-260.

Cummins, S., Curtis, S., Diez Roux, A.V and Macintyre, S. (2007) Understanding and representing ‘place’ in health research: A relational approach. Social Science & Medicine 65(9). P.1825-1838.

Curtis, S and Rees-Jones, I. (1998) Is there a place for geography in the analysis of health inequality? Sociology of Health and Illness 20(5). P.645-672.

Curtis, S., Southall, H., Congdon, P and Dodgeon, B. (2004) Area effects on health variation over the life-course: analysis of the longitudinal study sample in England using new data on are of residence in childhood. Social Science and Medicine 58 (1). P.57-74.

Cutter, S.L. (1995) Race, class and environmental justice. Progress in Human Geography 19(1). P.111-122.

Dalstra, J.A.A., Kunst, A.E and Mackenbach, J.P. (2005) A comparative appraisal of the relationship of education, income and housing tenure with less than good health among the elderly Europe. Social Science and Medicine 62(8). P. 2046-2060.

Davenport, C.F., Elley, K.M., Fry-Smith, A., Taylor-Weetman, C.L and Taylor, R.S. (2003) The effectiveness of routine dental checks: a systematic review of the evidence base. British Dental Journal 195(2). P.87-98.

Dawson, R.J., Peppe, R and Wang, M. (2011) An agent-based model for risk-based flood incident management. Natural Hazards 59(1). P.167-189.

Densmore, O. (2010) NetLogo Bag of Tricks - http://backspaces.net/wiki/NetLogo_Bag_of_Tricks

De Graaf-Ruizendaal, W.A and De Bakker, D.H. (2014) The construction of a decision tool to analyse local demand and local supply for GP care using a synthetic estimation model. Human Resources for Health 11(55).

Delgado-Angulo, E.K., Bernabe, E and Marcenes, W. (2015) Ethnic inequalities in dental caries among adults in East London. Journal of Public Health 38(2). P.55-62.

Department for Communities and Local Government (2004 & 2010) The English indices of deprivation - https://www.gov.uk/government/collections/english-indices-of-deprivation

Department of Health (2015) Choosing better oral health: An oral health plan for England - http://webarchive.nationalarchives.gov.uk/+/dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/dh_4123251

Department for Education (2017) FE data library: apprenticeships – Apprenticeships demographics data tool: starts 2011/12 to 2016/17 reported to data - https://www.gov.uk/government/statistical-data-sets/fe-data-library-apprenticeships

Desai, P., Kruger, E., Trolio, R and Tennant, M. (2015) Western Australian schools access to dentally optimal fluoridated water. Australian Dental Journal 60(1). P.112-118.

Dibben, C., Atherton, I., Cox, M., Watson, V., Ryan, M and Sutton, M. (2007) Investigating the impact of changing the weights that underpin the Index of Multiple Deprivation 2004. Department for Communities and Local Government – http://webarchive.nationalarchives.gov.uk/20100410180038/http:/www.communities.gov.uk/documents/communities/pdf/323211.pdf

Diez Roux, A.V. (2001) Investigating neighborhood and area effects on health. American Journal of Public Health 91 (11). P.1783-1789.

Diez Roux, A.V. (2004) Estimating neighborhood health effects: the challenges of causal inference in a complex world - Commentary. Social Science & Medicine 58(10). P.1953-1960.

Diez Roux, A.V and Mair, C. (2010) Neighborhoods and health. Annals of the New York Academy of Sciences 1186. P.125-145.

Digimap website (2016) - http://digimap.edina.ac.uk/webhelp/os/data_information/os_products/mastermap_topo.htm

Dijkstra, E. W. (1959) A note on two problems in connexion with graphs. Numerische Mathematik 1. P. 269–271. doi:10.1007/BF01386390.

Donaldson, A.N., Everitt, B., Newton, T., Steele, J., Sherriff, M and Bower, E. (2008) The effects of social class and dental attendance on oral health. Journal of Dental Research 87(1). P.60-64.

Dorling, D. (2001) Anecdote is the singular of data. Environment and Planning A 33. P.1335 – 1340.

Doyal, L and Nadioo, S. (2010) Why dentists should take a greater interest in sex and gender. British Dental Journal 209(7). P.335-337.

Drummond, A.M.A., Ferreira, E.F., Gomes, V.E and Marcenes, W. (2015) Inequality of experience of dental caries between different ethnic groups of Brazilians aged 15 to 19 years. Plos One 10(12). e0145553.

Dumas, A.S and Polk, D. (2015) Pediatric dental location and utilization in a high resource setting. Journal of Public Health Dentistry 75(3). P.183-190.

Dyck, I. (1989) Integrating Home and Wage Workplace: Women's Daily Lives in a Canadian Suburb. The Canadian Geographer 33(4). P. 329–341.

Eckersley, A.J and Blinkhorn, F.A. (2001) Dental attendance and dental health behaviour in children from deprived areas of Salford, North-West England. International journal of Paediatric Dentistry 11. P.103-109.

Edwards, K.L and Clarke, G. (2007) SimObesity: Combinatorial Optimisation (Deterministic) Model. In Tanton, R and Edwards, K.L. (eds.) Spatial microsimulation: a reference guide for users. Springer Netherlands. P.69-85.

Edwards, K.L and Clarke, G.P. (2009) The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. Social Science and Medicine 69(7). P. 1127-1134.

Edwards, K.L., Clarke, G,P., Ransley, J,K and Cade, J. (2010) The neighbourhood matters: studying exposures relevant to childhood obesity and the policy implications in Leeds, UK. Journal of Epidemiology and Community Health 64(3). P.194-201.

Edwards, K,L., Clarke, G,P., Thomas, J and Forman, D. (2011) Internal and external validation of Spatial Microsimulation models. Small area estimates of adult obesity. Applied Spatial Analysis and Policy 4(4). P. 281-300.

Ellen, I.G., Mijanovich, T and Dillman, K. (2001) Neighbourhood effects on health: exploring the links and assessing the evidence. Journal of Urban Affairs 23(3-4). P.391-408.

Erdem, O., Prins, R.G., Voorham, T.A.J.J., van Lenthe, F.J and Burdorf, A. (2015) Structural neighbourhood conditions, social cohesion and psychological distress in the Netherlands. European Journal of Public Health 25(6). P.995-1001.

Emami, E., Khiyani, M.F., Habra, C.P, Chasse, V and Rompre, P.H. (2016) Mapping the Quebec dental workforce: ranking rural oral health disparities. Rural and Remote Health 16(1):3630.

Entwisle, B. (2007) Putting people into place. Demography 44(4). P. 687-703.

Epstein, J,M and Axtell, R,L. (1996) Growing artificial societies: social science from the bottom up. The MIT Press, Cambridge MA.

ESRI website (2016) – GIS Dictionary - http://support.esri.com/en/knowledgebase/Gisdictionary/browse

Fejerskov, O. (1997) Concepts of dental caries and their consequences for understanding the disease. Community Dentistry and Oral Epidemiology 25(1). P.5-12.

Feng, X., Sambamoorthii, U and Wiener, R.C. (2016) Dental workforce availability and dental services utilization in Appalachia: a geospatial analysis. Community Dentistry and Oral Epidemiology 45(2). P.145-152.

Ferraro, M and Vieira, A.R. (2010) Explaining Gender Differences in Caries: A Multifactorial Approach to a Multifactorial Disease. International Journal of Dentistry, article ID 649643.

Ferguson, N.S., Lamb, K.E., Wang, Y., Ogilvie, D and Ellaway, A. (2013) Access to Recreational Physical Activities by Car and Bus: An Assessment of Socio-Spatial Inequalities in Mainland Scotland. Plos One 8(2).

Ferrell, J. (1995) Urban graffiti: crime, control and resistance. Youth and Society 27. P.73-92.

Field, A., Miles, J and Field, Z. (2012) Discovering statistics using R. Sage Publishing, London.

Finlayson, T.L., Williams, D.R., Siefert, K., Jackson, J.S and Nowjack-Raymer, R. (2010) Oral health disparities and psychosocial correlates of self-rated oral health in the national survey of American life. American Journal of Public Health 100(Supplement 1). P.246-255.

Flusty, S. (2000) Thrashing downtown: play as resistance to the spatial and representational regulation of Los Angeles. Cities 17(2). P.149-158.

Fone, D., Dunstan, F., Lloyd, K., Williams, G., Watkins, J and Palmer, S. (2007) Does social cohesion modify the association between area income deprivation and mental health? A multilevel analysis. International Journal of Epidemiology 36(2). P.338-345.

Fonseca, M.A.D (2012) The Effects of Poverty on Children's Development and Oral Health. Pediatric Dentistry. Volume 34(1). P.32-38.

Franceschi, S., Levi, F., La Vecchia, C., Conti, E., Dal Maso, L., Barzan, L and Talamini, R. (1999) Comparison of the effect of smoking and alcohol drinking between oral and pharyngeal cancer. International Journal of Cancer 83(1). P.1-4.

Freire, M.C.M., Reis, S.C.G.B., Figueiredo, N., Peres, K.G., Moreira, R.S and Antunes, J.L.F. (2013) Individual and contextual determinants of dental caries in Brazilian 12-years-olds in 2010. Revista de Saúde Pública 47(3). P.1-10.

Friche, A.A.L., Diez Roux, A.V., Cesar, C.C., Xavier, C.C., Proietta, F.A and Caiaffa, W.T. (2012) Assessing the psychometric and economic properties of neighborhood scales in developing countries: Saude Em Beaga study, Belo Horizonte, Brazil, 2008-2009. Journal of Urban Health 90(2). P.246-261.

Gardner, M. (1970) The fantastic combinations of John Conway's new solitaire game "Life". Scientific American 223. P.120–123.

Galobardes, B., Shaw, M., Lawlor, D.A., Lynch, J.W and Davey Smith, G. (2004) Indicators of socioeconomic position (Part 1). Journal of Epidemiology and Community Health 60(1). P.7-12.

Galster, G., Andersson, R and Musterd, S. (2010) Who Is Affected by Neighbourhood Income Mix? Gender, Age, Family, Employment and Income Differences. Urban Studies 47(14). P. 2915-2944.

Garland, J and Chakraborti, N. (2006) Race, space and place – examining identity and cultures of exclusion in rural England. Ethnicities 6(2). P.159-177.

Gastner, M.T and Newman, M.E.J. (2004) Diffusion-based method for producing density-equalizing maps. Proceedings of the National Academy of Sciences of the United States of America 101(20). P.7499-7504.

Geard, N., McCaw, J.M., Dorin, A., Korb, K.B and McVernon, J. (2013) Synthetic Population Dynamics: A Model of Household Demography. Journal of Artificial Societies and Social Simulation 16(1). P.8.

Gelman, A. (2005) Multilevel (hierarchical) modelling: what it can and can't do. Online PDF - http://www.cs.berkeley.edu/~russell/classes/cs294/f05/papers/gelman-2005.pdf

Geolytix (2016) Supermarket location data - http://blog.geolytix.net/tag/supermarket-location-data/

Gershuny, J and Sullivan, O. (2017) United Kingdom Time Use Survey, 2014-2015. Centre for Time Use Research, University of Oxford. [data collection]. UK Data Service. SN: 8128 - http://doi.org/10.5255/UKDA-SN-8128-1

Geyer, S., Schneller, T and Micheelis, W. (2010) Social gradients and cumulative effects of income and education on dental health in the fourth German Oral Health Study. Community Dentistry and Oral Epidemiology 38. P.120-128.

Giddens, A. (1984) The constitution of society. Berkeley, CA: University of California Press.

Gini, C. (1912) Variabilità e mutabilità. Reprinted in Pizetti, E and Salvemini, T. (eds) (1955) Memorie di metodologica statistica. Rome: Libreria Eredi Virgilio Veschi.

Gokhale, N and Nuvvula, S. (2016) Influence of socioeconomic and working status of the parents on the incidence of their children's dental caries. Journal of Natural Science, Biology and Medicine 7(2). P.127-129.

Google Maps (2016a) Map of buildings in Sheffield East [online]. Available from: https://www.google.co.uk/maps/place/Burngreave,+Sheffield+S4+7HA/@53.3930788,-1.4672431,2774m/data=!3m1!1e3!4m5!3m4!1s0x4879786e1acf73d5:0xf9854fab2402cf1!8m2!3d53.3930797!4d-1.4584883 [accessed August 29th 2016]

Google Maps (2016b) Map of buildings in Sheffield West [online]. Available from: https://www.google.co.uk/maps/place/Ranmoor,+Sheffield+S10+3HJ/@53.3718362,-1.5353158,2773m/data=!3m2!1e3!4b1!4m5!3m4!1s0x4879820290c23b2b:0x2e0c42dc6c5f24df!8m2!3d53.371837!4d-1.526561 [accessed August 29th 2016]

Google Maps (2017a) Map of buildings in Sheffield East [online]. Available from: https://www.google.co.uk/maps/place/Attercliffe,+Sheffield/@53.391431,-1.4397095,2772m/data=!3m1!1e3!4m5!3m4!1s0x4879781bf1b86999:0xac765274af5c6bb5!8m2!3d53.394858!4d-1.434197 [accessed January 25th 2017]

Google Maps (2017b) Map of buildings in Sheffield West [online]. Available from: https://www.google.co.uk/maps/place/Crosspool,+Sheffield+S10+5DP/@53.3769202,-1.5259218,2772m/data=!3m1!1e3!4m5!3m4!1s0x487982076ce78d53:0x1f15115263ab091e!8m2!3d53.376921!4d-1.517167 [accessed January 25th 2017]

Gorman, D.M., Mezic, J., Mezic, I and Gruenewald, P.J. (2006) Agent-based modelling of drinking behaviour: A preliminary model and potential applications to theory and practice. American Journal of Public Health 96(11). P.2055-2060.

Graham, H. (2002) Building an inter-disciplinary science of health inequalities: the example of life course research. Social Science and Medicine 55. P.2005–2016.

Gregory, D., Johnston, R., Pratt, G., Watts, M and Whatmore, S. (2009) The Dictionary of Human Geography (5th edition). Wiley-Blackwell. P.110-111.

Griffin, S.O., Regnier, E., Griffin, P.M., and Huntley, V. (2007) Effectiveness of fluoride in preventing caries in adults. Journal of Dental Research. Volume 86(5). P.410-415.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepson, J.U., Jorgensen, C., Mooij, W.M., Muller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R.A., Vabo, R., Visser, U and DeAngelis, D.L. (2006) A standard protocol for describing individual-based and agent-based models. Ecological Modelling 198(1-2). P.115-126.

Guiney, H., Woods, N., Whelton, H and Morgan, K. (2011) Predictors of utilisation of dental care services in a nationally representative sample of adults. Community Dental Health 28(4). P.269-273.

Haase, D., Lautenbach, S and Seppelt, R. (2010) Modeling and simulating residential mobility in a shrinking city using an agent-based approach. Environmental Modelling and Software 25(10). P.1225-1240.

Hagerstrand, T. (1974) Tidsgeografisk beskrivn-ing. Syfte och postulat. Svensk geografisk arsbok 50. P.86-94

Hancock, M., Calnan, M and Manley, G. (1999) Private or NHS General Dental Service care in the United Kingdom? A study of public perceptions and experiences. Journal of Public Health 21(4). P.415-420.

Harland, K., Heppenstall, A., Smith, D and Birkin, M. (2012) Creating realistic synthetic populations at varying spatial scales: A comparative critique of population synthesis techniques. Journal of Artificial Societies and Social Simulation 15(1).

Harris, C.K., Warnakulasuriya, K.A., Johnson, N.W., Gelbier, S and Peters, T.J. (1996) Oral health in alcohol misusers. Community Dental Health 13(4). P.199-203.

Harris, C., Warnakulasuriya, K.A., Gelbier, S., Johnson, N.W and Peters, T.J. (1997) Oral and dental health in alcohol misusing parents. Alcoholism – Clinical and Experimental Research 21(9). P.1707-1709.

Hart, P.E., Nilsson, N.J and Raphael, B. (1968) A Formal Basis for the Heuristic Determination of Minimum Cost Paths. IEEE Transactions on Systems Science and Cybernetics SSC4. 4(2). P.100-107. doi:10.1109/TSSC.1968.300136.

Hart, J.T. (1971) The inverse care law. The Lancet 297(7696). P. 405-412.

Harvey, D. (2000) Spaces of hope. Edinburgh University Press.

Hashimzade, N., Myles, G.D., Page, F and Rablen, M.D. (2015) The use of agent-based modelling to investigate tax compliance. Economics of governance 16(2). P.143-164.

Hassani-Mahmooei, B and Parris, B.W. (2012) Climate change and internal migration patterns in Bangladesh: an agent-based model**.** Environmental and Development Economics 17(6). P.763-780.

Haynes, R., Daras, K., Reading, R and Jones, A. (2007) Modifiable neighbourhood units, zone design and residents ' perceptions. Health and Place 13(4). P. 812-825.

Head, D.A., Marsh, P.D and Devine, D.A. (2014) Non-lethal control of the cariogenic potential of an agent-based model for dental plaque. Plos One 9(8): e105012.

Heard, D., Dent, G., Schifeling, T and Banks, D. (2015) Agent-based models and microsimulation. Annual Review of Statistics and its Application 2. P.259-272.

Heilmann, A., Tsakos, G and Watt, R.G. (2015) Oral health over the life course. In Burton-Jeangros, C., Cullati, S., Sacker, A and Blane, D. (eds.) A life course perspective on health trajectories and transitions. Springer. P. 39-59.

Heppenstall, A.J., Evans, A and Birkin, M.H. (2007) Genetic algorithm optimisation of an agent-based model for simulating a retail market. Environment and Planning B: Planning and Design 34(6). P.1051-1070.

Heppenstall, A.J., Malleson, N and Crooks, A. (2016) "Space, the final frontier": How good are agent-based models at simulating individuals and space in cities? Systems 4(1). P.9.

Hermes, K and Poulsen, M. (2012) Small area estimates of smoking prevalence in London. Testing the effect of input data. Health and Place 18(3). P. 630-638.

Hill, K.B., Chadwick, B., Freeman, R., O'Sullivan, I and Murray, J.J. (2013) Adult Dental Health Survey 2009: relationships between dental attendance patterns, oral health behaviour and the current barriers to dental care. British Dental Journal 214. P.25-32.

Hirsch, G.B., Edelstein, B.L., Frosh, M and Anselmo, T. (2012) A Simulation Model for Designing Effective Interventions in Early Childhood Caries. Preventing Chronic Disease 9(E66).

HM Revenue and Customs (2016) Soft Drinks Industry Levy - https://www.gov.uk/government/publications/soft-drinks-industry-levy/soft-drinks-industry-levy - Published 5th December 2016

Hobdell, M.H., Oliveira, E.R., Bautista, R., Myburgh, N.G., Lalloo, R., Narendran, S and Johnson, N.W. (2003) Oral diseases and socio-economic status (SES). British Dental Journal 194(2). P.91-96.

Hooley, M., Skouteris, H., Boganin, C., Satur, J and Kilpatrick, N. (2012) Parental influence and the development of dental caries in children aged 0-6 years: a systematic review of the literature. Journal of Dentistry 40(11). P.873-885.

Horner, M.W and Mascarenhas, A.K. (2007) Analyzing location-based accessibility to dental services: An Ohio case study. Journal of Public Health Dentistry 67(2). P.113-118.

Hudson, K., Stockard, J and Ramberg, Z. (2007) The impact of socioeconomic status and race-ethnicity on dental health. Sociological Perspectives 50(1). P.7-25.

Hunsinger, E. (2008) Iterative proportional fitting for a two-dimensional table. Alaska Department of Labor and Workforce Development - http://www.demog.berkeley.edu/~eddieh/IPFDescription/AKDOLWDIPFTWOD.pdf

Inagami, S., Cohen, D.A., Finch, B.K and Asch, S.M. (2006) You are where you shop: grocery store locations, weight, and neighborhoods. American Journal of Preventive Medicine 31(1). P.10-17.

Indices of Multiple Deprivation (2015a) Health domain data: Years of potential lost life - https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Indices of Multiple Deprivation (2015b) Crime domain data - https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Jager, R., van den Berg, N., Hoffmann, W., Jordan, R.A and Schwendicke, F. (2016) Estimating future dental services’ demand and supply: a model for Northern Germany. Community Dentistry and Oral Epidemiology 44(2). P.169-179.

Jamieson, L.M and Thomson, M.W (2006) Adult oral health inequalities described using area-based and household-based socioeconomic status measures. Journal of Public Health Dentistry 66(2). P.104-109.

Johansson, A and Kretz, T. (2012) Applied Pedestrian Modeling. In Heppenstall, A.J., Crooks, A.T., See, L.M and Batty, M. (eds.) Agent-based models of geographical systems. Springer. P.451-462.

Johnson, S.D and Groff, E.R. (2014) Strengthening Theoretical Testing in Criminology Using Agent-based Modeling. Journal of Research in Crime and Delinquency 51(4). P.509-525.

Jones, C.M., Taylor, G.O., Whittle, J.G., Evans, D and Trotter, D.P. (1997) Water fluoridation, tooth decay in 5 year olds, and social deprivation measured by the Jarman score: analysis of data from British dental surveys. British Medical Journal 315(7107). P.514-517.

Jones, C.M. (2001) Capitation registration and social deprivation in England. An inverse 'dental' care law? British Dental Journal. Volume 190(4). P.203-206.

Jones, M., Shah, T., Hayes, A., Uswak, G and Bell, S. (2016a) Dental service disparities in Canada: A Saskatoon, Saskatchewan case study. The Canadian Geographer 60(3). P.346-355.

Jones, P.M., Lovelace, R and Dumont, M. (2016b) rakeR: Easy spatial microsimulation (raking) in R. The Comprehensive R Archive Network (CRAN) - https://cran.r-project.org/web/packages/rakeR/index.html

Julihn, A., Agholme, M.B., Grindefjord, M and Modeer, T. (2006) Risk factors and risk indicators associated with high caries experience in Swedish 19-year-olds. Acta Odontologica Scandinavica 64(5). P.267-273.

Kant, A.K. (2004) Dietary patterns and health outcomes. Journal of the American Dietetic Association 104(4). P.615-635.

Kassebaum, N.J., Bernabe, E., Dahiya, M., Bhandari, B., Murray, C.J.L and Marcenens, W. (2015) Global Burden of Untreated Caries: A Systematic Review and Metaregression. Journal of Dental Research 94(5). P. 650-658.

Kassel, J.D., Stroud, L.R and Paronis, C.A. (2003) Smoking, stress, and negative affect: Correlation, causation, and context across stages of smoking. Psychological Bulletin 129(2). P. 270-304.

Kavroudakis, D., Ballas, D and Birkin, M. (2013) SimEducation: a dynamic spatial microsimulation model for understanding educational inequalities. In Tanton, R and Edwards, K.L. (eds.) Spatial microsimulation: a reference guide for users. Springer. P.209-222.

Kawachi, I and Kennedy, B.P. (1999) Income inequality and health: pathways and mechanisms. Health Services Research 34(1 part 2). P.215-227.

Kidd, E.A.M and Fejerskov, O. (2004) What Constitutes Dental Caries? Histopathology of Carious Enamel and Dentin Related to the Action of Cariogenic Biofilms. Journal of Dental Research 83(Special issue C). P.35-38.

Kim, J., Busogi, M., Ahn, C., Choi, M., Choi, E., Kim, K and Kim, N. (2014) Preliminary Research on an Agent-Based Modeling and Simulation in Demographic Study. Advanced Science and Technology Letter (58). P.44-47.

Kitchin, R. (1998) 'Out of place', 'knowing one's place': space, power and the exclusion of disabled people. Disability and Society 13(3). P.343-356.

Klabunde, A and Willekens, F. (2016) Decision-Making in Agent-Based Models of Migration: State of the Art and Challenges. European Journal of Population 31(1). P.73-97.

Koh, K, Grady, S.C and Vojnovic, I. (2015) Using simulated data to investigate the spatial patterns of obesity prevalence at the Census Tract level in metropolitan Detroit. Applied Geography 62. P.19-28.

Krause, D.D., May, W.L and Cossman, J.S. (2012) Overcoming data challenges examining oral health disparities in Appalachia. Online Journal of Public Health Informatics 4(3).

Krieger, N., Williams, D.R and Moss, N.E. (1997) Measuring social class in US public health research: concepts, methodologies, and guidelines. Annual Review of Public Health 18. P.341-378.

Kruger, E., Tennant, M and George, R. (2011) Application of geographic information systems to the analysis of private dental practices distribution in Western Australia. Rural Remote Health 11(3).

Kruger, E., Whyman R and Tennant, M. (2012) High-acuity GIS mapping of private practice dental services in New Zealand: does service match need? International Dental Journal 62(2). P.95-99.

Kuh, D and Wadsworth, M. (1991) Childhood influences on adult male earnings in a longitudinal-study. British Journal of Sociology 42(4). P.537-555.

Kumar, S., Tadakamadla, J and Johnson, N.W. (2016) Effect of tooth brushing frequency on incidence and increment of dental caries: A systematic review and meta-analysis. Journal of Dental Research 95(1). P.1230-1236.

Landes, D.P and Jardine, C. (2010) Targeting dental resources to reduce inequalities in oral health in the North East of England - a health equity audit methodology to evaluate the effects of practice location, practice population and deprivation. British Dental Journal 209(3).

Lang, I.A., Gibbs, S.J., Steel, N and Melzer, D. (2008) Neighbourhood deprivation and dental service use: a cross-sectional analysis of older people in England. Journal of Public Health 30(4). P.472-478.

LeDoux, T.F and Vojnovic, J. (2014) Examining the role between the residential neighborhood food environment and diet among low-income households in Detroit, Michigan. Applied Geography 55. P.9-18.

Lida, H and Rozier, R.G. (2013) Mother perceived social capital and children’s oral health and use of dental care in the United States. American Journal of Public Health 103(3). P.480-487.

Ligmann-Zielinska, A and Jankowski, P. (2007) Agent-based models as laboratories for spatially explicit planning policies. Environment and Planning B: Planning and Design 34(2). P.316-335.

Link, B.G and Phelan, J. (1995) Social conditions as fundamental causes of disease. Journal of Health and Social Behaviour. Extra Issue: Forty Years of Medical Sociology: The State of the Art and Directions for the Future. P.80-94.

Listl, S. (2012) Inequalities in dental attendance across the life-course. Journal of Dental Research 91(7s). P.91-97.

Lochner, K.A., Kawachi, I., Brennan, R.T and Buka, S.L. (2003) Social capital and neighborhood mortality rates in Chicago. Social Science and Medicine 56(8). P.1795-1805.

Locker, D. (1992) Smoking and oral health in older adults. Canadian Journal of Public Health 83(6). P.429-432.

Lopez, R., Frydenburg, M and Baelum, V. (2009) Contextual effects in the occurrence of periodontal attachment loss and necrotizing gingival lesions among adolescents. European Journal of Oral Sciences 117(5). P. 547-554.

Lorber, M.F., Slep, A.M., Heyman, R.E., Xu, S., Dasanayake, A.P and Wolff, M.S. (2014) Noxious family environments in relation to adult and childhood caries. Journal of the American Dental Association 145(9). P. 924-930.

Lovelace, R and Ballas, D. (2013) 'Truncate, replicate, sample': A method for creating integer weights for spatial microsimulation. Computers, Environment and Urban Systems. Volume 41. P.1-11.

Lovelace, R and Dumont, M. (2016) Spatial Microsimulation in R. CRC Press – Taylor and Francis Group.

Lucarotti, P.S.K and Burke, F.J.T. (2015) Factors influencing patients' continuing attendance at a given dentist. British Dental Journal 218(6). E13.

Lukacs, J.R and Largaespada, L.L. (2006) Exploring sex differences in dental caries prevalence: saliva, hormones, and "life-history" etiologies. American Journal of Human Biology 18(4). P.540-555.

Lukacs, J. R. (2008) Fertility and Agriculture Accentuate Sex Differences in Dental Caries Rates. Current Anthropology 49(5). P.901-914.

Luke, A.D and Stamatakis, K.A. (2012) Systems science methods in public health. Annual Review of Public Health 33. P.357-376.

Lymer, S., Brown, L., Yap, L and Harding, A. (2008) 2001 regional disability estimates for New South Wales, Australia, using Spatial Microsimulation. Applied Spatial Analysis and Policy 1(2). P. 99-116.

Lymer, S., Brown, L., Harding, A and Yap, M. (2009) Predicting the need for aged care services at the small area level: the CAREMOD spatial microsimulation model. International Journal of Microsimulation 2(2). P.27-42.

Macal, C.M and North, M.J. (2010) Tutorial on agent-based modelling and simulation. Journal of Simulation 4(3). P.151-162.

MacDonald, R., Shildrick, T., Webster, C and Simpson, D. (2005) Growing Up in Poor Neighbourhoods: The Significance of Class and Place in the Extended Transitions of 'Socially Excluded' Young Adults. Sociology 39(5). P.873-891.

Macdonald, L., Ellaway, A., Ball, K and Macintyre, S. (2011) Is proximity to a food retail store associated with diet and BMI in Glasgow, Scotland? BMC Public Health 11(464).

Macintyre, S., Maciver, S and Sooman, A. (1993) Area, Class and Health: Should we be Focusing on Places or People? Journal of Social Policy 22(2). P. 213-234.

Macintyre, S and Ellaway, A. (2000) Ecological approaches: Rediscovering the role of the physical and social environment. In Berkman, L.F and Kawachi, I. (eds.) Social Epidemiology. P.332-348. New York: Oxford University Press.

Macintyre, S., Ellaway, A and Cummins, S. (2002) Place effects on health: how can we conceptualise, operationalise and measure them? Social Science and Medicine 55 (1). P. 125-139.

Macintyre, S., Ellaway, A., Hiscock, R., Kearns, A., Der, G and McKay, L. (2003) What features of the home and the area might help to explain observed relationships between housing tenure and health? Evidence from the west of Scotland. Health & Place 9(3). P.207-218.

Macintyre, S., Macdonald, L and Ellaway, A. (2008) Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland. Social Science and Medicine. Volume 67(6). P.900-914.

Mackenbach, J.P. (2012) The persistence of health inequalities in modern welfare states: The explanation of a paradox. Social Science and Medicine 75(4). P.761-769.

Madan, C., Kruger, E., Perera, I and Tennant, M. (2010) Trends in demand for general anaesthetic care for paediatric caries in Western Australia: geographic and socio-economic modelling of service utilisation. International Dental Journal 60(3). P.190-196.

Malleson, N., Evans, A., Heppenstall, A and See, L. (2010) Evaluating an Agent-Based Model of Burglary. University of Leeds Working Paper - http://www.geog.leeds.ac.uk/fileadmin/downloads/school/research/wpapers/10_1.pdf

Malleson, N., Heppenstall, A., See, L and Evans A. (2013) Using an agent-based crime simulation to predict the effects of urban regeneration on individual household burglary risk. Environment and Planning B: Planning and Design 40(3). P.405-426.

Mamai-Homata, E., Topitsoglou, V., Oulis, C., Margaritis, V and Polychronopoulou, A. (2012) Risk indicators of coronal and root caries in Greek middle aged adults and senior citizens. BMS Public Health 12(484).

Manson, S.M., Sun, S and Bonsal, D. (2012) Agent-Based Modeling and Complexity. In Heppenstall, A., Crooks, A., See., L and Batty, M. (2012) Agent-Based Models of Geographical Systems. (P.125-141). Springer.

Marcenes, W.S and Sheiham, A. (1992) The relationship between work stress and oral health status. Social Science and Medicine 35(12). P.1511-1520.

Marinho, V.C., Higgins, J.P., Sheiham, A and Logan, S. (2003) Fluoride toothpastes for preventing dental caries in children and adolescents. The Cochrane Database of Systematic Reviews - CD002278.

Marmot, M.G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A and Davey Smith, M.G. (1991) Health inequalities among British civil servants: the Whitehall 2 study. The Lancet 337(8754). P. 1387-1393.

Marmot, M. (2005) Social determinants of health inequalities. The Lancet 365(9464). P.1099-1104.

Marmot, M., Allen, J., Goldblatt, P., Boyce, T., McNeish, D., Grady, M and Geddes, I. (2010) Fair Society, Healthy Lives: The Marmot Review – Strategic Review of Health Inequalities in England post-2010. London: Marmot Review Team.

Marshall, B.D.L., Paczkowski, M.M., Seemann, L., Tempalski, B., Pouget, E.R., Galea, S and Friedman, S.R. (2012) A Complex Systems Approach to Evaluate HIV Prevention in Metropolitan Areas: Preliminary Implications for Combination Intervention Strategies. Plos One 7(9).

Massey, D. (1985) New direction in space. In Gregory, D and Urry, J (1985) Social relations and spatial structures. Palgrave. P.9-19.

Massey, D. (1994) Space, place and gender. University of Minnesota Press. Minneapolis.

Matthews, R.B., Gilbert, N., Roach, A and Polhill, J.G. (2007) Agent-based land-use models: a review of applications. Landscape Ecology 22(10). P.1447-1459.

Mathur, M.R., Tsakos, G., Parmar, P., Millett, C.J and Watt, R.G. (2016) Socioeconomic inequalities and determinants of oral hygiene status among urban Indian adolescents. Community Dentistry and Oral Epidemiology 44. P.248-254.

Mazur, E and Urbanek, J. (1983) Space in geography. GeoJournal 7(2). P. 139-143.

McDonagh, M.S., Whiting, P.F., Wilson, P.M., Sutton, A.J., Chestnutt, I., Cooper, J., Misso, K., Bradley, M., Treasure, E and Kleijnen, J. (2000) Systematic review of water fluoridation. British Medical Journal 321. P.855-859.

McDonald, J.T and Conde, H. (2010) Does geography matter? The health service use and unmet health care need of older Canadians. Canadian Journal of Aging 29(1). P.23-37.

McCartney, G., Collins, C and Mackenzie, M. (2013) What (or who) causes health inequalities: Theories, evidence and implications? Health Policy 113(3). P.221-227.

McGrady, M.G., Ellwood, R.P., Maguire, A., Goodwin, M., Boothman, N and Pretty, I.A. (2012) The association between social deprivation and the prevalence and severity of dental caries and fluorosis in populations with and without water fluoridation. BMC Public Health 12(1222). P.1-17.

McGrath, C., Bedi, R and Dhawan, N. (2001) Who has difficulty in registering with an NHS dentist? – A national survey. British Dental Journal 191(12). P.682-685.

McGuire, S., Kruger, E and Tennant, M. (2011) Travel patterns for government emergency dental care in Australia: a new approach using GIS tools. Australian Dental Journal 56(4). P.389-393.

McKernan, S.C., Kuthy, R.A., Hanley, P.F., Jones, M.P., Momany, E.T., McQuistan, M.R and Damiano, P.C. (2015) Geographic variation of dental utilization among low income children. Health and Place 34. P.150-156.

McKernan, S.C., Pooley, M.J., Momany, E.T and Kuthy, R.A. (2016) Travel burden and dentist bypass among dentally insured children. Journal of Public Health Dentistry 76(3). P.220-227.

McMahon, A.D., Blair, Y., McCall, D.R and Macperhson, L.M.D. (2010) The dental health of three-year-old children in Greater Glasgow, Scotland. British Dental Journal 209, E5.

McNamara, J., Cassells, R., Wicks, P and Vidyattama, Y. (2010) Children in housing disadvantage in Australia. Development of a small area index. Housing Studies 25(5). P. 625-646.

McNeil, L.H., Kreuter, M.W and Subramanian, S.V. (2006) Social environment and physical activity: a review of concepts and evidence. Social Science and Medicine 63(4). P.1011-1022.

Merler, S., Ajelli, M., Fumanelli, L and Vespignani, A. (2013) Containing the accidental laboratory escape of potential pandemic influenza viruses. BMC Medicine 11(252).

Metcalf, S.S., Northridge, M.E., Widener, M.J., Chakraborty, B., Marshall, S.E and Lamster, I,B. (2013) Modelling social dimensions of oral health among older adults in urban environments. Health Education and Behaviour 40(1S). P.63 – 73.

Meyer, S.P. (2014) A spatial assessment of dentist supply in Ontario, Canada. The Canadian Geographer 58(4). P.481-498.

Mitchell, R., Dorling, D and Shaw, M. (2002) Population production and modelling mortality - an application of geographic information systems in health inequalities research. Health & Place 8(1). P. 15-24.

Mitchell, R and F. Popham. (2008) Effect of exposure to natural environment on health inequalities: an observational population study. Lancet 372(9650). P.1655-1660.

Mobley, C., Marshall, T.A., Milgrom, P and Coldwell, S.E. (2009) The Contribution of Dietary Factors to Dental Caries and Disparities in Caries. Academic Pediatrics 9(6). P.410-414.

Montero, J., Albaladejo, A and Zalba, J.I. (2014) Influence of the usual motivation for dental attendance on dental status and oral health-related quality of life. Medicina Oral Patologia Oral y Cirugia Bucal 19(3). P.225-231.

Moriarity, C and Scheuren, F. (2001) Statistical matching: a paradigm for assessing the uncertainty in the procedure. Journal of Official Statistics 17(3). P.407-422.

Morland, K., Diez Roux, A.V and Wing, S. (2006) Supermarkets, Other Food Stores, and Obesity: The Atherosclerosis Risk in Communities Study. American Journal of Preventive Medicine 30(4). P.333-339.

Morland, K.B and Evenson, K.R. (2009) Obesity prevalence and the local food environment. Health and Place 15(2). P.491-495.

Morning, A. (2005) Ethnic classification in global perspective – a cross national survey of the 2000 Census round - https://unstats.un.org/unsd/demographic/sconcerns/popchar/Morning.pdf

Morrissey, K., Clarke, G., Ballas, D., Hynes, S and O'Donoghue, C. (2008) Examining Access to GP Services in Rural Ireland Using Microsimulation Analysis. Area 40(3). P.354-364.

Morrissey, K., Hynes, S., Clarke, G and O'Donoghue, C. (2010) Examining the factors associated with depression at the small area level in Ireland using spatial microsimulation techniques. Irish Geography 43(1). P. 1-22.

Morrissey, K., O'Donoghue, C., Clarke, G and Li, J.J. (2013a) Using simulated data to examine the determinants of acute hospital demand at the small area level. Geographical Analysis 45(1). P.49-76.

Morrissey, K., Clarke, G and O'Donoghue, C. (2013b) Linking static spatial microsimulation modelling to meso-scale models: The relationship between access to GP services and long-term illness. In Edwards, K and Tanton, R. (eds.) Spatial Microsimualtion: A Reference Guide for Users. Springer, London. P.127-143.

Morrissey, K., Clarke, G., Williamson, P., Daly, A and O'Donoghue, C. (2015) Mental illness in Ireland: Simulating its geographical prevalence and the role of access to services. Environment and Planning B: Urban Analytics and City Science 42(2). P.338-353.

Moynihan, P and Petersen, P.E. (2004) Diet, nutrition and the prevention of dental diseases. Public Health Nutrition 7(1A). P.210-226.

Moyses, S. J., Moyses, S. T., McCarthy, M and Sheiham, A. (2006) Intra-urban differentials in child dental trauma in relation to Healthy Cities policies in Curitiba, Brazil. Health & Place 12(1). P.48-64.

Moyses, S.T., Camilloti, A.G., Vetorello, M and Moyses, S.J. (2008) Spatial analysis of dental trauma in 12-year-old schoolchildren in Curitiba, Brazil. Dental Traumatology 24(4). P. 449-453.

Muirhead, V and Marcenes, W. (2004) An ecological study of caries experience, school performance and material deprivation in 5-year-old state primary school children. Community Dentistry and Oral Epidemiology 32(4). P.265-270.

Muirhead, V.E., Quinonez, C., Figueiredo, R and Locker, D. (2009) Predictors of dental care utilization among working poor Canadians. Community Dentistry and Oral Epidemiology 37(3). P.199-208.

Nasseh, K., Eisenberg, Y and Vujicic, M. (2017) Geographic access to dental care varies in Missouri and Wisconsin. Journal of Public Health Dentistry 77(3). P.197-206.

National Diet and Nutrition Survey. (2014) NDNS results from years 1 to 4 combined of the rolling programme for 2008 and 2009 to 2011 and 2012: report - https://www.gov.uk/government/statistics/national-diet-and-nutrition-survey-results-from-years-1-to-4-combined-of-the-rolling-programme-for-2008-and-2009-to-2011-and-2012

National Health Service Choices website. (2015) How to keep your teeth clean - http://www.nhs.uk/Livewell/dentalhealth/Pages/Teethcleaningguide.aspx

National Institute of Dental and Craniofacial Research. (2014) Dental caries (tooth decay) in adults (age 20-64) – Table 3: Adults, Severity of Decay Measured by Number of Permanent Teeth Affected - https://www.nidcr.nih.gov/DataStatistics/FindDataByTopic/DentalCaries/DentalCariesAdults20to64.htm

Newton, J.T and Bower, E. J. (2006) The social determinants of oral health: new approaches to conceptualizing and researching complex causal networks. Community Dentistry and Oral Epidemiology 33(1). P. 25-34.

Ng, D.M and Jeffery, R.W. (2003) Relationships between perceived stress and health behaviors in a sample of working adults. Health Psychology 22(6). P.638-642.

Ngo, T.A and See, L. (2012) Calibration and Validation of Agent-Based Models of Land Cover Change. In Heppenstall, A., Crooks, A., See, L and Batty, M. (2012) Agent-Based Models of Geographical Systems. P.181-199. Springer.

Nicolau, B., Marcenes, W., Bartley, M and Sheiham, A. (2005) Association between socio-economic circumstances at two stages of life and adolescents' oral health status. Journal of Public Health Dentistry 65(1). P. 14-20.

Nkosi, T.M., Parent, M.E., Siemiatycki, J., Pintos, J and Rousseau, M.C. (2011) Comparison of indicators of material circumstances in the context of an epidemiological study. BMC Medical Research Methodology 11(108).

Noble, D., Smith, D., Mathur, R., Robson, J and Greenhalgh, T. (2012) Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. BMJ Open 2(1).

NOMIS – Official Labour Market Statistics. (2011) - https://www.nomisweb.co.uk/Census/2011

Norman, P. (1999) Putting Iterative Proportional Fitting on the Researcher's Desk. Working Paper (1999), University of Leeds - http://eprints.whiterose.ac.uk/5029/1/99-3.pdf

Office for National Statistics (2009) Income: Model-Based Estimates at MSOA Level, 2007/08 - http://www.neighbourhood.statistics.gov.uk/dissemination/datasetList.do?JSAllowed=true&Function=&%24ph=60_61_60&CurrentPageId=60&step=1&CurrentTreeIndex=-2&searchString=income&datasetFamilyId=266&Next.x=36&Next.y=18

Office for National Statistics (2010) SOC2010 volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010) - http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

Office for National Statistics – Neighbourhood Statistics (2011) 2011 Census: Aggregate data (England and Wales). Downloaded from http://www.neighbourhood.statistics.gov.uk/dissemination/

Office for National Statistics (2011a) 2011 Census of Population - http://www.ons.gov.uk/ons/rel/Census/2011-Census/index.html

Office for National Statistics (2011b) Neighbourhood Statistics website - http://www.neighbourhood.statistics.gov.uk/dissemination/Download1.do

Office for National Statistics (2011c) Area Classifications - http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/about-the-area-classifications/index.html

Office for National Statistics (2014) Output areas - http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/Census/output-area--oas-/index.html

Office for National Statistics (2015) House Price Statistics for Small Areas - http://www.neighbourhood.statistics.gov.uk/dissemination/datasetList.do?JSAllowed=true&Function=&%24ph=60&CurrentPageId=60&step=1&CurrentTreeIndex=-2&searchString=house+price&datasetFamilyId=2793&Next.x=28&Next.y=15

O'Hanlon, S., Forster, D.P and Lowry, R.J. (1997) Oral cancer in the North-East of England: incidence, mortality trends and the link with material deprivation. Community Dentistry and Oral Epidemiology 25. P.371-376.

O'Neil, C.A and Sattenspiel, L. (2010) Agent-Based Modeling of the Spread of the 1918-1919 Flu in Three Canadian Fur Trading Communities. American Journal of Human Biology 22(6). P.757-767.

Ojima, M., Amano, A and Kurata, S. (2015) Relationship between decayed teeth and metabolic syndrome: data from 4716 middle-aged male Japanese employees. Journal of Epidemiology 25(3). P.204-211.

Olivella-Rosell, P., Villafafila-Robles, R., Sumper, A and Bergas-Jane, J. (2015) Probabilistic Agent-Based Model of Electric Vehicle Charging Demand to Analyse the Impact on Distribution Networks. Energies 9(5). P.4160-4187.

Openshaw, S and Taylor, P.J. (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In Wrigley, N. (eds.) Statistical applications in spatial sciences. London: Pion. P.127–144.

Openshaw, S. (1984) Ecological fallacies and the analysis of areal Census data. Environment and Planning A 16(1). P.17-31.

OpenStreetMap (2016) – OpenStreetMap Official Website - https://www.openstreetmap.org/#map=6/54.910/-3.432

Orcutt, G. (1957) A new type of socio-economic system. Review of Economics and Statistics 39(2). P.116-123.

Ordinance Survey (2013) – The National Grid - http://www.ordnancesurvey.co.uk/resources/maps-and-geographic-resources/the-national-grid.html

O'Reilly, D and Stevenson, M. (2003) Selective migration from deprived areas in Northern Ireland and the spatial distribution of inequalities: Implications for monitoring health and inequalities in health. Social Science and Medicine 57. P.1455-1462.

O'Reilly, G., O'Reilly, D., Rosato, M and Connolly, S. (2007) Urban and rural variations in morbidity and mortality in Northern Ireland. BMC Public Health 7(123).

Ostberg, A.L., Halling, A and Lindblad, U. (1999) Gender differences in knowledge, attitude, behaviour and percieved oral health among adolescents. Acta Odontologica Scandinavica 57(4). P.231-236.

Owen, G., Harris, R and Jones, K. (2016) Under examination: multilevel models, geography and health research. Progress in Human Geography 40(3). P.394-412.

Pampalon, R., Duncan, C., Subramanian, S.V and Jones, K. (1999) Geographies of health perception in Quebec: a multilevel perspective. Social Science and Medicine 48 (10). P.1483-1490.

Pattussi, M.P., Marcenes, W., Croucher, R and Sheiham, A. (2001) Social deprivation, income inequality, social cohesion and dental caries in Brazilian school children. Social Science and Medicine 53. P. 915-925.

Pattussi, M.P., Hardy, R and Sheiham, A. (2006) The potential impact of neighborhood empowerment on dental caries among adolescents. Community Dentistry and Oral Epidemiology 34(5). P.344-350.

Peacock, M and Bissell, P. (2011) Reducing Health Inequalities Implementation Theme. Briefing Paper 3: The social determinants of health inequalities: implications for research and practice. National Institute for Health Research - http://clahrc-sy.nihr.ac.uk/images/health%20inequalities/CLAHRC%20SY%20Briefing%20Paper%203.pdf

Pearce, J., Witten, K and Bartie, P. (2006) Neighbourhoods and health: a GIS approach to measuring community resource accessibility. Journal of Epidemiology and Community Health 60(5). P.389-395.

Pearce, J., Hiscock, R., Blakely, T and Witten, K. (2009) A national study of the association between neighbourhood access to fast-food outlets and the diet and weight of local residents. Health & Place 15(1). P. 193-197.

Pearce, J. R., Richarsdson, A.E., Mitchell, R.J and Shortt, N.K. (2010). Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. Transactions of the Institute of British Geographers 35(4). P.522-539.

Pearce, J., Barnett, R and Moon, G. (2012) Sociospatial inequalities in health-related behaviours: pathways linking place and smoking. Progress in Human Geography 36(1). P.3-24.

Pearson, K. (1896) Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London, Series A 187. P.253–318.

Pearson, A.L., Pearce, J and Kingham, S. (2013) Deprived yet healthy: Neighbourhood-level resilience in New Zealand. Social Science and Medicine 91. P.238-245.

Peltzer, K and Pengpid, S. (2014) Oral health behaviour and social and health factors in university students from 26 low, middle and high income countries. International Journal of Environmental Research and Public Health 11(12). P.12247-12260.

Pereira, S.M., Ambrosano, G.M., Cortellazzi, K.L., Tagliaferro, E.P., Vettorazzi, C.A., Ferraz, S.F., Meneghim, M.C and Pereira, A.C. (2010) Geographic Information Systems (GIS) in Assessing Dental Health. International Journal of Environmental Research and Public Health 7(5). P.2423-2436.

Pereira, S.M., Pardi, V., Cortellazzi, K.L., Ambrosano, G.M.B., Vettorazzi, C.A., Ferrez, S.F.B., Meneghim, M.C and Pereira, A.C. (2014) Geographic information system and multilevel analysis: gingival status among 12-year-old schoolchildren in Sao Paulo, Brazil. Revista Panamericana de Salud Pública 35(2). P.136-143.

Perera, I., Kruger, E and Tennant, M. (2010) Rural public dental clinic distribution in three states of Australia: using spatial analysis to inform management and planning of services. Asia Pacific Journal of Health Management 5(1). P. 40-46.

Petersen, P.E and Lennon, M.A. (2004) Effective use of fluorides for the prevention of dental caries in the 21st century: the WHO approach. Community Dentistry and Oral Epidemiology 32. P.319-321.

Pickett, K.E and Pearl, M. (2001) Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. Journal of Epidemiology and Community Health 55(2). P.111-122.

Pihlstrom, B.L., Michalowicz, B.S and Johnson, N.W. (2005) Periodontal diseases. The Lancet 366(9499). P.1809-1820.

Pitts, N.B. (2004) Modern concepts of caries measurement. Journal of Dental Research 83(Special issue C). P.43-47.

Plane, J and Klodawsky, F. (2013) Neighbourhood amenities and health: examining the significance of a local park. Social Science and Medicine 99. P.1-8.

Polhill, G.J., Parker, D., Brown, D and Grimm, V (2008) Using the ODD Protocol for Describing Three Agent-Based Social Simulation Models of Land-Use Change. Journal of Artificial Societies and Social Simulation 11(23).

Popay, J., Thomas, C., Williams, G., Bennett, S., Gatrell, A and Bostock, L. (2003) A proper place to live: health inequalities, agency and the normative dimensions of space. Social Science and Medicine 57 (1). P.55-69.

Potter, M.A., Brown, S.T., Cooley, P.C., Sweeney, P.M., Hershey, P.B., Gleason, S.M., Lee, B.Y., Keane, C.R., Grefenstette, J and Burke, D.S. (2012) School closure as an influenza mitigation strategy: how variations in legal authority and plan criteria can alter the impact. BMC Public Health 12(977).

Poulton, R., Caspi, A., Milne, B.J., Thompson, W.M., Taylor, A., Sears, M.R and Moffitt, T.E. (2002) Association between children's experience of socioeconomic disadvantage and adult health: a life-course study. The Lancet 360(9346). P.1640-1645.

Poutanen, R., Lahti, S., Tolvanen, M and Hausen, H. (2006) Parental influence on children's oral health-related behaviour. Acta Odontologica Scandinavica 64(5). P.286-292.

Prendergast, M.J., Beal, J.F and Williams, S.A. (1997) The relationship between deprivation, ethnicity and dental health in 5-year-old children in Leeds, UK. Community Dental Health 14(1). P.18-21.

Proctor, K.L., Clarke, G.P., Ransley, J.K and Cade, J. (2008) Micro-level analysis of childhood obesity, diet, physical activity, residential socioeconomic and social capital variables: where are the obesogenic environments in Leeds? Area 40(3). P. 323-340.

Propper, C. (2000) The demand for private healthcare in the UK. Journal of Health Economics 19(6). P.855-876.

Public Health England (2015) South Yorkshire and Bassetlaw Oral Health Needs Assessment 2015 - https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/463067/South_Yorks___Bassetlaw_oral_health_needs_assessment.pdf

Public Health England (2016) National Dental Epidemiology Programme for England: oral health survey of five-year-old children 2015. A report on the prevalence and severity of dental decay - http://www.nwph.net/dentalhealth/14_15_5yearold/14_15_16/DPHEP%20for%20England%20OH%20Survey%205yr%202015%20Report%20FINAL%20Gateway%20approved.pdf

Quantum Geographic Information Systems – QGIS. Version 2.8 – Wien - QGIS Development Team (2015). Open Source Geospatial Foundation Project. http://www.qgis.org/en/site/

R Core Team (2015) R: A Language and Environment for Statistical Computing. Vienna, Austria - https://www.r-project.org/

Reibel, J. (2003) Tobacco and oral diseases – update on the evidence, with recommendations. Medical Principles and Practice 12(supplement 1). P.22-32.

Reid, B.C., Hyman, J.J and Macek, M.D. (2004) Race/ethnicity and untreated dental caries: the impact of material and behavioural factors. Community Dentistry and Oral Epidemiology 32(5). P.329-336.

Relph, E. (1976) Place and placelessness. London. Pion.

Richards, W., Ameen, J., Coll, A.M and Higgs, G. (2005) Reasons for tooth extraction in four general dental practices in South Wales. British Dental Journal 198(5). P.275-278.

Riley, J.L., Gilbert, G.H and Heft, M.W. (2006) Dental attitudes: proximal basis for oral health disparities in adults. Community Dentistry and Oral Epidemiology 34(4). P.289-298.

Riva, M., Gauvin, L and Barnett, T. (2007) Toward the next generation of research into small area effects in health a synthesis of multilevel investigations published since July 1998. Journal of Epidemiology and Community Health 61(10). P.853-861.

Riva, M and Smith, D.M. (2012) Generating small-area prevalence of psychological distress and alcohol consumption: validation of a spatial microsimulation method**.** Social Psychiatry and Psychiatric Epidemiology 47(5). P. 745-755.

Roberts-Thomson, K and Stewart, J.F. (2008) Risk indicators of caries experience among young adults. Australian Dental Journal 53. P.122-127.

Robinson, W.S. (1950) Ecological Correlations and the Behavior of Individuals. American Sociological Review 15(3). P.351-357.

Rocha, C.M., Kruger, E., McGuire, S and Tennant, M. (2013) The geographic distribution of patients seeking emergency dental care at the Royal Dental Hospital of Melbourne, Australia. Community Dental Health 30(3). P.149-154.

Rodgers, W.L. (1984) An evaluation of statistical matching. Journal of Business and Economic Statistics 2(1). P.91-102.

Roscigno, V.J., Tomaskovic-Devey, D and Crowley, M. (2006) Education and the inequalities of place. Social Forces 84(4). P. 2121-2145.

Rose, D and O’Reilly, K. (1998) The ESRC Review of Government Social Classifications. London and Swindon: ONS and ESRC.

Roudsari, M.S., Shariatpanahi, S.P., Ahmady, A.E and Khoshnevisan, M.H. (2016) Agent-based modeling: An innovative opportunity for population-based oral health promotion. Journal of Dentistry of Tehran University of Medical Sciences 13(2). P.73-76.

Rouxel, P.L., Heilmann, A., Aida, J., Tsakos, G and Watt, R.G. (2015) Social capital: theory, evidence, and implications for oral health. Community Dentistry and Oral Epidemiology 43(2). P.97-105.

Ryan, J., Maoh, H and Kanaroglou, P. (2009) Population synthesis: comparing the major techniques using a small, complete population of firms. Geographical Analysis 42(2). P.181-203.

Sabbah, W., Tsakos, G., Chandola, T., Sheiham, A and Watt, R, G. (2007) Social gradients in oral and general health. Journal of Dental Research 86(10). P. 992-996.

Sabbah, W., Tsakos, G., Sheiham, A and Watt, R. G. (2009) The role of health-related behaviors in the socioeconomic disparities in oral health. Social Science and Medicine 68. P.298-303.

Sabbah, W., Suominen, A.L., Vehkalahti, M.M., Aromaa, A and Bernabe, E. (2015) The role of behaviour in inequality in increment of dental caries among Finnish adults. Caries Research 49(1). P.34-40.

Sadeghipour, M., Shariatpanahi, P., Jafari, A., Khosnevisan, M.H and Ahmady, A.E. (2016) Oscillatory patterns in the amount of demand for dental visits: An agent-based modeling approach. Journal of Artificial Societies and Social Simulation 19(3).

Sadeghipour, M., Khoshnevisan, M.H., Jafari, A and Shariatpanahi, P. (2017) Friendship network and dental brushing behaviour among middle school students: an agent-based modeling approach. Plos One 12(1).

Saldunaite, K., Bendoraitiene, E.A., Slabsinskiene, E., Vasiliauskiene, V.A and Zubiene, J. (2014) The role of parental education and socioeconomic status in dental caries prevention among Lithuanian children. Medicina 50(3). P.156-161.

Saman, D.M., Arevalo, O and Johnson, A.O. (2009) The dental workforce in Kentucky: current status and future needs. Journal of Public Health Dentistry 70(3). P.188-196.

Saman, D.M., Johnson, A.O., Arevalo, O and Odi, A. (2011) Geospatially illustrating regional-based oral health disparities in Kentucky. Public Health Reports 126(4). P.612-618.

Sanders, A.E and Spencer, A.J. (2004) Social inequality in perceived oral health among adults in Australia. Australian and New Zealand Journal of Public Health 28(2). P.159-166.

Sanders, A.E and Spencer, A.J. (2005) Why do poor adults rate their oral health poorly? Australian Dental Journal 50(3). P.161-167. Brazilian Journal of Epidemiology 17(Supplement 2). P.15-28.

Sanders, A.E., Spencer, A.J and Slade, G.D. (2006) Evaluating the role of dental behaviour in oral health inequalities. Community Dentistry and Oral Epidemiology 34(1). P.71-79.

Santiago, B.M., Valenca, A.M and Vettore, M.V. (2014) The relationship between neighbourhood empowerment and dental caries experience: a multilevel study in adolescents and adults. Revista Brasileira de Epidemiologia 17 (suppl 2). P.15-28.

ScapeToad – cartogram software by the Choros laboratory (2008) - https://scapetoad.choros.ch/

Schelling, T.C. (1971) Dynamic models of segregation. Journal of Mathematical Sociology 1. P.143–186.

Schwendicke, F., Dorfer, C.E., Schlattmann, P., Foster Page, L., Thomson, W.M and Paris, S. (2015) Socioeconomic inequality and caries: a systematic review and meta-analysis. Journal of Dental Research 94(1). P.10-18.

Selwitz, R.H., Ismail, A.I and Pitts, N.B. (2007) Dental Caries. The lancet 369(9555). P.51-59.

Shearer, D.M., Thomson, W.M., Broadbent, J.M and Poulton, R. (2010) Maternal oral health predicts their children's caries experience in adulthood. Journal of Dental Research 90(5). P.672-677.

Sheffield City Council (2011) 2011 Census for Sheffield - https://www.sheffield.gov.uk/content/sheffield/home/your-city-council/census-2011.html

Sheiham, A., Maizels, J., Cushing, A and Holmes, J. (1985) Dental attendance and dental status. Community Dentistry and Oral Epidemiology 13(6). P.304-309.

Sheiham, A (2001) Dietary effects on dental diseases. Public Health Nutrition 4(2b). P.569-591.

Sheiham, A. (2006) Dental caries affects body weight, growth and quality of life in pre-school children. British Dental Journal 201(10). P. 625-626.

Shendarkar, A., Vasudevan, K., Lee, S and Son, Y.J. (2008) Crowd simulation for emergency response using BDI agents based on immersive virtual reality. Simulation Modelling Practice and Theory 16(9). P.1415-1429.

Shulman, H., Birkin, M and Clarke, G.P. (2015) A comparison of small-area hospitalisation rates, estimated morbidity and hospital access. Health and Place 36.

Siahpush, M and Carlin, J.B. (2005) Financial stress, smoking cessation and relapse: results from a prospective study of an Australian national sample. Addiction 101(1). P. 121-127.

SimSalud website (2017) - http://www.simsalud.org/

Singh, A., Rouxel, P., Watt R.G and Tsakos, G. (2013) Social inequalities in clustering of oral health related behaviours in a national sample of British adults. Preventive Medicine 57(2). P.102-106.

Sisson, K.L. (2007) Theoretical explanations for social inequalities in oral health. Community Dentistry and Oral Epidemiology 35(2). P.81-88.

Slack-Smith, L. (2012) How population-level data linkage might impact on dental research. Community Dentistry and Oral Epidemiology 40(Suppl.2). P.90-94.

Slade, G.D., Spencer, A.J., Davies, M.J and Stewart, J.F. (1996) Influence of exposure to fluoridated water on socioeconomic inequalities in children's caries experience. Community Dentistry and Oral Epidemiology. Volume 24(2). P.89-100.

Sloggett, A and Joshi, H. (1994) Higher mortality in deprived areas – community or personal disadvantage. British Medical Journal 309(6967). P.1470-1474.

Smith, D.M., Clarke, G and Harland, K.H. (2007) SimHealth: estimating small area populations using deterministic spatial microsimulation in Leeds and Bradford. Working Paper (06/07), University of Leeds - http://eprints.whiterose.ac.uk/4490/1/Simhealth.pdf

Smith, S.J and Easterlow, D. (2005) The strange geography of health inequalities. Transactions of the institute of British Geographers 30(2). P.173-190.

Smith, D.M., Pearce, J.R and Harland, K. (2011) Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. Health and Place 17(2). P.618-624.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan, D and Plunkett, E. (2015) The English Indices of Multiple Deprivation 2015. Department for Communities and Local Government - https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Sourcepole (2016) OpenLayers Plugin, version 1.4.0 - https://plugins.qgis.org/plugins/openlayers_plugin/

Speybroeck, N., Van Malderen, C., Harper, S., Muller, B and Devleesschauwer, B. (2013) Simulation models for socioeconomic inequalities in health: a systematic review. International Journal of Environmental Research and Public Health 10. P.5750-5780.

Stahlnacke, K., Soderfeldt, B., Unell, L., Halling, A and Axtelius, B. (2003) Perceived oral health: changes over 5 years in one Swedish age-cohort. Community Dentistry and Oral Epidemiology 31(4). P.292-299.

Steele, J., Shen, J., Tsakos, G., Fuller, E., Morris, S., Watt, R., Guarnizo-Herreno, C and Wildman, J. (2015) The interplay between socioeconomic inequalities and clinical oral health. Journal of Dental Research 94(1). P.19-26.

Tada, A and Hanada, N. (2004) Sexual differences in oral health behaviour and factors associated with oral health behaviour in Japanese young adults. Health and Place 118(2). P.104-109.

Tanton, R., Vidyattama, Y., Nepal, B and McNamara, J. (2011) Small area estimation using a reweighting algorithm. Journal of the Royal Statistical Society 174 (Part 4). P.931-951.

Tanton, R and Edwards, K., L. (2013) Spatial Microsimualtion: A reference Guide for Users. Springer.

Tanton, R. (2014) A Review of Spatial Microsimulation Methods. International Journal of Microsimulation 7(1). P. 4-25.

Tanton, R., Williamson, P and Harding, A. (2014) Comparing two methods for reweighting a survey field to small area data. International Journal of Microsimulation 7(1). P. 76-99.

Taylor, P.J and Carmichael, C.L. (1980) Dental-health and the application of geographical methodology. Community Dentistry and Oral Epidemiology 8(3). P.117-122.

Taylor, S.E and Seeman, T. E. (1999) Psychosocial resources and the SES-health relationship. Annals of the New York Academy of Sciences 896. P. 210-225.

Tellez, M., Sohn, W., Burt, B.A and Ismail, A.I. (2006) Assessment of the relationship between neighbourhood characteristics and dental caries severity among low-income African-Americans: A multilevel approach. Journal of Public Health Dentistry 66(1). P.30-36.

Teweldemedhin, E., Marwala, T and Mueller, C. (2004) Agent-based modelling: A case study in HIV epidemic. Fourth International Conference on Hybrid Intelligent Systems. P.154-159.

The British Fluoridation Society (2012) One in a million: the facts about water fluoridation – The extent of water fluoridation - http://www.bfsweb.org/onemillion/onemillion.htm

Thiele, J.C and Grimm, V. (2010) NetLogo meets R: Linking agent-based models with a toolbox for their analysis. Environmental Modelling and Software 25(8). P. 972-974.

The Scottish Government (2006) Scottish Index of Multiple Deprivation - http://www.scotland.gov.uk/Publications/2006/10/13142739/0

The Logo foundation - http://el.media.mit.edu/logo-foundation/

The Welsh Indices of Multiple Deprivation (2011) StatsWales - https://statswales.wales.gov.uk/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/Archive/WIMD-2011

Thomas, B., Pritchard, J., Ballas, D., Vickers, D and Dorling, D. (2009) A Tale of Two Cities – The Sheffield Project. Social and Spatial Inequalities Research group - http://sasi.group.shef.ac.uk/research/sheffield/

Thomson, W.M., Poulton, R., Milne, B.J., Caspi, A., Broughton, J.R and Ayers, K.M. (2004) Socioeconomic inequalities in oral health in childhood and adulthood in a birth cohort. Community Dentistry and Oral Epidemiology 32(5). P.345-353.

Thomson, W.M., Williams, S.M., Broadbent, J.M., Poulton, R and Locker, D. (2010) Long-term dental visiting patterns and adult oral health. Journal of Dental Research 89(3). P.307-311.

Thrift, N. J. (1983) On the determination of social action in space and time. Environment and Planning D: Society and Space 1: P.23-57.

Thrift, N.J. (2003) Space: the fundamental stuff of human geography. In Clifford, N., Holloway, S., Rice, S.P and Valentine, G. (2003) Key concepts in Geography. Sage Publishing. P.95-107.

Tickle, M., Williams, N., Jenner, T and Blinkhorn, A. (1999) The effects of socioeconomic status and dental attendance on dental caries experience, and treatment patterns in 5-year-old children. British Dental Journal 186(3). P.135-137.

Tickle, M., Moulding, G., Milson, K and Blinkhorn, A. (2000) Dental caries, contact with dental services and deprivation in young children: Their relationship at a small area level. British Dental Journal 189(7). P.376-379.

Tickle, M., Milson, K.M., Humphris, G.M and Blinkhorn, A.S. (2003) Parental attitudes to the care of the carious primary dentition. British Dental Journal 195(8). P.451-455.

Timonen, P., Niskanen, M., Suominen-Taipale, L., Jula, A., Knuuttila, M and Ylostalo, P. (2010) Metabolic syndrome, periodontal infection, and dental caries. Journal of Dental Research 89(10). P.1068-1073.

Tisue, S and Wilensky, U. (2004) NetLogo: Design and implementation of a multi-agent modelling environment. Center for Connected Learning and Computer-Based Modeling Northwestern University, Evanston, Illinois - https://ccl.northwestern.edu/papers/2013/netlogo-agent2004c.pdf

Tomintz, M.N., Clarke, G.P and Rigby, J.E. (2008) The geography of smoking in Leeds: estimating individual smoking rates and the implications for the location of stop smoking services. Area 40(3). P. 341-353.

Tomintz, M., Kosar, B and Clarke, G. (2016) smokeSALUD: exploring the effect of demographic change on the smoking prevalence at municipality level in Austria. International Journal of Health Geographics 15(36).

Torrens, P.M and McDaniel, A.W. (2013) Modeling geographic behaviour in riotous crowds. Annals of the Association of America Geographers 103(1). P.20-46.

Townsend, P., Phillimore, P and Beattie, A. (1988) Health and Deprivation: Inequality and the North. London: Croom Helm.

Tracy, J.A., Bevins, S.N., VandeWoude, S and Crooks, K.R. (2014) An agent-based movement model to assess the impact of landscape fragmentation on disease transmission. Ecosphere 5(9). Article 119.

Tsakos, G., Sheiham, A., Lliffe, S., Kharicha, K., Harari, D., Swift, C.G., Gillman, G and Stuck, A.E. (2009) The impact of educational level on oral health-related quality of life in older people in London. European Journal of Oral Sciences 117(3). P. 286-292.

Tuan, Y.F. (1977) Space and place: The perspective of experience. University of Minnesota Press, Minneapolis.

Tully, S., Cojocaru, M and Bauch, C.T. (2013) Coevolution of risk perception, sexual behaviour, and HIV transmission in an agent-based model. Journal of Theoretical Biology 337. P. 125-133.

Tunstall, R., Bevan, M., Bradshaw, J., Croucher, K., Duffy, S., Hunter, C., Jones, A., Rugg, J., Wallace, A and Wilcox, S. (2013) Joseph Rowntree Foundation – The links between housing and poverty: an evidence review - https://www.york.ac.uk/media/chp/documents/2013/poverty-housing-options-full.pdf

Turrell, G., Sanders, A.E., Slade, G.D., Spencer, A.J and Marcenes, W. (2007) The independent contribution of neighborhood disadvantage and individual-level socioeconomic position to self-reported oral health: a multilevel analysis. Community Dentistry and Oral Epidemiology 35(3). P.195-206.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. Adult Dental Health Survey: 2009. 3rd Edition. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. British Social Attitudes Survey: 2013. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. The General Lifestyle Survey: 2011. 13th Edition. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. The Living Costs and Food Survey: 2012. 12th Edition. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. The Health Survey for England: 2012. 21st Edition. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. Poverty and Social Exclusion Living Standards Survey, 2012. Colchester, Essex: UK Data Archive.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research. Understanding Society: Wave 4, 2012-2013. 4th Edition. Colchester, Essex: UK Data Archive.

UK Data Service (2011a) Census Microdata - https://Census.ukdataservice.ac.uk/get-data/microdata

UK Data Service (2011b) Census boundary data - https://Census.ukdataservice.ac.uk/get-data/boundary-data

UK Data Service Census Support (2016) - http://Census.edina.ac.uk/bds.html

Vallee, J., Le Roux, G., Chaix, B., Kestens, Y and Chauvin, P. (2015) The 'constant size neighbourhood trap' in accessibility and health studies. Urban Studies 52(2). P.338-357.

Vanobberge, J.N., Martens, L.C., Lesaffre, E and Declerck, D. (2001) Parental occupational status related to dental caries experience in 7-year-old children in Flanders (Belgium). Community Dental Health 18(4). P. 256-262.

Vellappally, S., Fiala, Z., Smejkalova, J., Jacob, V and Shriharsha, P. (2007) Influence of tobacco use in dental caries development. Central European Journal of Public Health 15(3). P. 116-121.

Walby, S. (1990) Theorizing Patriarchy. Basil Blackwell, Oxford.

Walker, R.O., Beagrie, G.S., Jones, M.S., Liptrot, T.H., Naylor, M.N and Yardley, R.M. (1965a) Dentistry in the United Kingdom. A survey of the present position and future trends. British Dental Journal 118. P.510-515.

Walker, R.O., Beagrie, G.S., Jones, M.S., Liptrot, T.H., Naylor, M.N and Yardley, R.M. (1965b) Dentistry in the United Kingdom. A survey of the present position and future trends. British Dental Journal 119. P.139-146.

Wallace, B.B and Macentee, M.I. (2012) Access to dental care for low-income adults: perceptions of affordability, availability and acceptability. Journal of Community Health 37(1). P.32-29.

Wang, H., Northridge, M.E., Kunzel, C., Zhang, Q., Kum, S.S., Gilbert, J.L., Jin, Z and Metcalf, S.S. (2016) Modeling social capital as dynamic networks to promote access to oral healthcare. Social Computing, Behavioural-Cultural Modeling, and Prediction 9708. P.117-130.

Warren, J.J., Weber-Gasparoni, K., Marshall, T.A., Drake, D.R., Dehkordi-Vakil, F., Dawson, D.V and Tharp, K.M. (2009) A longitudinal study of dental caries risk among very young low SES children. Community Dentistry and Oral Epidemiology 37(2). P.116-122.

Warren, K.R., Postolache, T.T., Groer, M.E., Pinjari, O., Kelly, D.L and Reynolds, M.A. (2014) Role of chronic stress and depression in periodontal diseases. Periodontol 2000 64(1). P.127-138.

Watt, R.G. (2002) Emerging theories into the social determinants of health: implications for oral health promotion. Community Dentistry and Oral Epidemiology 30(4). P.241-247.

Watt, R.G. (2007) From victim blaming to upstream action: tackling the social determinants of oral health inequalities. Community Dentistry and Oral Epidemiology 35(1). P.1-11.

Watt, R.G. (2012) Social determinants of oral health inequalities: implications for action. Community Dentistry and Oral Epidemiology 40(Suppl.2). P.44-48.

Watt, R and Sheiham, A. (1999) Inequalities in oral health: a review of the evidence and recommendations for action. British Dental Journal 187(1). P.6-12.

Wilensky, U. (1999) NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL - https://ccl.northwestern.edu/netlogo/

Willie-Stephens, J., Kruger, E and Tennant, M. (2014) Public and private dental services in NSW: a geographic information system analysis of access to care for 7 million Australians. New South Wales Public Health Bulletin 24(4). P.164-170.

Wilkinson, R.G. (1996) Unhealthy societies. The afflictions of inequality. Routledge, London.

Wilkinson, R.G. (1997) Health inequalities: relative or absolute material standards? British Medical Journal 314. P.591-595.

Wilkinson, R and Marmot, M. (2003) Social determinants of health: The solid facts. World Health Organisation: Geneva - http://www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf

Wilkinson, R.G and Pickett, K.E. (2006) Income inequality and population health: A review and explanation of the evidence. Social Science and Medicine 62. P.1768-1784.

Williams, N.J., Whittle, J.G and Gatrell, A.C. (2002) The relationship between socio-demographic characteristics and dental health knowledge and attitudes of parents with young children. British Dental Journal 193(11). P.651-654.

Williamson, P., Birkin, M and Rees, P.H. (1998) The estimation of population microdata by using data from small area statistics and samples of anonymised records. Environment and Planning A 30(5). P. 785-816.

Williamson, P. (2013) An evaluation of two synthetic small-area microdata simulation methodologies: Synthetic reconstruction and combinatorial optimisation. In Tanton, R and Edwards, K. (eds.) Spatial microsimulation: A reference guide for users. Springer Netherlands. P.19-47.

Winn, D.M. (2001) Tobacco use and oral disease. Journal of Dental Education 65(4). P.306-312.

World Health Organisation (2003) The World Oral Health Report 2003 - http://www.who.int/oral_health/media/en/orh_report03_en.pdf

World Health Organisation. (2010) A Conceptual Framework for Action on the Social Determinants of Health. Social determinants of health discussion paper 2. Geneva -

http://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf

Wu, B.M., Birkin, M.H and Rees, P.H. (2008) A spatial microsimulation model with student agents. Computers, Environments and Urban Systems 32(6). P. 440-453.

Wu, B.M and Birkin, M.H. (2012) Agent-Based Extensions to a Spatial Microsimulation Model of Demographic Change. In Heppenstall, A. J., Crooks, A.T., See, L.M and Batty, M. (2012) Agent-Based Models of Geographical Systems. Springer. London.

Yang, Y., Diez Roux, A.V., Auchincloss, A.H., Rodriguez, D.A and Brown, D.G. (2011) A spatial agent-based model for the simulation of adults' daily walking within a city. American Journal of Preventative Medicine 40(3). P.353-361.

Yeung, C.A. (2008) A systematic review of the efficacy and safety of fluoridation. Evidence-Based Dentistry 9(2). P.39-43.

Xiang, X., Kennedy, R, and Madey, G. (2005) Verification and validation of agent-based scientific simulation models - http://www.cse.nd.edu/~nom/Papers/ADS019_Xiang.pdf

Zainab, U., Kruger, E and Tennant, M. (2015) Major metropolis rail system access to dental care for the retired and elderly: a high resolution geographic study of Sydney, Australia. Gerodontology 32. P.302-308.

Zero, D.T. (2006) Dentifrices, mouthwashes, and remineralization/caries arrestment strategies. BMC Oral Health 6(Suppl 1):S9.

Zhang, D.L., Giabbanelli, P.J., Arah, O.A and Zimmerman, F.J. (2014) Impact of Different Policies on Unhealthy Dietary Behaviors in an Urban Adult Population: An Agent-Based Simulation Model. American Journal of Public Health 104(7). P.1217-1222.

Zhou, Y. (2016) Path finding model using the A-star algorithm in NetLogo. Geospatial Computational Social Science blog - http://geospatialcss.blogspot.co.uk/2016/01/path-finding-model-using-a-star.html

Appendices

Appendix A – Post-hoc (Bonferroni) test results for Sheffield East

Multiple Comparisons

Measure: score

Bonferroni

(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	2	-1831.32	893492.297	1.000	-3170352.42	3166689.78
	3	-5513.49	893492.297	1.000	-3174034.59	3163007.61
	4	-3254.63	893492.297	1.000	-3171775.73	3165266.47
	5	-3250.92	893492.297	1.000	-3171772.02	3165270.18
	6	-5815.27	893492.297	1.000	-3174336.37	3162705.83
	7	-11506594.33*	893492.297	.000	-14675115.43	-8338073.24
	8	-11780651.30*	893492.297	.000	-14949172.40	-8612130.21
	9	-11890550.75*	893492.297	.000	-15059071.85	-8722029.66
	10	-11222652.09*	893492.297	.000	-14391173.18	-8054130.99
	11	-10365423.53*	893492.297	.000	-13533944.62	-7196902.43
	12	-8848737.26*	893492.297	.000	-12017258.36	-5680216.17
	13	-11249272.40*	893492.297	.000	-14417793.50	-8080751.30
	14	-11924842.80*	893492.297	.000	-15093363.90	-8756321.71
2	1	1831.32	893492.297	1.000	-3166689.78	3170352.42
	3	-3682.17	893492.297	1.000	-3172203.27	3164838.93
	4	-1423.31	893492.297	1.000	-3169944.41	3167097.79
	5	-1419.60	893492.297	1.000	-3169940.70	3167101.50
	6	-3983.95	893492.297	1.000	-3172505.05	3164537.15
	7	-11504763.01*	893492.297	.000	-14673284.11	-8336241.92
	8	-11778819.98*	893492.297	.000	-14947341.08	-8610298.89
	9	-11888719.43*	893492.297	.000	-15057240.53	-8720198.34
	10	-11220820.77*	893492.297	.000	-14389341.86	-8052299.67
	11	-10363592.21*	893492.297	.000	-13532113.30	-7195071.11
	12	-8846905.94*	893492.297	.000	-12015427.04	-5678384.85
	13	-11247441.08*	893492.297	.000	-14415962.18	-8078919.98
	14	-11923011.48*	893492.297	.000	-15091532.58	-8754490.39
3	1	5513.49	893492.297	1.000	-3163007.61	3174034.59
	2	3682.17	893492.297	1.000	-3164838.93	3172203.27
	4	2258.86	893492.297	1.000	-3166262.24	3170779.96
	5	2262.57	893492.297	1.000	-3166258.53	3170783.67

	6	-301.78	893492.297	1.000	-3168822.88	3168219.32
	7	-11501080.84*	893492.297	.000	-14669601.94	-8332559.75
	8	-11775137.81*	893492.297	.000	-14943658.91	-8606616.72
	9	-11885037.26*	893492.297	.000	-15053558.36	-8716516.17
	10	-11217138.60*	893492.297	.000	-14385659.69	-8048617.50
	11	-10359910.04*	893492.297	.000	-13528431.13	-7191388.94
	12	-8843223.77*	893492.297	.000	-12011744.87	-5674702.68
	13	-11243758.91*	893492.297	.000	-14412280.01	-8075237.81
	14	-11919329.31*	893492.297	.000	-15087850.41	-8750808.22
4	1	3254.63	893492.297	1.000	-3165266.47	3171775.73
	2	1423.31	893492.297	1.000	-3167097.79	3169944.41
	3	-2258.86	893492.297	1.000	-3170779.96	3166262.24
	5	3.71	893492.297	1.000	-3168517.39	3168524.81
	6	-2560.64	893492.297	1.000	-3171081.74	3165960.46
	7	-11503339.70*	893492.297	.000	-14671860.80	-8334818.61
	8	-11777396.67*	893492.297	.000	-14945917.77	-8608875.58
	9	-11887296.12*	893492.297	.000	-15055817.22	-8718775.03
	10	-11219397.46*	893492.297	.000	-14387918.55	-8050876.36
	11	-10362168.90*	893492.297	.000	-13530689.99	-7193647.80
	12	-8845482.63*	893492.297	.000	-12014003.73	-5676961.54
	13	-11246017.77*	893492.297	.000	-14414538.87	-8077496.67
	14	-11921588.17*	893492.297	.000	-15090109.27	-8753067.08
5	1	3250.92	893492.297	1.000	-3165270.18	3171772.02
	2	1419.60	893492.297	1.000	-3167101.50	3169940.70
	3	-2262.57	893492.297	1.000	-3170783.67	3166258.53
	4	-3.71	893492.297	1.000	-3168524.81	3168517.39
	6	-2564.35	893492.297	1.000	-3171085.45	3165956.75
	7	-11503343.41*	893492.297	.000	-14671864.51	-8334822.32
	8	-11777400.38*	893492.297	.000	-14945921.48	-8608879.29
	9	-11887299.83*	893492.297	.000	-15055820.93	-8718778.74
	10	-11219401.17*	893492.297	.000	-14387922.26	-8050880.07
	11	-10362172.61*	893492.297	.000	-13530693.70	-7193651.51
	12	-8845486.34*	893492.297	.000	-12014007.44	-5676965.25
	13	-11246021.48*	893492.297	.000	-14414542.58	-8077500.38
	14	-11921591.88*	893492.297	.000	-15090112.98	-8753070.79
6	1	5815.27	893492.297	1.000	-3162705.83	3174336.37
	2	3983.95	893492.297	1.000	-3164537.15	3172505.05
	3	301.78	893492.297	1.000	-3168219.32	3168822.88
	4	2560.64	893492.297	1.000	-3165960.46	3171081.74
	5	2564.35	893492.297	1.000	-3165956.75	3171085.45
	7	-11500779.06*	893492.297	.000	-14669300.16	-8332257.97

	8	-11774836.03*	893492.297	.000	-14943357.13	-8606314.94
	9	-11884735.48*	893492.297	.000	-15053256.58	-8716214.39
	10	-11216836.82*	893492.297	.000	-14385357.91	-8048315.72
	11	-10359608.26*	893492.297	.000	-13528129.35	-7191087.16
	12	-8842921.99*	893492.297	.000	-12011443.09	-5674400.90
	13	-11243457.13*	893492.297	.000	-14411978.23	-8074936.03
	14	-11919027.53*	893492.297	.000	-15087548.63	-8750506.44
7	1	11506594.33*	893492.297	.000	8338073.24	14675115.43
	2	11504763.01*	893492.297	.000	8336241.92	14673284.11
	3	11501080.84*	893492.297	.000	8332559.75	14669601.94
	4	11503339.70*	893492.297	.000	8334818.61	14671860.80
	5	11503343.41*	893492.297	.000	8334822.32	14671864.51
	6	11500779.06*	893492.297	.000	8332257.97	14669300.16
	8	-274056.97	893492.297	1.000	-3442578.06	2894464.13
	9	-383956.42	893492.297	1.000	-3552477.51	2784564.68
	10	283942.25	893492.297	1.000	-2884578.85	3452463.34
	11	1141170.80	893492.297	1.000	-2027350.29	4309691.90
	12	2657857.07	893492.297	.320	-510664.02	5826378.17
	13	257321.93	893492.297	1.000	-2911199.16	3425843.03
	14	-418248.47	893492.297	1.000	-3586769.56	2750272.63
8	1	11780651.30*	893492.297	.000	8612130.21	14949172.40
	2	11778819.98*	893492.297	.000	8610298.89	14947341.08
	3	11775137.81*	893492.297	.000	8606616.72	14943658.91
	4	11777396.67*	893492.297	.000	8608875.58	14945917.77
	5	11777400.38*	893492.297	.000	8608879.29	14945921.48
	6	11774836.03*	893492.297	.000	8606314.94	14943357.13
	7	274056.97	893492.297	1.000	-2894464.13	3442578.06
	9	-109899.45	893492.297	1.000	-3278420.54	3058621.65
	10	557999.22	893492.297	1.000	-2610521.88	3726520.31
	11	1415227.77	893492.297	1.000	-1753293.32	4583748.87
	12	2931914.04	893492.297	.122	-236607.05	6100435.14
	13	531378.90	893492.297	1.000	-2637142.19	3699900.00
	14	-144191.50	893492.297	1.000	-3312712.60	3024329.60
9	1	11890550.75*	893492.297	.000	8722029.66	15059071.85
	2	11888719.43*	893492.297	.000	8720198.34	15057240.53
	3	11885037.26*	893492.297	.000	8716516.17	15053558.36
	4	11887296.12*	893492.297	.000	8718775.03	15055817.22
	5	11887299.83*	893492.297	.000	8718778.74	15055820.93
	6	11884735.48*	893492.297	.000	8716214.39	15053256.58
	7	383956.42	893492.297	1.000	-2784564.68	3552477.51
	8	109899.45	893492.297	1.000	-3058621.65	3278420.54

	10	667898.67	893492.297	1.000	-2500622.43	3836419.76
	11	1525127.22	893492.297	1.000	-1643393.87	4693648.32
	12	3041813.49	893492.297	.081	-126707.60	6210334.59
	13	641278.35	893492.297	1.000	-2527242.74	3809799.45
	14	-34292.05	893492.297	1.000	-3202813.15	3134229.04
10	1	11222652.09*	893492.297	.000	8054130.99	14391173.18
	2	11220820.77*	893492.297	.000	8052299.67	14389341.86
	3	11217138.60*	893492.297	.000	8048617.50	14385659.69
	4	11219397.46*	893492.297	.000	8050876.36	14387918.55
	5	11219401.17*	893492.297	.000	8050880.07	14387922.26
	6	11216836.82*	893492.297	.000	8048315.72	14385357.91
	7	-283942.25	893492.297	1.000	-3452463.34	2884578.85
	8	-557999.22	893492.297	1.000	-3726520.31	2610521.88
	9	-667898.67	893492.297	1.000	-3836419.76	2500622.43
	11	857228.56	893492.297	1.000	-2311292.54	4025749.65
	12	2373914.82	893492.297	.811	-794606.27	5542435.92
	13	-26620.31	893492.297	1.000	-3195141.41	3141900.78
	14	-702190.72	893492.297	1.000	-3870711.81	2466330.38
11	1	10365423.53*	893492.297	.000	7196902.43	13533944.62
	2	10363592.21*	893492.297	.000	7195071.11	13532113.30
	3	10359910.04*	893492.297	.000	7191388.94	13528431.13
	4	10362168.90*	893492.297	.000	7193647.80	13530689.99
	5	10362172.61*	893492.297	.000	7193651.51	13530693.70
	6	10359608.26*	893492.297	.000	7191087.16	13528129.35
	7	-1141170.80	893492.297	1.000	-4309691.90	2027350.29
	8	-1415227.77	893492.297	1.000	-4583748.87	1753293.32
	9	-1525127.22	893492.297	1.000	-4693648.32	1643393.87
	10	-857228.56	893492.297	1.000	-4025749.65	2311292.54
	12	1516686.27	893492.297	1.000	-1651834.83	4685207.36
	13	-883848.87	893492.297	1.000	-4052369.97	2284672.22
	14	-1559419.27	893492.297	1.000	-4727940.37	1609101.82
12	1	8848737.26*	893492.297	.000	5680216.17	12017258.36
	2	8846905.94*	893492.297	.000	5678384.85	12015427.04
	3	8843223.77*	893492.297	.000	5674702.68	12011744.87
	4	8845482.63*	893492.297	.000	5676961.54	12014003.73
	5	8845486.34*	893492.297	.000	5676965.25	12014007.44
	6	8842921.99*	893492.297	.000	5674400.90	12011443.09
	7	-2657857.07	893492.297	.320	-5826378.17	510664.02
	8	-2931914.04	893492.297	.122	-6100435.14	236607.05
	9	-3041813.49	893492.297	.081	-6210334.59	126707.60
	10	-2373914.82	893492.297	.811	-5542435.92	794606.27

	11	-1516686.27	893492.297	1.000	-4685207.36	1651834.83
	13	-2400535.14	893492.297	.745	-5569056.23	767985.96
	14	-3076105.54	893492.297	.071	-6244626.64	92415.55
13	1	11249272.40*	893492.297	.000	8080751.30	14417793.50
	2	11247441.08*	893492.297	.000	8078919.98	14415962.18
	3	11243758.91*	893492.297	.000	8075237.81	14412280.01
	4	11246017.77*	893492.297	.000	8077496.67	14414538.87
	5	11246021.48*	893492.297	.000	8077500.38	14414542.58
	6	11243457.13*	893492.297	.000	8074936.03	14411978.23
	7	-257321.93	893492.297	1.000	-3425843.03	2911199.16
	8	-531378.90	893492.297	1.000	-3699900.00	2637142.19
	9	-641278.35	893492.297	1.000	-3809799.45	2527242.74
	10	26620.31	893492.297	1.000	-3141900.78	3195141.41
	11	883848.87	893492.297	1.000	-2284672.22	4052369.97
	12	2400535.14	893492.297	.745	-767985.96	5569056.23
	14	-675570.40	893492.297	1.000	-3844091.50	2492950.69
14	1	11924842.80*	893492.297	.000	8756321.71	15093363.90
	2	11923011.48*	893492.297	.000	8754490.39	15091532.58
	3	11919329.31*	893492.297	.000	8750808.22	15087850.41
	4	11921588.17*	893492.297	.000	8753067.08	15090109.27
	5	11921591.88*	893492.297	.000	8753070.79	15090112.98
	6	11919027.53*	893492.297	.000	8750506.44	15087548.63
	7	418248.47	893492.297	1.000	-2750272.63	3586769.56
	8	144191.50	893492.297	1.000	-3024329.60	3312712.60
	9	34292.05	893492.297	1.000	-3134229.04	3202813.15
	10	702190.72	893492.297	1.000	-2466330.38	3870711.81
	11	1559419.27	893492.297	1.000	-1609101.82	4727940.37
	12	3076105.54	893492.297	.071	-92415.55	6244626.64
	13	675570.40	893492.297	1.000	-2492950.69	3844091.50

Based on observed means.

The error term is Mean Square(Error) = 3991642427229.323.

*. The mean difference is significant at the .05 level.

Appendix B – Post-hoc (Bonferroni) test results for Sheffield West

Multiple Comparisons

Measure: score

Bonferroni

(I) model	(J) model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	2	-2592.29	252498.609	1.000	-898008.14	892823.56
	3	-6773.63	252498.609	1.000	-902189.48	888642.22
	4	-4482.97	252498.609	1.000	-899898.82	890932.88
	5	-2605.71	252498.609	1.000	-898021.56	892810.14
	6	-6270.96	252498.609	1.000	-901686.81	889144.89
	7	-25231260.11*	252498.609	.000	-26126675.96	-24335844.26
	8	-25251828.54*	252498.609	.000	-26147244.40	-24356412.69
	9	-25134794.56*	252498.609	.000	-26030210.42	-24239378.71
	10	-25195897.49*	252498.609	.000	-26091313.35	-24300481.64
	11	-25336312.02*	252498.609	.000	-26231727.87	-24440896.16
	12	-25303953.35*	252498.609	.000	-26199369.20	-24408537.49
	13	-25892112.99*	252498.609	.000	-26787528.84	-24996697.13
	14	-25281559.39*	252498.609	.000	-26176975.24	-24386143.54
2	1	2592.29	252498.609	1.000	-892823.56	898008.14
	3	-4181.34	252498.609	1.000	-899597.19	891234.51
	4	-1890.68	252498.609	1.000	-897306.53	893525.17
	5	-13.42	252498.609	1.000	-895429.27	895402.43
	6	-3678.67	252498.609	1.000	-899094.52	891737.18
	7	-25228667.82*	252498.609	.000	-26124083.67	-24333251.97
	8	-25249236.25*	252498.609	.000	-26144652.11	-24353820.40
	9	-25132202.27*	252498.609	.000	-26027618.13	-24236786.42
	10	-25193305.20*	252498.609	.000	-26088721.06	-24297889.35
	11	-25333719.73*	252498.609	.000	-26229135.58	-24438303.87
	12	-25301361.06*	252498.609	.000	-26196776.91	-24405945.20
	13	-25889520.70*	252498.609	.000	-26784936.55	-24994104.84
	14	-25278967.10*	252498.609	.000	-26174382.95	-24383551.25
3	1	6773.63	252498.609	1.000	-888642.22	902189.48
	2	4181.34	252498.609	1.000	-891234.51	899597.19
	4	2290.66	252498.609	1.000	-893125.19	897706.51
	5	4167.92	252498.609	1.000	-891247.93	899583.77
	6	502.67	252498.609	1.000	-894913.18	895918.52
	7	-25224486.48*	252498.609	.000	-26119902.33	-24329070.63
	8	-25245054.91*	252498.609	.000	-26140470.77	-24349639.06

	9	-25128020.93*	252498.609	.000	-26023436.79	-24232605.08
	10	-25189123.86*	252498.609	.000	-26084539.72	-24293708.01
	11	-25329538.39*	252498.609	.000	-26224954.24	-24434122.53
	12	-25297179.72*	252498.609	.000	-26192595.57	-24401763.86
	13	-25885339.36*	252498.609	.000	-26780755.21	-24989923.50
	14	-25274785.76*	252498.609	.000	-26170201.61	-24379369.91
4	1	4482.97	252498.609	1.000	-890932.88	899898.82
	2	1890.68	252498.609	1.000	-893525.17	897306.53
	3	-2290.66	252498.609	1.000	-897706.51	893125.19
	5	1877.26	252498.609	1.000	-893538.59	897293.11
	6	-1787.99	252498.609	1.000	-897203.84	893627.86
	7	-25226777.14*	252498.609	.000	-26122192.99	-24331361.29
	8	-25247345.57*	252498.609	.000	-26142761.43	-24351929.72
	9	-25130311.59*	252498.609	.000	-26025727.45	-24234895.74
	10	-25191414.52*	252498.609	.000	-26086830.38	-24295998.67
	11	-25331829.05*	252498.609	.000	-26227244.90	-24436413.19
	12	-25299470.38*	252498.609	.000	-26194886.23	-24404054.52
	13	-25887630.02*	252498.609	.000	-26783045.87	-24992214.16
	14	-25277076.42*	252498.609	.000	-26172492.27	-24381660.57
5	1	2605.71	252498.609	1.000	-892810.14	898021.56
	2	13.42	252498.609	1.000	-895402.43	895429.27
	3	-4167.92	252498.609	1.000	-899583.77	891247.93
	4	-1877.26	252498.609	1.000	-897293.11	893538.59
	6	-3665.25	252498.609	1.000	-899081.10	891750.60
	7	-25228654.40*	252498.609	.000	-26124070.25	-24333238.55
	8	-25249222.83*	252498.609	.000	-26144638.69	-24353806.98
	9	-25132188.85*	252498.609	.000	-26027604.71	-24236773.00
	10	-25193291.78*	252498.609	.000	-26088707.64	-24297875.93
	11	-25333706.31*	252498.609	.000	-26229122.16	-24438290.45
	12	-25301347.64*	252498.609	.000	-26196763.49	-24405931.78
	13	-25889507.28*	252498.609	.000	-26784923.13	-24994091.42
	14	-25278953.68*	252498.609	.000	-26174369.53	-24383537.83
6	1	6270.96	252498.609	1.000	-889144.89	901686.81
	2	3678.67	252498.609	1.000	-891737.18	899094.52
	3	-502.67	252498.609	1.000	-895918.52	894913.18
	4	1787.99	252498.609	1.000	-893627.86	897203.84
	5	3665.25	252498.609	1.000	-891750.60	899081.10
	7	-25224989.15*	252498.609	.000	-26120405.00	-24329573.30
	8	-25245557.58*	252498.609	.000	-26140973.44	-24350141.73
	9	-25128523.60*	252498.609	.000	-26023939.46	-24233107.75
	10	-25189626.53*	252498.609	.000	-26085042.39	-24294210.68

	11	-25330041.06*	252498.609	.000	-26225456.91	-24434625.20
	12	-25297682.39*	252498.609	.000	-26193098.24	-24402266.53
	13	-25885842.03*	252498.609	.000	-26781257.88	-24990426.17
	14	-25275288.43*	252498.609	.000	-26170704.28	-24379872.58
7	1	25231260.11*	252498.609	.000	24335844.26	26126675.96
	2	25228667.82*	252498.609	.000	24333251.97	26124083.67
	3	25224486.48*	252498.609	.000	24329070.63	26119902.33
	4	25226777.14*	252498.609	.000	24331361.29	26122192.99
	5	25228654.40*	252498.609	.000	24333238.55	26124070.25
	6	25224989.15*	252498.609	.000	24329573.30	26120405.00
	8	-20568.43	252498.609	1.000	-915984.29	874847.42
	9	96465.55	252498.609	1.000	-798950.31	991881.40
	10	35362.62	252498.609	1.000	-860053.24	930778.47
	11	-105051.91	252498.609	1.000	-1000467.76	790363.95
	12	-72693.24	252498.609	1.000	-968109.09	822722.62
	13	-660852.88	252498.609	.905	-1556268.73	234562.98
	14	-50299.28	252498.609	1.000	-945715.13	845116.57
8	1	25251828.54*	252498.609	.000	24356412.69	26147244.40
	2	25249236.25*	252498.609	.000	24353820.40	26144652.11
	3	25245054.91*	252498.609	.000	24349639.06	26140470.77
	4	25247345.57*	252498.609	.000	24351929.72	26142761.43
	5	25249222.83*	252498.609	.000	24353806.98	26144638.69
	6	25245557.58*	252498.609	.000	24350141.73	26140973.44
	7	20568.43	252498.609	1.000	-874847.42	915984.29
	9	117033.98	252498.609	1.000	-778381.87	1012449.83
	10	55931.05	252498.609	1.000	-839484.80	951346.90
	11	-84483.47	252498.609	1.000	-979899.33	810932.38
	12	-52124.80	252498.609	1.000	-947540.66	843291.05
	13	-640284.44	252498.609	1.000	-1535700.30	255131.41
	14	-29730.85	252498.609	1.000	-925146.70	865685.01
9	1	25134794.56*	252498.609	.000	24239378.71	26030210.42
	2	25132202.27*	252498.609	.000	24236786.42	26027618.13
	3	25128020.93*	252498.609	.000	24232605.08	26023436.79
	4	25130311.59*	252498.609	.000	24234895.74	26025727.45
	5	25132188.85*	252498.609	.000	24236773.00	26027604.71
	6	25128523.60*	252498.609	.000	24233107.75	26023939.46
	7	-96465.55	252498.609	1.000	-991881.40	798950.31
	8	-117033.98	252498.609	1.000	-1012449.83	778381.87
	10	-61102.93	252498.609	1.000	-956518.78	834312.92
	11	-201517.45	252498.609	1.000	-1096933.31	693898.40
	12	-169158.78	252498.609	1.000	-1064574.64	726257.07

	13	-757318.42	252498.609	.297	-1652734.28	138097.43
	14	-146764.83	252498.609	1.000	-1042180.68	748651.03
10	1	25195897.49*	252498.609	.000	24300481.64	26091313.35
	2	25193305.20*	252498.609	.000	24297889.35	26088721.06
	3	25189123.86*	252498.609	.000	24293708.01	26084539.72
	4	25191414.52*	252498.609	.000	24295998.67	26086830.38
	5	25193291.78*	252498.609	.000	24297875.93	26088707.64
	6	25189626.53*	252498.609	.000	24294210.68	26085042.39
	7	-35362.62	252498.609	1.000	-930778.47	860053.24
	8	-55931.05	252498.609	1.000	-951346.90	839484.80
	9	61102.93	252498.609	1.000	-834312.92	956518.78
	11	-140414.52	252498.609	1.000	-1035830.38	755001.33
	12	-108055.85	252498.609	1.000	-1003471.71	787360.00
	13	-696215.49	252498.609	.609	-1591631.35	199200.36
	14	-85661.90	252498.609	1.000	-981077.75	809753.96
11	1	25336312.02*	252498.609	.000	24440896.16	26231727.87
	2	25333719.73*	252498.609	.000	24438303.87	26229135.58
	3	25329538.39*	252498.609	.000	24434122.53	26224954.24
	4	25331829.05*	252498.609	.000	24436413.19	26227244.90
	5	25333706.31*	252498.609	.000	24438290.45	26229122.16
	6	25330041.06*	252498.609	.000	24434625.20	26225456.91
	7	105051.91	252498.609	1.000	-790363.95	1000467.76
	8	84483.47	252498.609	1.000	-810932.38	979899.33
	9	201517.45	252498.609	1.000	-693898.40	1096933.31
	10	140414.52	252498.609	1.000	-755001.33	1035830.38
	12	32358.67	252498.609	1.000	-863057.18	927774.52
	13	-555800.97	252498.609	1.000	-1451216.82	339614.88
	14	54752.63	252498.609	1.000	-840663.23	950168.48
12	1	25303953.35*	252498.609	.000	24408537.49	26199369.20
	2	25301361.06*	252498.609	.000	24405945.20	26196776.91
	3	25297179.72*	252498.609	.000	24401763.86	26192595.57
	4	25299470.38*	252498.609	.000	24404054.52	26194886.23
	5	25301347.64*	252498.609	.000	24405931.78	26196763.49
	6	25297682.39*	252498.609	.000	24402266.53	26193098.24
	7	72693.24	252498.609	1.000	-822722.62	968109.09
	8	52124.80	252498.609	1.000	-843291.05	947540.66
	9	169158.78	252498.609	1.000	-726257.07	1064574.64
	10	108055.85	252498.609	1.000	-787360.00	1003471.71
	11	-32358.67	252498.609	1.000	-927774.52	863057.18
	13	-588159.64	252498.609	1.000	-1483575.49	307256.21
	14	22393.96	252498.609	1.000	-873021.90	917809.81

13	1	25892112.99*	252498.609	.000	24996697.13	26787528.84
	2	25889520.70*	252498.609	.000	24994104.84	26784936.55
	3	25885339.36*	252498.609	.000	24989923.50	26780755.21
	4	25887630.02*	252498.609	.000	24992214.16	26783045.87
	5	25889507.28*	252498.609	.000	24994091.42	26784923.13
	6	25885842.03*	252498.609	.000	24990426.17	26781257.88
	7	660852.88	252498.609	.905	-234562.98	1556268.73
	8	640284.44	252498.609	1.000	-255131.41	1535700.30
	9	757318.42	252498.609	.297	-138097.43	1652734.28
	10	696215.49	252498.609	.609	-199200.36	1591631.35
	11	555800.97	252498.609	1.000	-339614.88	1451216.82
	12	588159.64	252498.609	1.000	-307256.21	1483575.49
	14	610553.60	252498.609	1.000	-284862.26	1505969.45
14	1	25281559.39*	252498.609	.000	24386143.54	26176975.24
	2	25278967.10*	252498.609	.000	24383551.25	26174382.95
	3	25274785.76*	252498.609	.000	24379369.91	26170201.61
	4	25277076.42*	252498.609	.000	24381660.57	26172492.27
	5	25278953.68*	252498.609	.000	24383537.83	26174369.53
	6	25275288.43*	252498.609	.000	24379872.58	26170704.28
	7	50299.28	252498.609	1.000	-845116.57	945715.13
	8	29730.85	252498.609	1.000	-865685.01	925146.70
	9	146764.83	252498.609	1.000	-748651.03	1042180.68
	10	85661.90	252498.609	1.000	-809753.96	981077.75
	11	-54752.63	252498.609	1.000	-950168.48	840663.23
	12	-22393.96	252498.609	1.000	-917809.81	873021.90
	13	-610553.60	252498.609	1.000	-1505969.45	284862.26

Based on observed means.

The error term is Mean Square(Error) = 318777736873.749.

*. The mean difference is significant at the .05 level.