Socioeconomic Inequalities in Risk of and Exposure to Gastrointestinal Infections in the UK

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by

Natalie Louise Adams

December 2017

Department of Public Health and Policy

University of Liverpool

Declaration

This thesis is my own work. The material contained in this thesis has not been presented, nor is currently being presented, either wholly or in part for any other degree or qualification.

I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Chapter 4 contains results previously published PLOS ONE as *Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review and meta-analysis* by Natalie L. Adams and Tanith C. Rose (students and joint first authors), Victoria J.K. Howard (student) and Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Margaret Whitehead, Ross Harris and David C. Taylor-Robinson (supervisors). All authors contributed to the conception and design of the study. NA and TR performed the literature searches and NA, TR and VH performed the data extraction. NA, TR, BB, DTR and RH performed the analyses and all authors interpreted the data. NA and TR drafted the manuscript which was revised critically by all authors. All authors approved the final version of the manuscript.

Chapter 5 contains results previously published in The European Journal of Public Health as *Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 Study)* by Natalie L. Adams (student and first author) and Tanith C. Rose (student) and Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Margaret Whitehead, Benjamin Barr and David C Taylor-Robinson (supervisors). All authors contributed to the conception and design of the study. NA performed the analyses with guidance from BB and DTR. NA drafted the manuscript which was critically revised by all authors. All authors approved the final version of the manuscript.

Acknowledgements

The research was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections at University of Liverpool in partnership with Public Health England (PHE), in collaboration with University of East Anglia, University of Oxford and the Quadram Institute. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England. I would like to express my gratitude to the NIHR HPRU in Gastrointestinal Infections for providing the funding which made it possible for me to undertake this work and for the many development opportunities.

This work would not have been possible without the support of a great many people who all contributed to making this a wonderful experience. To each I owe a particular debt of gratitude for their support and encouragement to undertake this research. I would like to thank my supervisors Dame Professor Margaret Whitehead, Professor Sarah O'Brien, Professor David Taylor-Robinson, Professor Jeremy Hawker, Dr Benjamin Barr and Dr Mara Violato for the unfailing support, wisdom and patience they have shown me. I have been extremely fortunate to have had the opportunity to benefit from their expertise and experience.

Undertaking this research would have been all the more challenging without the support and friendship of colleagues past and present in the wider community of the National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections, the Department of Public Health and Policy at the University of Liverpool and Public Health England. There are too many to mention individually but I would like to thank in particular Professor Bob Adak, Dr John Harris, Dr André Charlett, Professor Finn Diderichsen, Ross Harris, James Lewis, Gillian Smith, Alex Elliot, Paul Loveridge, Vivianne Buller, Lisa Byrne, Kirsten Glen, Ros Lewis and Cletha Fialho.

To my family and friends, particularly Debbie, Martin, Calvin, Jo, Emma, Jennie, Val, Tim, Charlotte and Sian - who have shown me such support and who each found their own way to help, I am immensely grateful. Phil, thank you for all your support - for the early starts to drop me at the station, rescuing me from various locations when I was trying to get home, weekends spent proof-reading, the opportunity to work on my thesis in Bangkok and for ensuring I was well fed! Tanith, with whom I have shared every moment of this, thank you for your support; I could not imagine having done this without you.

Abstract

Introduction: Gastrointestinal (GI) infections are a significant burden both to the NHS and to society; affecting around a quarter of the UK population each year at an estimated cost of £1.5 billion. Socioeconomic inequalities in health are a serious problem and reducing such inequalities is high on the public health agenda in this country. Many infections are socially patterned but the role of socioeconomic inequalities in the risk of and exposure to GI infections is unclear, with published studies providing conflicting results. This thesis aims to investigate whether risk of or exposure to GI infections in the UK is socially patterned and if so, which sectors of society experience a greater burden of infection and through what mechanisms.

Methods: I undertook novel analyses of existing UK-based high-quality and comprehensive secondary data on GI infections to explore the relationship between GI infections and socioeconomic status (SES) using a variety of analytical techniques. Study 1 explores the role of SES in risk of GI infections in high income countries through a systematic review and meta-analysis of 102 published studies using random- and fixed-effects meta-analysis and random-effects meta-regression. Study 2 assesses the association between SES and GI infections in a community cohort of 6,836 participants, using a Cox proportional hazards survival analysis approach. Study 3 presents results from an observational study utilising two NHS telephone-based services to explore the role of SES amongst individuals accessing remote health advice. Finally, Study 4 presents results of a case-study of a severe GI infection, Shiga toxin-producing *Escherichia coli* (STEC), to investigate socioeconomic patterning of risk factors for infection and to explore the role of demographic and socioeconomic factors in progression from STEC to a severe outcome, Haemolytic Uraemic Syndrome (HUS) in a separate cohort of paediatric HUS cases.

Results: In high income countries, disadvantaged children but not adults had a significantly higher risk of GI infection compared to less disadvantaged children. In England, odds of calls to NHS helplines about GI infection symptoms in disadvantaged children and adults were significantly higher compared to their less disadvantaged counterparts. Disadvantaged adults were found to have lower risk of GI infections in the community cohort and both children and adults were less likely to be reported as having STEC infection or developing HUS. Overall, the results provide strong evidence to suggest that risk of GI infection differs by SES across the life course, with disadvantaged children at highest risk of GI infections.

Conclusions: Disadvantaged children are at greater risk of GI infections compared to their more advantaged counterparts in the UK. The relationship between deprivation and risk of GI infection in adults is less clear. This thesis found that increased risk may relate to differential exposure, vulnerability or healthcare-seeking behaviours, including symptom recognition, across socioeconomic groups. This work has provided further insight into relationship between SES and GI infections and sets the direction for policies to reduce inequalities in GI illness in children and for more focussed research to deepen the understanding of the relationship particularly between SES and GI infection in adults.

Contents

Tables

Figures

Abbreviations

A&E	Accident and Emergency
BPSU	British Paediatric Surveillance Unit
CI	Confidence interval
E. coli	*Escherichia coli*
ESQ	Enhanced surveillance questionnaire
GBRU	Gastrointestinal Bacteria Reference Unit
GI	Gastrointestinal
GLM	Generalised linear models
GP	General practice/practitioner (primary care)
HES	Hospital Episode Statistics
HPA	Health Protection Agency
HPRU	Health Protection Research Unit
HR	Hazard ratio
HUS	Haemolytic Uraemic Syndrome
ICD	International Statistical Classification of Diseases and Related Health Problems
IID	Infectious intestinal disease
IID2 Study	The Second Study of Infectious Intestinal Disease in the UK
IRR	Incidence rate ratio
IMD	Index of multiple deprivation
LSOA	Lower layer Super Output Area
NHS	National Health Service
NIHR	National Institute for Health Research
NS-SEC	National statistics socioeconomic classification
OECD	Organisation for Economic Co-operation and Development
ONS	Office for National Statistics
OR	Odds ratio
PCR	Polymerase chain reaction
PHE	Public Health England
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RR	Relative risk, rate ratio, risk ratio, ratio of risk ratio
SEC	Socioeconomic conditions
SES	Socioeconomic status
STEC	Shiga toxin-producing *Escherichia coli*(also known as VTEC)
Stx	Shiga toxin
VTEC	Vero cytotoxin-producing *Escherichia coli*(also known as STEC)
WHO	World Health Organisation

Chapter 1
Introduction

1.1 Relevance of the issue

Gastrointestinal (GI) infections, which cause diarrhoea, vomiting and occasionally more serious complications, are a public health concern. Globally, diarrhoeal disease is the second leading cause of death in children under five and is responsible for around 525,000 deaths each year in this age group (World Health Organisation, 2017a) although this burden is felt particularly in low income countries. In high income countries, GI infections are common and mostly self-limiting; with morbidity, rather than mortality, being the most pressing public health concern.

In the UK, which is the country on which this thesis focuses, GI infections are a significant burden both to the National Health Service (NHS) and society. Previous estimates suggest around 25% of people in the UK suffer an episode of infectious intestinal disease (IID) per year (Tam et al., 2011a), which equates to approximately 17 million cases, and around one million General Practice (GP) consultations, annually (Tam et al., 2011a). It is estimated that 19 million days are lost each year: 11 million working days and 8 million absences from school (Tam et al., 2011a). Foodborne illness, a subset of IID, in England and Wales costs individuals, the economy and the NHS around £1.5 billion annually (Tam et al., 2011a). A large proportion of the burden of GI infection remains hidden; it is estimated that there are 147 cases in the community for every one case that is reported to national surveillance (Tam et al., 2011a); many individuals do not present to healthcare.

Socioeconomic inequalities in health, which are "*differences in health status between different socioeconomic groups that are avoidable and unjust and as such are amenable to concerted action*" (Whitehead, 1990), are increasingly recognised as significantly contributing to both the causes and consequences of morbidity and mortality (Dahlgren and Whitehead, 2007a, Wilkinson and Marmot, 2003). Indeed, using the London Underground as an analogy, area-level socioeconomic differences mean travelling just 20 minutes on the Central line from Lancaster Gate to Mile End represents a decrease in life expectancy of 12 years (Cheshire, 2012). The specific role of these socioeconomically-driven inequalities is well-understood for many diseases of both an infectious and non-infectious nature – with increasing disadvantage placing individuals at increased risk of, exposure to and consequences of diseases including cancer, coronary heart disease (Graham, 2009), human immunodeficiency virus and tuberculosis (Biering-Sørensen et al., 2012, Hughes and

Gorton, 2015, Semenza, 2010). Despite this, the role of socioeconomically-driven inequalities is not well understood for GI infections.

Despite interventions aimed at reducing the burden of GI infections, levels of GI infections have remained relatively stable over time, suggesting that new approaches to understand and tackle these infections must be sought. This thesis aims to investigate whether risk of or exposure to GI infections in the UK is socially patterned and if so, which sectors of society experience a greater burden of infection and by which mechanisms. The studies in this thesis are needed to identify effective policy entry points and understand the role played by structurally determined lifestyles in generating any observed inequalities in GI infections.

1.2 Previous research and gaps in the literature

Several studies have sought to understand the role of socioeconomic inequalities in risk of and exposure to GI infections in high income countries and a focus on socioeconomic inequalities as a driver of differential risk of GI infection has increased in recent years. Studies that have examined the relationship between socioeconomic status (SES) and risk of GI infection in high income countries have found conflicting results (Newman et al., 2015), with multiple studies finding no differential risk, and others reporting higher risk in more disadvantaged groups and lower risk in more disadvantaged groups. These studies are highly heterogeneous in terms of study design as well as in data sources and analytical methods, making it challenging to identify trends within this literature. It is also unclear whether the differing socioeconomic patterns reported in the literature reflect socioeconomic differences in risk of infection or rather differences in reporting of disease, exposure to risk factors and interaction with or access to healthcare services. Studies finding lower risk in more disadvantaged groups often cite differential exposure or ascertainment bias as the likely reason for this finding (de Wit et al., 2001a, Gillespie et al., 2008, Jalava et al., 2011, Simonsen et al., 2008, Spencer et al., 2012). Studies finding higher risk in more disadvantaged groups often also cite differential exposures as the main factor explaining results (Gillespie et al., 2010a, Simonsen et al., 2008), although others suggest findings could relate to differential healthcare interaction or differential disease severity (Beale et al., 2010, Herikstad et al., 2002).

Whilst there is, as yet, no clear consensus on whether the risk of, or exposure to, GI infections varies by socioeconomic position, there is worrying evidence in the UK suggesting that the consequences of GI infections are greater amongst more disadvantaged groups in society, with children of parents from unskilled/manual social classes found to have twice the odds of norovirus-associated gastroenteritis compared to children of parents from non-manual social classes (Phillips et al., 2011). Furthermore, hospital admissions for children with gastroenteritis were found to be almost twice as high in the most disadvantaged areas compared to the least disadvantaged (Olowokure et al., 1999).

Like many diseases, GI infections arise from a complex interplay between host factors, individual behaviours and a varied set of exposures – all of which may interact to contribute to differential risk of disease. It is epidemiologically challenging to explore these individual and potentially multiplicative risks and exposures, particularly at the population level. Improvements in the amount and quality of both health and socioeconomic data in recent years, as well as the ability to perform linkage between numerous and vast datasets has increased the power and ability to study this topic. As our ability to link and research socioeconomic factors in more depth increases, so do the opportunities to better understand this topic.

Given the high social and economic burden of GI infections, there is a need to understand whether socioeconomic inequalities in GI infections exist and what the extent and nature of any inequality identified is. The role of socioeconomic inequalities in risk of and exposure to GI infections in the UK is the primary focus of this thesis, with the aim to fill some of the gaps in the knowledge on this relationship. A variety of novel analyses of existing UK-based high-quality and comprehensive secondary data on GI infections (including community cohort data, NHS telephone-based healthcare advice data and data from national enhanced surveillance systems) were conducted to assess the association between GI infections and SES. This thesis provides estimates of the role of socioeconomic inequalities in the risk of GI infection, suggests explanations for inequalities identified and presents the implications for policy of the findings. The results of this work contribute to enhance the evidence base on the role of SES in GI infections with the aim to ultimately inform policies and interventions to reduce the risk, vulnerability and social, economic and healthcare consequences of GI infections in the UK.

1.3 Aim and objectives of this thesis

The overall aim of this thesis is to investigate the existence, extent and nature of socioeconomic inequalities in risk of GI infections in the UK. The objectives of this thesis are therefore:

1. To conduct a systematic review of existing evidence of socioeconomic inequalities in risk of GI infections in high income countries.
2. To investigate the extent and nature of socioeconomic inequalities in risk of GI infections in the community in the UK, with estimates derived from the most up-to-date population-based household survey.
3. To analyse the extent of, and mechanisms underlying, socioeconomic inequalities in risk of GI infections in the community, with estimates derived from routine data on members of the public seeking telephone-based healthcare advice in England.
4. To explore the social patterning of clinical outcomes, healthcare contact and risk factors for a laboratory-confirmed, potentially severe, GI infection (Shiga toxin-producing *Escherichia* coli; STEC) and socio-demographic inequalities in risk of development of a serious sequela (Haemolytic Uraemic Syndrome; HUS) in England.
5. To draw out policy implications and recommendations for further research into the role of socioeconomic inequalities in GI infections.

1.4 Overview of the thesis

Figure 1 provides an overview of the four empirical studies which comprise this thesis. In what follows, I summarise the content of each chapter of the thesis.

- Chapter 2 (Literature review) provides a summary of the causes of health inequalities and an overview of GI epidemiology. A review of the existing literature on the role of socioeconomic inequalities in GI infections is presented and key gaps in the literature are highlighted.

- Chapter 3 (Methods) describes each of the primary data sources and provides an overview of the analytical methods used in the subsequent empirical research chapters.

- Chapter 4 (Study 1) presents the results of a systematic review and meta-analysis of the evidence on the relationship between socioeconomic status and GI infections in high income countries. It discusses the findings, the limitations of the available evidence and recommendations for future studies investigating the role of socioeconomic inequalities in risk of GI infections. This addresses objective 1: to establish the current levels of knowledge.

- Chapter 5 (Study 2) presents the results of the latest longitudinal population-based survey of IID in the community in the UK. A survival analysis approach was undertaken to assess the role of SES in the risk of developing IID, accounting for differential follow-up time. This addresses objective 2: to provide the most up-to-date robust estimates of risk by SES within the community, derived from household survey data.

- Chapter 6 (Study 3) presents the results of an observational study investigating socioeconomic inequalities in symptoms of diarrhoea and vomiting in the community accessing telephone-based advice using the NHS Direct and NHS 111 telephone services. This addresses objectives 3: to explore the role of socioeconomic inequalities in IID risk amongst members of the public seeking telephone-based healthcare advice.

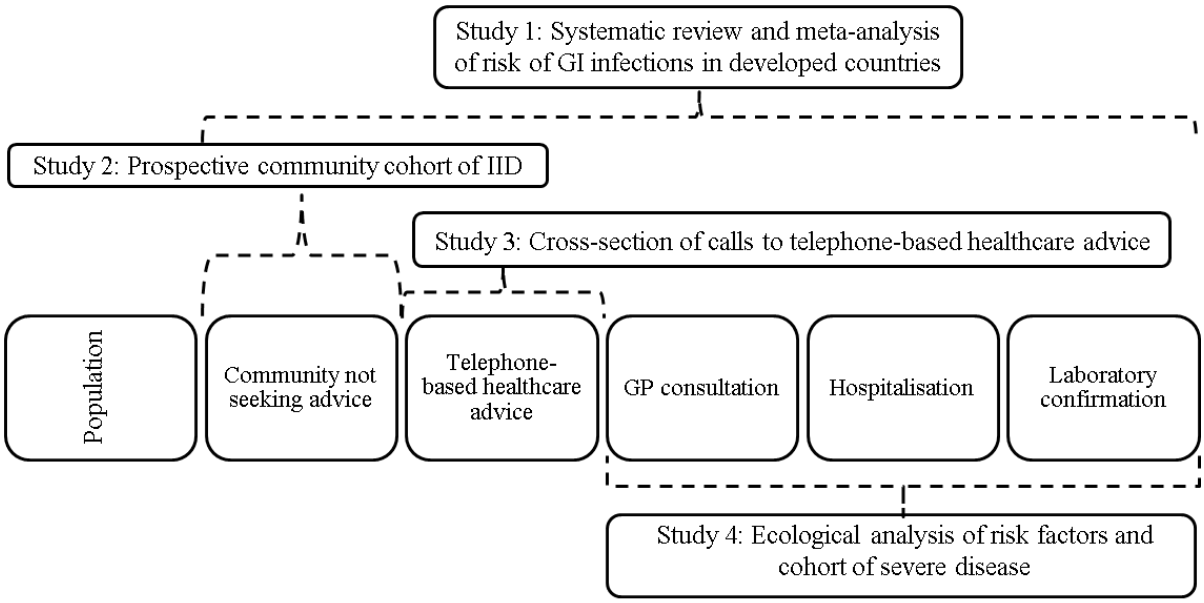
- Chapter 7 (Study 4) presents the results of an observational study of STEC cases in England, exploring the social patterning of exposures and risk factors for infection and the social patterning of the development of HUS. This addresses objectives 4: to explore the social patterning of exposures to GI infection and the sociodemographic inequalities in risk of progression to a severe sequela.

- Chapter 8 (Discussion) provides a discussion of the results and key findings from the preceding four study chapters (Chapters 4-7). Strengths and limitations of the overall thesis are discussed. It also provides reflections on policy implications and recommendations for future research.

- Additional appendices corresponding to each chapter are provided. Details of ethical approvals, search terms, questionnaires, additional analysis tables and publications from this thesis are provided.

Through this thesis, I consolidate existing knowledge on risk of GI infections in high income countries, including syndromic definitions of GI infections which allows for any bias in terms of inequalities in healthcare interaction to be minimised. I explore inequalities in GI infections in England in terms of individuals who would not necessarily be captured by any national surveillance system using novel analyses of two comprehensive community-level datasets. Finally, I analyse data from a large national enhanced surveillance system for a specific laboratory-diagnosed GI pathogen to explore the social patterning of exposures in England. The role of socioeconomic status in the consequences of GI infections is outside the scope of this thesis. I do not address inequalities in the risk of chronic or non-acute GI infections nor do I address inequalities in GI infections in low income countries.

Figure 1.1: Overview of the studies in this thesis



Chapter 2
Literature Review

2.1 Introduction

Despite evidence to support the association between SES and health inequalities (Dahlgren and Whitehead, 2007a), the relationship between socio-economic status and individual or population-level risk for GI infections in high income countries is unclear. In particular, whilst there is some evidence to suggest that the consequences of GI infections are greater for more disadvantaged groups (Conway et al., 1990, Olowokure et al., 1999, Phillips et al., 2011) there is little evidence to link socio-economic status to risk of or exposure to GI infections; where evidence does exist, the results are conflicting. Yet understanding the relationship between SES as a potentially modifiable factor and GI infections, as well as the mediating pathways from SES to risk or outcome is important to identify effective policy entry points.

This chapter first describes the historical context to socioeconomic inequalities in infectious diseases in the UK. I then present some general theories as to the causes of health inequalities and the different mechanisms that may be operating to generate socioeconomic inequalities in health as proposed in the Diderichsen model (Diderichsen et al., 2001). I then present an overview of GI epidemiology relevant to this thesis before focussing on the key literature exploring socioeconomic inequalities in the risk of GI infections in high income countries.

2.2 Historical socioeconomic inequalities in infectious diseases

There is a long tradition of acknowledgment of, and research into, health inequalities in the UK (Goldblatt and Whitehead, 2000), and socioeconomic inequalities in infectious diseases in particular. This section aims to provide an overview of some of the key events which have occurred in more recent history from the 18th century that have contributed to our understanding of inequalities in infectious diseases in the UK.

The first reliable marker of population fluctuations dates back to the Doomsday book of 1086 AD (Whitehead, 1998) and since then strong evidence to suggest that poverty and ill-health are inextricably linked has emerged. During the Industrial Revolution in Britain, which began in the late 1700s, overcrowding in large towns and in workhouses leading to poor hygiene and unsanitary conditions became commonplace particularly for the most disadvantaged and led to the spread of

infectious diseases, including cholera, typhoid and general ill-health (Royal Commission on the Health of Towns, 1844).

While the role of poverty in increasing the risk of specific infectious diseases was not well understood at the time, it was recognised that those in poverty were experiencing an unequal burden of infectious diseases. As such, from 1834 the Poor Law Amendment Act, designed to improve the health of the public and particularly of the poor, was passed by Parliament. In 1836, The Registrar General's Office was established by Parliament as part of this process in order to track births and deaths in England and Wales for the purposes of property rights (Lilienfeld, 2007). William Farr, a British epidemiologist, was appointed as Compiler of Abstracts within the Registrar General's Office. He laid the foundations for the use of these national statistics for surveillance and epidemiological analyses, including investigations of cholera and smallpox outbreaks (Dunn, 2002, Lilienfeld, 2007), and emphasised the importance of not simply recording deaths but of reviewing data recorded by the national office in order to try and understand underlying associations between inequalities and ill-health to improve outcomes for future generations (Acheson et al., 1998). Similarly, Dr William Henry Duncan, Liverpool's first Medical Officer of Health between 1847 and 1863 and the first in the country, was a strong proponent of improving the health of the public and with particular regard for the living conditions of the most disadvantaged (Ashton, 1988), linking poor sanitation as the underlying cause of infectious diseases such as cholera (Halliday, 2003).

With improvements in sanitation and hygiene, infectious diseases declined and more recently the focus for research and policy recommendations has been on the relationship between socioeconomic status and health more generally. The establishment of the National Insurance Act 1911, the founding of the NHS in 1948 and a variety of inquiries set up to investigate and assess health inequalities in the UK (Acheson et al., 1998, Black et al., 1980, Marmot et al., 2010, Whitehead, 1992, Whitehead et al., 2014), have also highlighted that socioeconomic inequalities in health are still prevalent in the UK and there is also evidence to suggest that inequalities in infectious diseases such as tuberculosis and sexually transmitted infectious remain a concern (Parliamentary Office of Science and Technology, 2017).

Historically, poorer groups in high, middle and low income countries have had much higher rates of infectious diseases and greater mortality from them compared to their more affluent counterparts and in most countries, these inequalities remain glaringly evident (World Health Organisation, 2012). Furthermore, these historical studies of socioeconomic inequalities in infectious diseases have led to breakthroughs in understanding causal pathways and the role played by the conditions in which disadvantaged families live, including sub-standard housing, overcrowding, malnutrition, sanitation and contaminated water supplies. In light of this, it is therefore logical to be concerned about socioeconomic inequalities in GI infections and the focus on inequalities in GI infections in this thesis is just as relevant today as such studies may help reveal factors that are helping or hindering prevention and control of these infections in different groups in society. As such, the studies in this thesis have been devised to determine what the current social patterning of GI infections is in high income countries such as the UK and what the underlying mechanisms might be.

Causes of health inequalities

Socioeconomic inequalities in health are "*differences in health status between socioeconomic groups that are avoidable and unjust and as such are amenable to concerted action*" (Whitehead, 1990). There are several theories which seek to explain the causes of health inequalities such as the artefact explanation and theories of social selection and social causation.

i) The artefact explanation suggests that biases in response or measurement of SES may give a misleading relationship between health and SES (Black et al., 1980).
ii) Natural and social selection explanations suggest that health can determine subsequent social position: people in poor health tend to move down the social scale, while people in good health tend to move up into higher classes, thereby perpetuating the observed gap in health between the top and bottom of the social scale (Black et al., 1980).
iii) Two main social causation explanations suggest that social circumstances can affect subsequent health, but differ in where they put the emphasis. The structural/material explanation emphasises the role of the external conditions under which people live and work: people in lower social positions suffer poorer

living and working conditions which lead to poorer health (Black et al., 1980). The cultural/behavioural explanation emphasises differences between socioeconomic groups in lifestyle preferences or behavioural decisions that lead to differences in their subsequent health. Arguably, the two social causation explanations are linked as the lifestyles of different socioeconomic groups are shaped by the social and economic environments in which they live (Dahlgren and Whitehead, 2007a).

iv) The life course perspective suggests that early life experiences or exposure to risk factors shape the health of an individual as an adult. Such life experiences or risk factors may also cluster together; exposure to one factor may increase the likelihood of experiencing other factors.

Social determinants of health

The social determinants of health approach proposes that a combination of the economic and social factors described above lead to socioeconomic inequalities in healthcare access and health outcomes. Social determinants of health can be seen as the

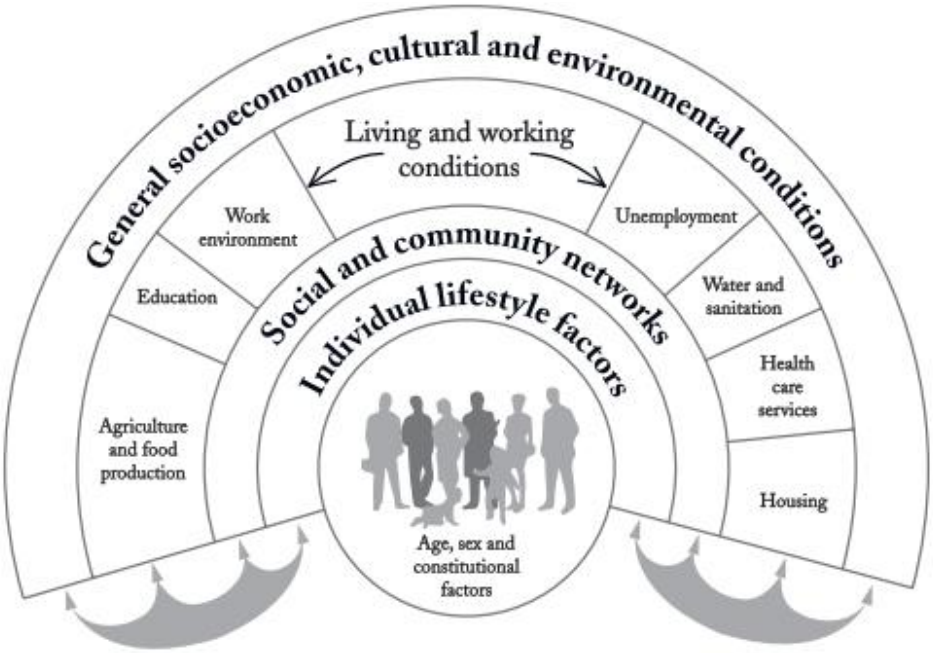
> "*conditions in which people are born, grow, work, live and age, and the wider set of forces and systems shaping the conditions of daily life. These forces and systems include economic policies and systems, development agendas, social norms, social policies and political systems*".
>
> (World Health Organisation, 2017b)

The Dahlgren and Whitehead model illustrating the social determinants of health (Figure 2.1) (Dahlgren and Whitehead, 1993), represents the layers of interacting factors which contribute to and influence the health of populations. At the centre, individuals possess age, sex and constitutional characteristics that influence their health but are largely fixed and therefore not amenable to change. Surrounding the core are layers of influence that are theoretically modifiable to policy. First are individual lifestyle factors such as food consumption habits and smoking which influence health, but do not take place in a vacuum – they are affected by the layers of influence beyond. Second, a person's health is influenced by social and community networks such as friends and family, among whom they live and work. These in turn are influenced by factors in the third layer: a person's living and

working conditions and access to essential goods and services such as education, employment, effective healthcare, and food supply. Finally, the fourth layer represents the overarching mediators of population health: the general socioeconomic, cultural and environmental conditions prevailing in society. Each layer has an impact on the previous layer and can influence individuals and groups in society.

Figure 2.1 Social determinants of health approach

Source: Dahlgren and Whitehead (2007b)



The Diderichsen Model

The Diderichsen model is an analytic framework which can be used to distinguish between different mechanisms that may be operating to generate socioeconomic inequalities in health (Diderichsen et al., 2001). Within this model there are a number of mechanisms through which inequalities can be generated (Figure 2.2). Firstly, social stratification (I) is the mechanism by which individuals in society are sorted into different social positions based on their social contexts. These contexts can then lead to individuals in these different social positions experiencing differential exposure to conditions which may lead to ill health (II). Differential vulnerability (III) is a mechanism through which the same level of exposure could have different

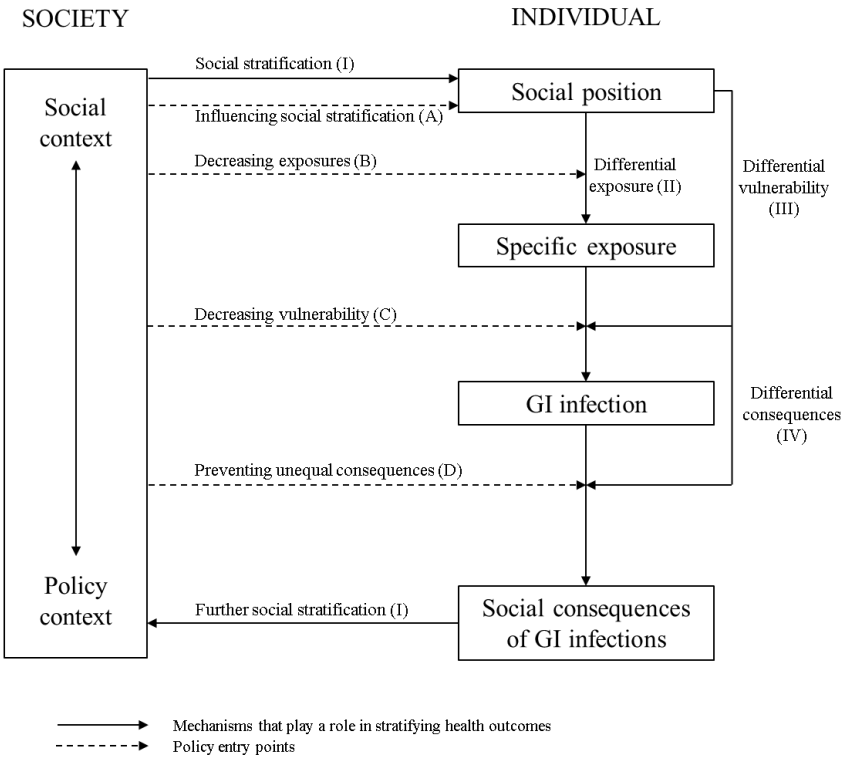
effects across social strata. Finally, individuals across different social strata may then experience differential social and economic consequences of ill health (IV).

This model also highlights several policy entry points (A-D) (Diderichsen et al., 2001). The first policy entry point (A) represents an opportunity to influence social stratification through policies designed to reduce social inequalities through education, labour market or family welfare policies; or through impact assessments of social and economic policies to reduce their impact on social stratification (Diderichsen et al., 2001). The second and third policy entry points (B and C) represent an opportunity to introduce policies to reduce exposure and vulnerability; for example through reducing excess exposure to potential risk factors according to social position (B) or by attempting to modify the effect of interacting exposures (C) (Diderichsen et al., 2001). The final policy entry point (D) seeks to prevent unequal consequences of ill-health and further social inequalities. Although not the focus of this thesis, policies in this category concentrate on reducing differential outcomes of disease such as through a focus on equity of healthcare resource allocation and consideration of additional resources according to social position (Diderichsen et al., 2001).

Through this model we can begin to understand the ways in which health inequalities are generated and, by applying this to GI infections, help identify policy areas which may help to reduce inequalities in GI infections. This model will be used as the basis on which to explore the pathways leading to socioeconomically driven inequalities in risk of and exposure to GI infections.

Figure 2.2 Diderichsen Model

Source: Diderichsen et al. (2001)



2.3 GI infections

Gastrointestinal infections cause a variety of symptoms, most commonly including diarrhoea and/or vomiting, and are caused by a range of pathogens including viruses, bacteria and protozoa (Adak et al., 2002, Musher and Musher, 2004) each with differing epidemiology and management, making control and prevention challenging. In healthy individuals, illness is usually short-lived and most recover within a few days but have the potential in some cases to cause serious sequelae such as HUS (Byrne et al., 2015), Guillain-Barré syndrome (McCarthy and Giesecke, 2001) and reactive arthritis (Dworkin et al., 2001), and are an important public health concern in terms of their potential clinical, economic and social burden.

GI infections are common and are a significant social and economic burden. Despite interventions designed to control and reduce transmission, the incidence of many GI infections in the UK has remained relatively constant. It is clear that new efforts to understand acquisition and transmission of diarrhoeal illness in the UK are required

to tackle this issue. Whilst a better understanding of the role of pathogens, hosts and their interactions, as well as biological and microbiological processes is crucial, it is also essential to understand socioeconomic mechanisms that may be contributing to the failure to reduce these illnesses. With a better understanding of the mechanisms leading to socioeconomic inequalities in GI infections, it may be possible to identify important links in the causal chain that can be addressed more effectively than at present.

Epidemiology of GI infections in the UK

In the Second Study of Infectious Intestinal Disease in the UK (IID2 study), the overall community-level incidence of IID was estimated to be 274 cases per 1,000 person-years (95% confidence interval (CI): 254-296), which equates to 25% of the population, or 17 million people, in the UK experiencing an episode of IID each year (Food Standards Agency, 2000, Tam et al., 2011a). Children, particularly those under 5, are regarded as experiencing the highest burden of community-level IID, with children under one year of age experiencing an average of at least one episode of IID per year (1,079 per 1,000 person-years, 95% CI 750.1-1553.3) (Tam et al., 2011a).

In the UK, it has been estimated that the annual incidence of acute gastroenteritis presenting to healthcare services in children under five years of age is 5.78 per 100 children (95% CI 5.30-6.32) (Van Damme et al., 2007), accounting for 16% of paediatric accident and emergency (A&E) presentations in one study (Armon et al., 2001). Women are also regarded as experiencing a higher burden of community-level GI infection compared to men across most age groups, but particularly amongst women aged 25-34 years (387/1,000 person-years, 95% CI 304.9 - 492.2) compared to men of the same age (184/1,000 person-years, 95% CI 102 - 332.5) (Tam et al., 2011a). Whilst infants undoubtedly experience the highest burden, individuals aged 65 years and above have also been identified as experiencing a greater burden of IID, particularly in terms of recurrent episodes of IID (Tam et al., 2013). Amongst this age group, each additional IID episode was found to increase the risk of experiencing a subsequent IID episode three-fold, compared to 1.7- fold in infants (Tam et al., 2013). This is thought likely to be due to a greater proportion of individuals with impaired immune systems or chronic gastrointestinal conditions which could increase risk of illness (Tam et al., 2013). Being immunocompromised (Roy et al.,

2004), the use of proton pump inhibitors to suppress gastric acid (Doorduyn et al., 2010, Freeman et al., 2016, Jensen et al., 2017) and having comorbidities such as chronic intestinal illness (Doorduyn et al., 2010) have been shown to increase risk of GI infection.

Transmission of GI infection

Transmission of GI infection is generally via the faeco-oral route. This can be direct; as the result of poor hand hygiene and person-to-person spread via contact with other symptomatic individuals; or indirect via contaminated food vehicles, water or the environment. In the UK, and other high income countries, a variety of exposures and risk factors for the acquisition of a GI pathogen and subsequent development of a symptomatic GI infection have been established, largely as the result of outbreak investigations.

Environmental exposures, such as farming activities (Doorduyn et al., 2010, O'Brien et al., 2001), exposure to animals/ruminants and their faeces (Adams et al., 2016b, Byrne et al., 2015, Doorduyn et al., 2010, Gillespie et al., 2003, Kapperud, 2003, Locking et al., 2001, Neimann et al., 2003), domestic animals and pets (Doorduyn et al., 2010, Gillespie et al., 2003, Little et al., 2008), reptile-associated exposures (Harker et al., 2014) and foreign travel (Friedman et al., 2004, Gillespie et al., 2003, Neimann et al., 2003) have all been associated with infection. In the IID2 Study, between 8% and 12% of participants with symptoms had travelled outside the UK in the ten days prior to onset (Tam et al., 2011a). Waterborne exposures such as private or contaminated water supplies have been reported (Adams et al., 2016b, Byrne et al., 2015, Kapperud, 2003), including swimming (Doorduyn et al., 2010).

A wide variety of food vehicles has been implicated in outbreaks, ranging from vegetables (Gillespie et al., 2010b, Launders et al., 2015), salad items (Byrne et al., 2016, Jenkins et al., 2015, Long et al., 2002) and fruit (Byrne et al., 2014a, Neimann et al., 2003) to meat such as beef, ham, pork and lamb (Adams et al., 2016b, Gillespie et al., 2010b, Gillespie et al., 2003, Kapperud, 2003, Little et al., 2008, Neimann et al., 2003). The consumption of eggs (Harker et al., 2014, Lane et al., 2014, Molbak and Neimann, 2002), chicken or poultry (Doorduyn et al., 2010, Friedman et al., 2004, Harker et al., 2014, Little et al., 2010, Neimann et al., 2003),

particularly when either is consumed or handled raw or undercooked, and fish and shellfish are also potential risk factors (Gillespie et al., 2010b, Inns et al., 2013).

Unpasteurised milk and pasteurisation failures have also been associated with risk of GI infection, particularly *Campylobacter* (Evans et al., 1996, Fahey et al., 1995, Gillespie et al., 2010b, Gillespie et al., 2003, Neimann et al., 2003, Southern et al., 1990). Consumption of ready-to-eat food, such as sandwiches (Gillespie et al., 2010b), eating outside of the home (Byrne et al., 2014a, Doorduyn et al., 2010, Friedman et al., 2004, Gillespie et al., 2003) and eating at barbeques (Doorduyn et al., 2010, Gillespie et al., 2003, Kapperud, 2003, Neimann et al., 2003) have also been reported.

Identifying GI infections in the UK

There are multiple ways of measuring GI infections in the UK although the only way to be certain that an individual is infected with a GI pathogen is to identify such a pathogen via diagnostic testing in a laboratory. Gastrointestinal pathogens causing food poisoning, heamolytic uraemic syndrome (HUS) or infectious bloody diarrhoea are statutorily notifiable therefore the details of cases and the positive results of any diagnostic tests will be reported via laboratory reporting systems to national surveillance. For GI infections not caused by food poisoning, positive results are voluntarily reported from diagnostic laboratories via the same reporting systems.

For some specific GI pathogens, namely *Listeria monocytogenes* and STEC, enhanced surveillance systems are in place due to the potential severity of illness resulting from these pathogens. In both of these enhanced surveillance systems, standard forms are collected which record in-depth epidemiological and clinical data such as patient demographics, symptoms and food, water and environmental exposures. These are used to supplement the microbiological results to provide more detailed information on risk factors and exposures than could be gleaned from diagnostic results alone. In outbreak situations, standard forms are used to record details of cases and of the outbreak itself. Measurement of individuals with STEC, and HUS, will be described in more detail in Chapter 7, which presents the results of an analytical study using these data.

GI infections are usually self-limiting and therefore many instances of illness go unreported to national surveillance as individuals may not seek care and as such these cases are not recorded in surveillance systems. There are systems which can be used to identify syndromic definitions of GI infections, such as the NHS 111 telephone advice service. This system records data on individuals utilising the telephone service including information on the syndromes of diarrhoea and vomiting. Further detail on the use of NHS 111 and its predecessor, NHS Direct, will be provided in Chapter 6 which details the use of these data in an analytical study.

As GI infections are often relatively mild and self-limiting, contact with healthcare services is not always necessary. Attempts to measure the true incidence and burden of GI infections in the UK began in the late 1980s. Following national epidemics of foodborne illness associated with *Salmonella* enteritidis phage type 4 and *Listeria* monocytogenes (Food Standards Agency, 2000), it was decided by the Secretary of State for Health and the Minister of Agriculture, Fisheries and Food that there was a need to establish a Committee on the Microbiological Safety of Food, chaired by Professor Mark Richmond in 1989. The recommendations of the committee included setting up a study to determine the true incidence of IID in the community (Food Standards Agency, 2000).

The first such study (the IID1 Study) was conducted in England between 1993 and 1996 and estimated that a fifth of all individuals suffered from an episode of IID each year (Food Standards Agency, 2000), thus affecting around 9.5 million individuals at that time. It estimated that for every one case notified to national surveillance, there were 136 unreported cases in the community (Food Standards Agency, 2000). It also quantified the cost of IID in England as being approximately £750 million (Food Standards Agency, 2000). This study emphasised the magnitude of the burden of GI infections, in particular the 'hidden' burden in the community. It raised awareness of the importance of trying to tackle GI infections in order to reduce the costs both to society and the NHS. A decade later a second study (the IID2 Study), which is used in Study 2 of this thesis, was commissioned to determine whether the incidence of IID had changed since the mid-1990s and to recalibrate national surveillance data (Tam et al., 2011a). The IID2 study, unlike the IID1 study, covered the UK and revised the estimate from the first study to suggest that a quarter of all individuals in the UK suffer from an episode of IID each year and that for every one case that is

reported to national surveillance, there are 147 cases in the community (Tam et al., 2011a). The estimated cost of foodborne IID, a proportion of all IID, in the UK was £1.5 billion (Tam et al., 2011a). The results of the IID2 study indicated that the burden of GI infections remained high, and the estimate was higher still than that of the IID1 Study. This highlights the need to further understand risk factors for GI infections and the need for the work planned within this thesis which seeks to provide a deeper understanding of the role of SES in the risk of GI infections in order to generate policies which could reduce this burden.

2.4 Socioeconomic inequalities in GI infections

This section provides a critical review of illustrative examples of the literature identifying and analysing the current knowledge on socioeconomic inequalities in the risk of and exposure to GI infections. This review is not exhaustive and will be built upon further in the systematic review presented in Chapter 4 (Study 1). Particular attention is paid to methodological and analytical approaches as well as potential confounding, mediating or moderating variables in order to inform the studies in this thesis. Where relevant, comparisons and contrasts are drawn between studies conducted in the UK and in other high income countries including Denmark, Germany, France, Australia, Canada and the USA although the primary focus is on the UK.

Studies conducted in the UK

In the UK, several studies have explored the relationship between GI infections and SES for both adults and children. For adults, studies using laboratory report data present an unclear picture with several studies finding no association between incidence and SES for Giardia, all *Salmonella* (Hughes and Gorton, 2015) and *Salmonella* Typhimurium (Banatvala et al., 1999) although lower incidence of *Salmonella* Enteritidis was associated with more disadvantaged areas (Banatvala et al., 1999). Lower risk of *Campylobacter* (Bessell et al., 2010, Gillespie et al., 2008, Hughes and Gorton, 2015, Nichols et al., 2012), *Cryptosporidium* (Hughes and Gorton, 2015, Lake et al., 2007) and *Shigella* (Hughes and Gorton, 2015) was also identified in more disadvantaged areas. In contrast, for *Listeria*, incidence was found to be higher in more disadvantaged areas (Gillespie et al., 2010a).

Similarly, studies using syndromic definitions of GI infections such as diarrhoea or vomiting also present conflicting findings for adults. No association was identified between symptoms of GI infection in four studies (Beale et al., 2010, Evans et al., 2006, McAteer et al., 2011, Stone et al., 1994) and one study (Scallan et al., 2004) found rates of acute gastroenteritis were significantly lower in skilled or unskilled manual workers compared to professional or non-manual workers. In contrast, reports of diarrhoea were found to be significantly higher in more disadvantaged areas although there was no significant difference for vomiting (Beale et al., 2010) and hospital admissions were significantly higher in more disadvantaged areas (Olowokure et al., 1999).

For children, fewer studies were identified, but these found more consistent results, reporting higher risk of GI infections in more disadvantaged children. Risk of norovirus infection was significantly higher among disadvantaged children (Phillips et al., 2011). Hospital admissions for children with GI infections were found to be significantly higher in more disadvantaged areas (Olowokure et al., 1999, Pockett et al., 2011) and infants of mothers with lower levels of education were found to be more likely to have diarrhoea (Baker et al., 1998). One study explored GP presentation for rotavirus and found higher rates of rotavirus in children living in rented council accommodation and in accommodation with fewer rooms (Sethi et al., 2001).

Studies conducted in Europe

In European countries there is a similarly unclear picture of the relationship between GI infections and SES in adults using laboratory reports, with findings differing by pathogen. There were also differences between the use of income and education to measure SES. In Denmark, for example, no association was found between income or education level and STEC or other *Salmonella* infection (Simonsen et al., 2008), between income and *Salmonella* Typhimurium (Simonsen et al., 2008) and between education and *Salmonella* Enteritidis. Another study conducted in Finland found lower incidence of STEC infection in areas with lower education (Jalava et al., 2011) but a higher incidence of STEC infection in areas with a higher proportion of low income households with children. The Danish study also found lower income to be associated with lower risk of infection with *Campylobacter*, *Salmonella* Enteritidis

and *Yersinia enterocolitica* (Simonsen et al., 2008) and a higher risk of *Shigella* but lower education was associated with a lower risk of *Campylobacter* and a higher risk of *Salmonella* Typhimurium, *Shigella* and *Yersinia enterocolitica* (Simonsen et al., 2008). The relationship was similarly unclear for viral GI pathogens, with one study conducted in the Netherlands finding a lower risk of norovirus in individuals with lower education levels (De Wit et al., 2003) and no association between education and rotavirus infection (De Wit et al., 2003) and a German study finding increased incidence of rotavirus hospitalisation in areas with high levels of unemployment (Wilking et al., 2012).

The findings from studies using syndromic definitions of GI infections in adults were more consistent, with studies conducted in the Netherlands, France and Germany finding no association (Doorduyn et al., 2012, Van Cauteren et al., 2012, Wilking et al., 2013) or lower risk of GI infection in more disadvantaged individuals in the Netherlands and France (de Wit et al., 2001a, Van Cauteren et al., 2012), with the French study finding no association when using occupation but lower risk in disadvantaged individuals when using education (Van Cauteren et al., 2012).

For children, inconsistent results were found when using laboratory reports. In the Danish study reported above for adults, there was no association identified between income or education and the incidence of *Salmonella* Enteritidis (Simonsen et al., 2008), between income and *Yersinia enterocolitica* or between education and other *Salmonella*, STEC or *Shigella* for children (Simonsen et al., 2008). Lower income was associated with higher incidence of *Campylobacter, Salmonella* Typhimurium, other *Salmonella* and *Shigella* and a lower incidence of STEC infection (Simonsen et al., 2008). Lower education was associated with higher incidence of *Yersinia enterocolitica* but lower incidence of *Campylobacter* and *Salmonella* Typhimurium (Simonsen et al., 2008). In Italy, another study found significantly higher incidence of non-typhoidal *Salmonella* infection in children of unemployed or unskilled fathers (Borgnolo et al., 1996).

As with the UK studies, European studies exploring the relationship between syndromically-defined GI infection and SES in children present more consistent findings towards higher risk in more disadvantaged children or no association. Studies conducted in Portugal and Italy found no association between diarrhoea and

parental education (Barros and Lunet, 2003, Iacono et al., 2005) and studies conducted in Sweden and Denmark found higher risk of diarrhoea (Ludvigsson, 2006) and of hospitalisation for GI infections (Biering-Sørensen et al., 2012) in children of parents with low education levels.

Studies conducted outside Europe

Outside Europe, studies using laboratory reports and syndromic definitions of GI infections generally found lower risk in disadvantaged adults or no association. Several studies conducted in Australia and Canada found no association between SES and *Campylobacter,* non-typhoidal *Salmonella* infections (Gibney et al., 2017), *Salmonella* Enteritidis (Varga et al., 2013) and STEC (Pearl et al., 2009). Other studies, conducted in the USA, Canada and New Zealand, have found significantly lower rates of *Campylobacter* (Bemis et al., 2014, Green et al., 2006, Pyra et al., 2012, Rind and Pearce, 2010, Spencer et al., 2012), STEC, *Salmonella, Shigella* (Chang et al., 2009) and G*iardia* (Cohen et al., 2008) in disadvantaged areas. As identified in the European studies, different results using income and education as measures of SES were also identified, with lower education being associated with higher risk of *Salmonella* (Younus et al., 2007) and *Cryptosporidium* (Cohen et al., 2008) but no association identified for income (Cohen et al., 2008, Younus et al., 2007) in the USA.

Studies using syndromic definitions of GI infections found lower risk of GI infections was associated with more disadvantaged adults in the USA and Australia, (Fein et al., 1995, Hall et al., 2006, Herikstad et al., 2002, Jones et al., 2007, Pollard et al., 2014) or no association between SES and GI infection in studies conducted in New Zealand, Australia, the USA and Canada (Adlam et al., 2011, Bytzer et al., 2001, Herikstad et al., 2002, Majowicz et al., 2004, Sargeant et al., 2008).

Few studies outside of Europe were identified for children. One study conducted in Australia found no association between diarrhoea and vomiting and SES (Eaton-Evans and Dugdale, 1987) and studies conducted in the USA and Japan found higher incidence of *Campylobacter* (Bemis et al., 2014) and STEC (Sakuma et al., 2006) in disadvantaged areas.

2.5 Theoretical explanations

The literature on the association between SES and risk of GI infection presents mixed results. There is great heterogeneity across the studies in terms of design, analytical methods, exposure variables used and even definitions of what constitutes a GI infection, which presents a challenge to draw any definitive conclusions. Studies included in this review also suffer from severe limitations and speculate on reasons for observations without providing empirical evidence. In what follows, I will firstly discuss factors that may influence the results in the studies in this review. There will then follow a discussion of potential hypotheses generated by this review.

Artefact explanations

Few studies found lower risk of GI infection among more disadvantaged children. Those that did speculated that this finding was likely due to artefact explanations. The studies included in this review were observational. While cohort, case-control and cross-sectional studies are robust study designs for exploring the relationship between SES and GI infections they do introduce potential biases through selection bias, loss to follow-up and ecological fallacy. Studies requiring participation through surveys often reported that disadvantaged individuals were underrepresented and suffered from low response rates, while response rates for advantaged individuals were much higher and as such they were overrepresented in the samples.

Differential case ascertainment could result in disadvantaged individuals being less likely to be recorded in healthcare systems or have a laboratory confirmation of a GI pathogen. This could be related to greater access to healthcare for those in higher education and income groups, particularly in countries with healthcare systems similar to that in the USA where insurance status may determine whether an individual will seek care and, if care is sought, whether the tests and medication required will be covered by insurance. This effect may be particularly so for adults which may, again, explain the difficulty in identifying clear trends in SES and risk in adults. Due to the way in which cases of GI infection are detected, there will undoubtedly be some relationship between wishing to seek healthcare and having the means to do so; a trend which will be more exaggerated in counties such as the USA where healthcare is not freely available to all but which may still exist to some extent in countries where there is theoretically equitable access. Foreign travel was

identified as a potential 'confounder' in several studies and therefore either controlled for in the analysis or excluded cases on this basis which may influence ascertainment of cases.

The use of SES measure was also variable in the studies included in this review. Several studies used measures which would not be sensitive enough to detect any differences, such as employed individuals compared to those who were not employed; the latter would likely include economically inactive individuals, those keeping home and the chronically sick and disabled. Some studies reported recording of occupation via free-text fields which would result in very imprecise indications of occupational class. Income and education were frequently used as proxies for SES. Education is easier and more reliably ascertained than self-reported income and thus may be a more reliable indicator of SES in some countries, but suffers from technical difficulties in the UK.

It has also been speculated that those with a higher level of education may be more likely to seek care (Younus et al., 2007) and also that they may be more likely to seek care with mild or moderate symptoms than those with lower education (Younus et al., 2007). Conversely, there may be differences in healthcare interaction amongst those of a lower SES or differential testing and treatment practices by SES which may increase the likelihood of seeking care, being recorded in healthcare systems such as hospital records or national surveillance records through laboratory confirmation. Tam et al. (2003) explored the potential biases in individuals who presented to primary care in the UK, and found that those with a lower level of education, those living in flats or maisonettes and those living in rented accommodation were more likely to present to healthcare. Tam et al. (2003) adjusted for illness severity. It is also speculated that the perceived increased risk amongst more disadvantaged groups may also reflect greater disease severity or worse consequences of GI infection in these groups. These hypotheses will be tested in the empirical studies in this thesis.

Several studies speculated that there may be regional differences in testing, treatment or reporting of GI infections. In countries where there is a fee to see a GP, such as New Zealand, it is plausible that rates of GP-reported GI infections would differ by area-level disadvantage and, in insurance-led healthcare systems such as the USA,

the middle classes may be least likely to be able to access healthcare because they are not eligible for Medicaid (and other similar schemes) which cover the most disadvantaged sections of the population and they are less able to afford private healthcare fees than their more wealthy counterparts or those in jobs where healthcare insurance is included in employee contracts.

Possible explanations for why there may be lower actual risk of GI infections in disadvantaged adults

It is also possible that studies finding lower risk of GI infection in disadvantaged individual are detecting a genuine effect which could be related to differential exposure. Many studies finding this association speculated that higher SES is likely to lead to higher rates of foreign travel and therefore higher risk of GI infection. Foreign travel may be a marker of wealth and therefore higher SES, so excluding such cases may result in controlling out the effect of SES. It is also possible that foreign travel exposes an individual to a greater risk of GI infection.

Other studies speculated that lower risk of GI infection in disadvantaged individuals may be as a result of lower pet ownership, less likelihood of eating outside of the home and lower exposure to environmental risk factors than those with higher levels of disposable income. Another possible explanation is that more disadvantaged individuals have better hand and food hygiene which could reduce their risk of acquiring a GI infection.

Possible explanations for why there may be higher actual risk of GI infections in disadvantaged adults

For those studies reporting a higher risk of GI infection in disadvantaged adults, factors such as overcrowding may result in poorer quality or shared facilities for cooking and washing or higher proximity to and contact with pets, all of which may increase the risk and spread of GI infection for disadvantaged adults. The relationship between disadvantaged areas and exposure to nutritionally-poor takeaways and fast-food outlets (Saunders et al., 2015) could result in higher risk through poor diet but also potentially through poor food hygiene in such premises. One study speculated that increased risk of *Listeria* among disadvantaged individuals was potentially due to disadvantaged individuals being less likely to store food safely and more likely to keep and consume food which has expired (Gillespie et al.,

2010b). A poorer immune system, greater likelihood of comorbidities, poorer diet and poorer hygiene may all be contributing factors to a decreased resistance to and therefore increased risk of infection. Comorbidity is more prevalent and arises at a younger age in lower SES groups (Barnett et al., 2012).

Possible explanations for why there may be higher actual risk of GI infections in disadvantaged children

Studies exploring the relationship between GI infection and SES found more consistent results for children; with disadvantaged children at higher risk of GI infections than more advantaged children. There are several hypotheses for this finding. One explanation for why this relationship is seen in children but not adults is that low income parents may not seek care for themselves but may ensure that care is sought promptly for their children, thereby ensuring that GI infections in disadvantaged children are more accurately recorded in the routine infectious disease data. Once care is sought, there may be differences in the way that the healthcare system interacts with disadvantaged compared to advantaged children. GPs, for example, may be more likely to test and physicians may be more likely to admit disadvantaged children to hospital if there are potential concerns about effective rehydration at home. Disadvantaged children may also be differentially exposed to GI infections compared to more advantaged children. It is also possible that disadvantaged individuals may be exposed earlier in life than more advantaged individuals; differential exposure by SES in childhood may explain the higher risk of GI infection in disadvantaged children and the lower risk of GI infection in disadvantaged adults. A study conducted in the UK found that seropositivity to *Helicobacter pylori* in adults was associated with lower SES and adverse housing conditions in childhood (Pearce et al., 2013). Furthermore, within low income countries, C*ampylobacter* is almost exclusively seen in disadvantaged children (Fernández et al., 2008, Kakai et al., 1995, Lloyd-Evans et al., 1983, Quetz et al., 2010) while adults are rarely infected or identified (Coker et al., 2002).

2.6 Gaps in the literature

This literature review has identified several gaps in the literature of the role of SES in the risk of and exposure to GI infections.

- A systematic review of the literature is warranted to make sense of the conflicting and inconsistent findings in the existing literature. This gap is filled through the systematic review and meta-analysis conducted in Study 1 (Chapter 4).
- The first national population-based survey, the IID1 Study, was conducted in the 1990s and many of the existing studies, particularly in the UK are relatively old. It is therefore necessary to provide an up to date assessment of the relationship between GI infections and SES at the population-level. This gap is filled by analysing a prospective community cohort from the IID2 study, conducted in 2008-2009 and analysis of calls to NHS telephone helplines (Studies 2 and 3, Chapters 5 and 6).
- Little is known about the role of differential exposure leading to differential risk. Understanding the mechanisms through which certain population groups may be at greater risk of exposure to a GI pathogen may improve our understanding of differential disease risk. This is particularly true for differences between SES groups and hypothesised risk factors such as pet ownership, eating out, foreign travel and healthcare interaction for which no empirical evidence has been identified. Differential exposure and healthcare contact explanations are explored in Study 4 (Chapter 7).
- Existing studies have generally relied on laboratory confirmation of a GI pathogen in order to explore the role of SES in risk and exposure. This could lead to a number of potential biases, particularly if the pathway to a laboratory diagnosis differs by SES. Understanding the association between SES and GI infections defined in other ways, such as syndromically, without reliance on laboratory testing is therefore important as differential risk may be masked by the way in which cases of GI infection are defined. Studies 1-3 (Chapters 4-6) address this gap by using data which are not reliant on an individual seeking care for inclusion in the studies.
- Previous studies have found different results across the life course. Understanding the association between GI infection and SES for adults and children separately is crucial. The risk of GI infections by SES is explored throughout the life course in each of the studies in this thesis.

2.7 Summary

Gastrointestinal infections are common and socioeconomic inequalities are an important public health concern. Despite this, little is known about the relationship between SES and risk of or exposure to GI infections, with published literature presenting conflicting findings. The burden of GI infections in the UK is high, and there is no consensus on whether this is socially patterned in adults, though there is some evidence in children. Despite the undisputed association between socioeconomic inequalities and ill- health, the evidence on the role of such inequalities in the risk of and exposure to GI infections is inconsistent and contradictory. Existing studies have presented hypotheses to explain the observed results but have provided limited, or no, empirical evidence to support or refute these hypotheses.

In the subsequent chapters, I seek to fill these gaps in the literature and try to consolidate the inconsistent findings seen. A systematic review and meta-analysis was warranted to analyse the inconsistent existing evidence and the gaps in knowledge in a more structured way in order to inform the subsequent empirical analyses in this thesis (Study 1). Making use of existing datasets which do not rely on laboratory confirmation, I explore the role of socioeconomic inequalities in GI infections via a population-based prospective cohort study and two telephone-based healthcare advice services to assess whether socioeconomic inequalities in GI infections exist and provide updated estimates of the extent and nature of any inequality identified. (Studies 2 and 3) Finally, to explore the role of differential exposure, I analyse data for a laboratory-confirmed severe GI pathogen (STEC) and subsequent development of a serious sequela (HUS) (Study 4).

Chapter 3
Methods

3.1 Introduction

This chapter provides a description of the various data sources and analytical methods used in the analyses presented in Chapters 4-7. Firstly, the five main datasets used will be described; the Second Study of Infectious Intestinal Disease in the community (IID2); the syndromic surveillance datasets – NHS111 and NHS Direct; the National Enhanced Surveillance System for STEC (NESSS); and the British Paediatric Surveillance Unit (BPSU) Haemolytic Uraemic Syndrome (HUS) Study; alongside supplementary datasets used to generate the socioeconomic and rurality variables. There will then follow a description of the various methods used in each of the four study chapters, including details of ethical approvals (where necessary), inclusion and exclusion criteria, primary outcomes and covariates, data extraction and data analysis. Finally, I will summarise by exploring how the different datasets and analytical methods, when brought together, answer the research questions.

As described in Chapter 1, the objectives of this thesis are:

1. To conduct a systematic review of existing evidence of socioeconomic inequalities in risk of GI infections in high income countries.
2. To investigate the extent and nature of socioeconomic inequalities in risk of GI infections in the community in the UK, with estimates derived from the most up-to-date population-based household survey.
3. To analyse the extent of, and mechanisms underlying, socioeconomic inequalities in risk of GI infections in the community, with estimates derived from routine data on members of the public seeking telephone-based healthcare advice in England.
4. To explore the social patterning of clinical outcomes, healthcare contact and risk factors for a laboratory-confirmed, potentially severe, GI infection (STEC) and socio-demographic inequalities in risk of development of a serious sequela (HUS) in England.
5. To draw out policy implications and recommendations for further research into the role of socioeconomic inequalities in GI infections.

The methods described in the subsequent sections within this chapter were each selected to address the relevant objectives in the most appropriate way.

3.2 Description of data sources

The Second Study of Infectious Intestinal Disease in the community (IID2 Study)

The IID2 Study was commissioned to determine whether the incidence of IID had changed since first Study of Infectious Intestinal Disease in the community was undertaken in the mid-1990s and to recalibrate national surveillance data (Tam et al., 2011a). There were seven separate but linked studies which took place between 28th April 2008 and 31st August 2009, with the exception of the Telephone Survey which took place from 1st February 2008 to 31st August 2009. This study was undertaken across England, Wales and Scotland.

For the purposes of this thesis, data from the prospective population-based cohort study were used. In this study, a cohort of 7,033 randomly-selected individuals was recruited from 88 General Practices across the UK, resulting in 6,836 participants who were eligible for inclusion in the study. The participants were followed up weekly for up to one year to investigate how many experienced symptoms consistent with IID during this period. Those who reported symptoms were asked to complete a questionnaire about their illness, as well as details of their interaction with healthcare services and were sent a stool specimen kit and asked to submit a sample. Analyses using this dataset will be presented in Chapter 5 (Study 2).

NHS Direct

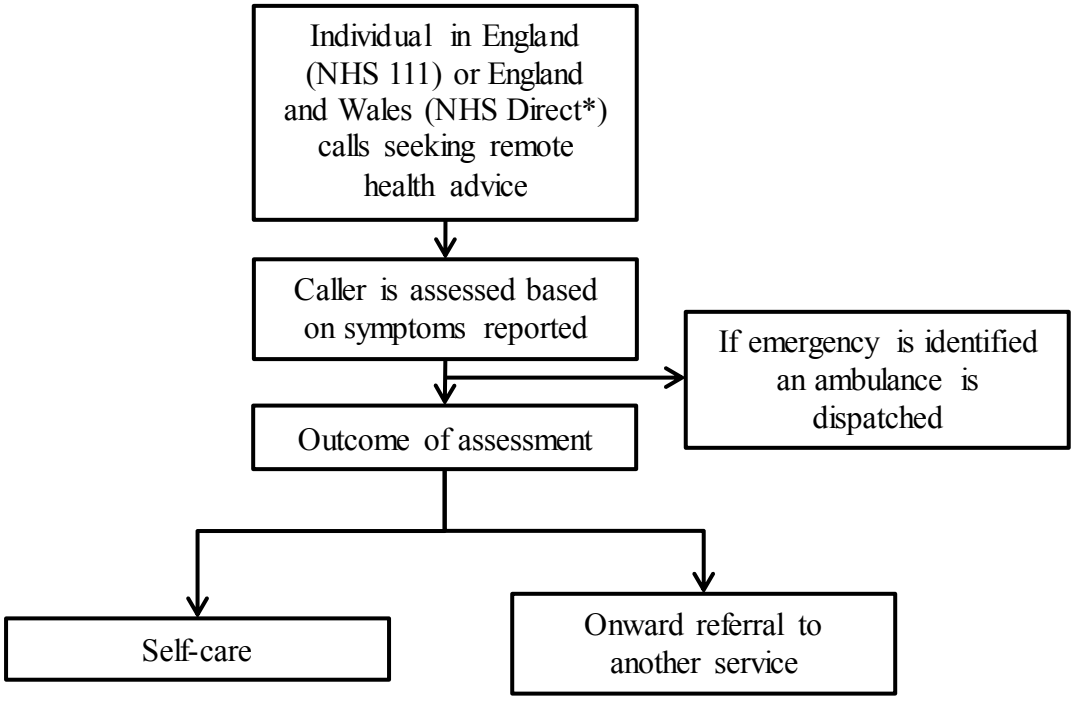
Telephone-accessed healthcare encompasses advice delivered to individuals over the telephone or online as opposed to face-to-face consultation. The first service, NHS Direct (NHS Direct, 2014), was a nurse-led telephone health helpline for non-emergencies established in 1999 and was available across England and Wales by 2000. The service operated 24 hours a day, 365 days a year, responding to approximately 5 million calls per year (Cooper et al., 2003). A computerised decision making system through which clinical algorithms could be followed was used by the NHS Direct nurses to assess each caller based on their symptoms and provide healthcare advice or referral to other NHS services (Cooper et al., 2003). Calls to NHS Direct were charged as a business rate number.

NHS 111

In 2013, a new service replaced NHS Direct. This new system, NHS 111, is operated by non-clinical call takers using with access to clinical advice if required. As with NHS Direct, this service is available 24 hours a day, 365 days a year, although unlike NHS Direct, this service is for callers from England only and is a Freephone service. Furthermore, NHS 111 also acts as an out of hours GP service which means call volumes are higher than for NHS Direct. Figure 3.1 describes the procedure which individuals may go through when accessing the telephone-accessed healthcare. Analyses using both NHS Direct and NHS 111 call data from the HPA/PHE NHS Direct/111 syndromic surveillance systems, based upon data routinely collected and used by PHE for routine public health surveillance, will be presented in Chapter 6 (Study 3).

Figure 3.1: Procedure for telephone-accessed healthcare

Source: National Health Service (2015)



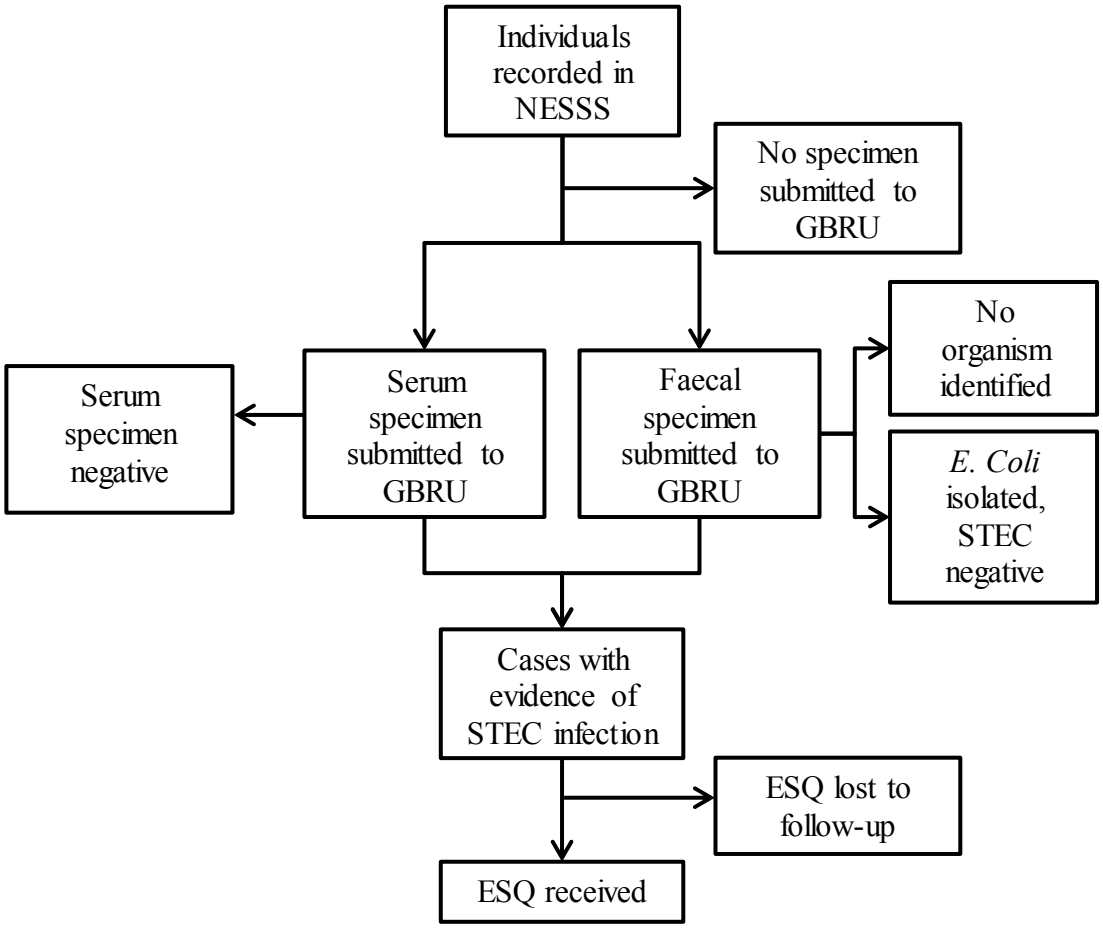
*NHS Direct is no longer operational and was succeeded by the NHS 111 system; therefore these two systems do not co-exist

National Enhanced Surveillance System for Shiga toxin-producing Escherichia coli (NESSS)

In January 2009, Public Health England (PHE) introduced a national enhanced surveillance system for STEC in England (NESSS). This surveillance system collects standardised information for each laboratory confirmed case of STEC reported in England. Following identification of a presumptive STEC at a local laboratory or the identification of a case of HUS, specimens are sent to the PHE Gastrointestinal Bacteria Reference Unit (GBRU) for testing. The relevant PHE region is notified for follow-up which includes completion of an enhanced surveillance questionnaire (ESQ) (Appendix 5.1) collecting detailed information on patient demographics, symptoms, food and water exposures and UK and non-UK travel during the exposure period (the week prior to illness onset). Data from ESQs are reconciled with microbiological results. This surveillance system is described in detail elsewhere (Byrne et al., 2015). Risk factor questions included in the ESQ are evidence-based, using known-exposures documented in the literature as well as exposures identified as part of outbreak investigations (Al-Jader et al., 1999, Gillespie et al., 2005, Goh et al., 2002, Ihekweazu et al., 2012, Launders et al., 2015, Locking et al., 2001, Parry et al., 1998, Pennington, 2009). Approximately, 1,000 cases of STEC are reported to NESSS each year (Public Health England Gastrointestinal Bacteria Reference Unit, 2017). The most commonly reported serotype is STEC O157, but since December 2012 some frontline laboratories have been able to detect non-O157 STEC due to the implementation of multiplex polymerase chain reaction (PCR) testing leading to an increase in the detection of these serotypes (Byrne et al., 2014b). Figure 3.2 shows the flow of STEC cases reported to this system. Analyses using this dataset, along with further detail on STEC, will be presented in Chapter 7 (Study 4).

Figure 3.2: Flowchart of STEC cases reported via NESSS

Source: Adapted from Byrne et al. (2015; p.4)



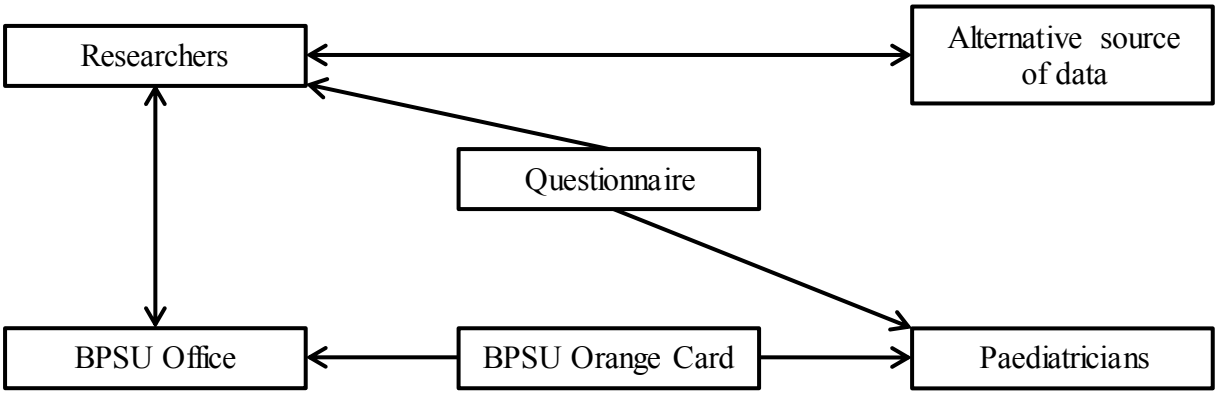
STEC – Shiga toxin-producing *Escherichia coli* (*E. coli*); NESSS – National Enhanced Surveillance System for STEC; GBRU – Gastrointestinal Bacteria Reference Unit; ESQ – Enhanced Surveillance Questionnaire

British Paediatric Surveillance Unit (BPSU) Haemolytic Uraemic Syndrome (HUS) Study

Between 1st October 2011 and 31st October 2015 a surveillance study of HUS was conducted by the British Paediatric Surveillance Unit (BPSU), in conjunction with PHE. The BPSU facilitates studies on rare childhood diseases, or rare complications of childhood diseases, where the necessity for large sample sizes requires a national approach (British Paediatric Surveillance Unit, 2017). The BPSU collects data via their 'orange card' surveillance system which is an electronic reporting card through which doctors participating in the scheme are requested to indicate if they have seen any patient with any condition listed on the form within the preceding month (British Paediatric Surveillance Unit, 2016). This reporting card is sent to over 3,300 doctors each month (British Paediatric Surveillance Unit, 2016). Figure 3.3 shows the flow of reporting to the BPSU.

Figure 3.3: Flowchart of reporting to BPSU

Source: British Paediatric Surveillance Unit (2016)



BPSU – British Paediatric Surveillance Unit

For the purposes of this study, the following reporting definition for HUS was used:

> *"Any child up to and including 15 years of age seen for the first time in the last month and who has been diagnosed with HUS, excluding children who have; septicaemia; malignant hypertension; chronic uraemia; primary vascular disease"*
>
> *(British Paediatric Surveillance Unit, 2014)*

A case of HUS meeting this reporting definition triggered notification to PHE investigators who sent a standardised surveillance questionnaire (Appendix 5.2) to the paediatrician requesting information on; case demography, diarrhoeal prodrome; household contacts; treatment history; microbiological investigations; clinical parameters of illness; clinical management of illness and status of the case at the time of data capture. Analyses using this dataset, along with further detail on HUS, will be presented in Chapter 7 (Study 4).

The National Statistics Socio-economic Classification

The Office for National Statistics (ONS) produces a National Statistics Socio-economic classification (NS-SEC) which uses an individual's occupation and employment status as a measure of the socioeconomic position of individuals within society (Office for National Statistics, 2010). The NS-SEC evolved from earlier measures of socioeconomic position (Social Class and Socioeconomic Group) in order to better explain the role of socioeconomic position in explaining the patterning of outcomes of interest (Office for National Statistics, 2010), be that in health, economics or crime. Since 2001, the NS-SEC has been available and used in all official statistics and surveys (Office for National Statistics, 2010), including the 2001 and 2011 Censuses.

Analyses making use of this classification system will be presented in Chapter 5. In the IID2 Study the five-class self-coded method was used. This version of NS-SEC was used as it is suitable for use in studies such as those collecting data via postal surveys for which it may not be practical to collect and code detailed occupation data (Office for National Statistics, 2010). For the purposes of the analyses in Chapter 5, this five-class version was recoded into the three-class version (Office for National

Statistics, 2010). The relationship between the five- and three-class versions can be seen in Table 3.1.

Table 3.1: Overview of the five-class and three-class NS-SEC

Source: Office for National Statistics (2010)

Class	NS-SEC five-class self-coded version	NS-SEC three-class self-coded version
1	Managerial, administrative and professional occupations	Higher managerial, administrative and professional occupations
2	Intermediate occupations	Intermediate occupations
3	Small employers and own account workers	
4	Lower supervisory and technical occupations	Routine and manual occupations
5	Semi-routine and routine occupations	
	Never worked and long-term unemployed*	Never worked and long-term unemployed*

*Not classifiable

English Index of Deprivation 2010

The English Index of Deprivation 2010 is a measure of multiple deprivation used in England at a small area level (Department for Communities and Local Government, 2011). They comprise 38 indicators across seven domains (Income; Employment; Health and Disability; Education Skills and Training; Barriers to Housing and Other Services; Crime; Living Environment) which, combined and weighted (Table 3.2), are used to calculate the Index of Multiple Deprivation 2010 (IMD 2010) (Department for Communities and Local Government, 2011). The score for each domain, and the combined overall score, are calculated and assigned to each individual living in a small area (Lower layer Super Output Areas; LSOA). There are 34,753 LSOAs in England and Wales and each has a population of approximately 1,500 residents and 650 households (Neighbourhood Statistics, n.d.). Postcodes of individuals can be linked to the IMD score for the LSOA in which they reside.

Table 3.2: IMD domains and weighting

Source: Department for Communities and Local Government (2011)

Domain	Weight
Income deprivation	22.5%
Employment deprivation	22.5%
Health deprivation and disability	13.5%
Education, skills and training deprivation	13.5%
Barriers to housing and services	9.3%
Crime	9.3%
Living environment deprivation	9.3%

Further details of how the IMD score was used in each study will be provided in the relevant sections below and in each study chapter. For ease of interpretation and presentation, the IMD scores were also categorised into quintiles, as set out below, based on the distribution of IMD scores across England (Table 3.3).

Table 3.3: IMD quintiles

Source: University of Oxford (n.d.)

Quintile	IMD score range
1 (Least disadvantaged)	≤ 8.49
2	8.5 - 13.79
3	13.8 - 21.35
4	21.36 - 34.17
5 (Most disadvantaged)	≥ 34.18

Rural Urban Classification

The Rural-Urban Classification (RUC) is an official statistic used to determine the rurality of small areas. As with IMD, RUC can be assigned to LSOAs. Areas outside of settlements of over 10,000 residents are considered rural (Department for Environment, Food and Rural Affairs, 2016). Each LSOA can be assigned to one of eight rural or urban categories (Table 3.4).

Analyses making use of this classification system will be presented in Chapter 5, 6 and 7 (Studies 2, 3 and 4). In these studies, the RUC was recoded to either rural or urban, as detailed in Table 3.4.

Table 3.4: Overview of Rural-Urban Classification

Source: Government Statistical Service (2013)

	Category
Urban	Major Conurbation
	Minor Conurbation
	City and Town
	City and Town in a Sparse Setting
Rural	Town and Fringe
	Town and Fringe in a Sparse Setting
	Village and Dispersed
	Village and Dispersed in a Sparse Setting

Postcode Headcounts and Household Estimates

The Postcode headcounts dataset, from the 2011 census provides estimates of the number of usual residents for each unit postcode in England and Wales (Office for National Statistics, 2011). Analysis making use of this reference population and details of how this dataset was used will be provided in the relevant study methods section below and will be presented in Chapter 6 (Study 3).

Mid-year Population Estimates

Mid-year population estimates, and population estimates stratified by age, sex and IMD Quintile from the Office for National Statistics Office for National Statistics [dataset] (2011), were downloaded and used to provide population denominator data for the calculation of incidence rates in Chapter 7 (Study 4).

Human Development Index

The Human Development Index (HDI) (United Nations Development Programme, 2016) is a composite measure of average achievement across three dimensions measuring life expectancy at birth; years of schooling; and Gross National Income

per capita. The score for each of these dimensions is aggregated to calculate the overall HDI value. In Chapter 4 (Study 1), the HDI value was used to categorise the countries from which the studies originated by relative level of development into tertiles of low, medium or high HDI.

Köppen Climate Classification System

A simplified version of the Köppen system for classifying climates by the Met Office, the national weather service for the UK, was used to classify the climate of each country included in Chapter 4 (Study 1) to assess whether climate was a potential moderating factor for risk of GI infection. This system classifies the climate of each country based on annual and monthly averages of temperature and precipitation into six broad categories; equatorial; arid; Mediterranean; snow; polar; and temperate (Met Office, 2015).

3.3 Ethical approval

Ethical approval for the research was obtained from the NHS Health Research Authority South East Coast – Surrey Research Ethics Committee on 4th December 2015 (REC Reference number 15/LO/2138; Appendix 1.1). Further details of ethical approval covering individual studies and datasets can be found in each subsequent section below.

3.4 General overview of methods

Systematic reviews and meta-analyses

A systematic review is a robust and evidence-based method of collating empirical evidence based on pre-specified criteria to address a specific research question (Higgins and Green, 2011). While literature reviews are able to identify and summarise important themes within the literature, a systematic review aims to minimise bias by using explicit and systematic methods which results in more reliable and robust findings (Uman, 2011). Within the process of a systematic review, specific consideration is given to the reliability and validity of included studies through assessment of the risk of bias and detailed synthesis and presentation of result (Higgins and Green, 2011, Uman, 2011). Included studies may be drawn from published and unpublished literature, including grey literature. In addition, meta-analysis can be used alongside systematic reviews. Meta-analysis is designed to

critically evaluate and statistically combine results from studies identified in a systematic review to provide an overall quantifiable effect estimate.

Regression methods

Regression modelling, used throughout this thesis, models the effect of an explanatory variable, or variables, on a response variable. The explanatory variable may be, for example, an exposure, risk factor or characteristics of participants. In multiple regression models, potentially confounding variables can be included in order to provide an adjusted effect estimate. I will describe the regression models used in this thesis in general below and explain the specific uses within the description of methods for each study.

Logistic regression models describe the effect of predictor variables on categorical outcomes which can be binary, ordinal or multinomial. Logistic regression produces estimates of odds ratios and is a more powerful and flexible approach than other traditional methods such as Mantel-Haenszel as multiple confounding variables, interaction terms and continuous, binary or categorical exposure variables may also be included within the model. Logistic regression can also be described in terms of generalised linear models (GLM) which are extensions of linear regression which can be used to model binary data, as well as continuous, count or survival data. GLMs allow for response variables which have non-normal error distributions and consist of three components: a random or systematic component, a linear predictor and a link function (Fox, 2016). The random component specifies the conditional distribution of the response variable, such as Gaussian (normal), binomial, multinomial, Poisson, gamma or inverse-Gaussian (Fox, 2016). The linear predictor is a linear function of the regressors and the link function transforms the expectation of the response variable (Fox, 2016). GLMs can be estimated using maximum-likelihood estimations, providing estimates of regression coefficients as well as estimated asymptotic standard errors of the coefficients (Fox, 2016). There are several assumptions of a GLM; the observations of the dependent variable are assumed to be independent, the dependent variable is assumed to be a member of the exponential family of distributions; the variance is assumed to be constant across observations; and a linear relationship between the response and the linear predictor is assumed (Dunteman and Ho, 2006, Grafen and Hails, 2002). These assumptions

can be tested through visual inspection of the residuals to assess homogeneity of the variance and linearity.

The Cox proportional hazards model is a regression model for the analysis of survival data, allowing for the testing of differences in survival times in two or more groups of interest while also allowing for the control of other covariates. The hazard function gives the risk or hazard that the event will occur, per time unit, given that an individual has survived to the specified time. As such, it provides an estimate of relative risk, not absolute risk, by estimating the hazard ratio for an outcome of interest adjusted for the other covariates in the model rather than an odds ratio as in logistic regression. The Cox proportional hazards model is "non-parametric" as no assumptions are made about the form of the baseline hazard. In addition, there are a number of important assumptions that need to be tested. Firstly, Cox regression assumes that there is no multicollinearity among covariates. Secondly, the censoring of participants must not be related to the outcome. Finally, the proportional hazards assumption assumes that the survival curves have hazard functions that are proportional over time. These assumptions can be assessed using log-log plots to test the proportional hazards assumption for each predictor and using a global test to assess whether the overall model overall met the proportional hazards assumption (UCLA Institute for Digital Research and Education, 2017b). Kaplan-Meier survival curves can be constructed to estimate the survival function from the data. An alternative to Cox Proportional Hazards is Poisson regression, which is used when count of events over time is the desired outcome rather than survivorship/failure as for Cox Proportional Hazards.

Meta-regression (Berkey et al., 1995, Thompson and Sharp, 1999), similar to simple regression where an outcome is predicted according to one or more explanatory variables, is a regression method used in meta-analysis. The explanatory variables are study characteristics which may influence the effect estimate (Higgins and Green, 2011). Meta-regression differs from simple regression as studies are weighted by the precision of their effect estimate and, in random-effects meta-regression, residual heterogeneity among study effects is allowed for (Higgins and Green, 2011). Fixed-effects meta-regression does not take account of this residual heterogeneity and is therefore likely to produce misleading results (Higgins and Thompson, 2004).

Missing data methods

Multiple Imputation using Chained Equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) can be used to statistically simulate missing covariates such that the records with missing covariates may be included in analyses. This approach uses the distribution of the observed data to estimate multiple imputations, rather than single imputations, reflecting the uncertainty of the true values before combining these imputed values for analysis (UCLA Institute for Digital Research and Education, 2017a). All other variables were included in the imputation in order to produce the most accurate imputation possible.

The main assumption of MICE is that the missing data are missing at random (MAR). This assumption can be tested by assessing the distribution of the missing variable by other variables, such as age and sex. Including all variables thought to be associated with or predict missingness in multiple imputation models can help to provide more accurate and stable estimates and including these variables can increase power through a more inclusive analysis strategy.

There are alternative methods for dealing with missing data. Complete case analysis, whereby only individuals with no missing data in any of the variables required in the analysis are included, is often used but this can lead to biased results, particularly because there can be a cumulative effect of missing data leading to a substantial proportion of data being excluded and a resultant loss of power and precision (Sterne et al., 2009). It is also possible to use single, as opposed to multiple, imputation, although as this method does not account for uncertainty in the imputations generated, single imputation was rejected in preference of multiple imputation.

3.5 Description of methods specific to each study

This section reports on the study design, data collection, data handling and analytical methods used in each of the subsequent study chapters.

Study 1: Relationship between socioeconomic status and gastrointestinal infections in high income countries: A systematic review and meta-analysis (Chapter 4)

In Chapter 4 I present results from a systematic review and meta-analysis assessing the association between SES and risk of GI infections in high income countries. The methods used in this study chapter are detailed below. This systematic review and

meta-analysis is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (Moher et al., 2009).

Rationale

The review of the literature, presented and discussed in Chapter 2, demonstrated the conflicting results from studies investigating the relationship between SES and GI infection. Due to the heterogeneity in study design, measure of SES, measure of GI infection and populations investigated, it was challenging to draw any conclusions on magnitude and direction of the SES impact on GI infection using a non-statistical approach. A systematic review and meta-analysis was therefore warranted to collate and summarise these diverse studies in a meaningful way using a systematic and statistical approach.

Objectives

This systematic review and meta-analysis aimed to consolidate and evaluate the current level of knowledge on the relationship between GI infections and socioeconomic inequalities and identify any gaps in the literature in order to shape future studies into this relationship.

Table 3.5: Statement of question being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)

Research question: For individuals from high income countries, is lower compared to higher SES associated with the incidence or prevalence of GI infection?	
Participants	Individuals from high income countries
Interventions	Low SES
Comparisons	High SES
Outcomes	Incidence or prevalence of GI infection
Study design	Observational studies

The exposure of interest was "lower compared to higher SES", measured at the individual or aggregate level by income, education, occupation, employment or deprivation of area of residence. The primary outcome of interest was the incidence or prevalence of any symptomatic GI infection measured using population level surveys, routine surveillance systems, laboratory or hospitalisation data, and included

syndromic definitions of GI infections without a laboratory diagnosis. Syndromic definitions of GI infections were included as it is possible that various socioeconomic or healthcare seeking behavioural factors could influence whether an individual is diagnosed with a GI infection.

Protocol and registration

This systematic review is registered in the PROSPERO International prospective register of systematic reviews (PROSPERO 2015:CRD42015027231(Adams et al., 2015)). The protocol for this systematic review can be found in Appendix 2.1 (Rose et al., 2016).

Eligibility criteria

For inclusion in the systematic review, studies were required to be observational (cross-sectional; ecological; case-control; cohort - prospective and retrospective), reporting quantitative results and analysis of empirical data on the prevalence or incidence of any symptomatic GI infection by SES, in a representative population sample. Socioeconomic status could be measured by occupation, income, education, employment or deprivation at individual or aggregate level in order to be inclusive and assess how measures of SES were used in the literature.

Only studies conducted in high income countries, defined as being a member country of the Organisation for Economic Co-operation and Development (OECD) (Organisation for Economic Co-operation and Development, 2015), written in or translated into English, reporting on human subjects and using data collected after 1980 and up to the search date of 13th October 2015 were included. For countries that joined the OECD after 1980, data collection must have occurred after the date the country became a member of the OECD. The focus was on high income countries as they were thought more likely to have similar public health infrastructures, social characteristics and food production and safety standards which would be comparable to both the UK and to each other. As social conditions within countries change over time through development, and methods of classifying SES may also change over time, restricting to publications using data from 1980 onwards ensured that the results were as relevant as possible to the present day.

Studies not meeting the above criteria, including case studies, case series or literature reviews, or studies reporting on outbreaks of GI infection, travel associated illness only or asymptomatic individuals only were excluded. Studies conducted solely in a specific population subgroup without a general population comparator group, or studies conducted in institutional settings such as nurseries, hospitals or the military were excluded. Table 3.6 reports inclusion and exclusion criteria in detail.

Table 3.6: Inclusion and exclusion criteria

Inclusion criteria
- Studies quantitatively measuring the prevalence or incidence of any symptomatic gastrointestinal infection in a representative population sample
- Studies quantitatively measuring socioeconomic status at an individual or aggregate level by occupation, income, education, employment or area deprivation
- Studies reporting a quantitative association between the first two inclusion criteria i.e. reporting an association between gastrointestinal infection and socioeconomic status
- Studies written or translated into English language
- Studies reporting on human subjects
- Subjects selected from the populations of countries that are members of the OECD, reporting data after 1980 or the date that they became a member of the OECD
- Studies reporting on data collected after 1980
- Observational studies
Exclusion criteria
- Studies not providing a quantitative measure
- Unrepresentative population sample
- Outbreak reports
- Studies analysing travel related cases only
- Review studies
- Case reports
- Studies not available in English

Information sources

The pilot application of search terms was conducted in May 2015. Systematic searches of three databases were undertaken (MEDLINE (Ovid); Scopus and Web of Science Core Collection) on 13th October 2015. Hand searching via reference lists of studies deemed eligible for inclusion was undertaken. Experts in the field of GI infections were consulted to identify any additional studies that were not found using

the above information sources. In addition, grey literature searches were performed using the Google Scholar search application and the Google internet search engine on 11th and 12th January 2016 respectively. For included studies which lacked quantitative data to allow for inclusion in meta-analyses, authors were also contacted for additional information where possible.

Search

Three search strategies were employed. A pilot review was conducted prior to the final systematic review, in May 2015. The aim was to identify relevant search terms and to test these search terms and the eligibility criteria used for screening in the review. In order to identify all relevant literature on the relationship between socioeconomic inequalities and gastrointestinal infections from biomedical, scientific and social science literature, three literature databases (MEDLINE, Scopus and Web of Science) were systematically searched using the piloted search terms. MeSH Subject Heading terms were used and these were adapted for each database as appropriate (Appendix 2.3). Using database filters, the results were restricted to publications which analysed data from 1980 to the date of the search only (13th October 2015). Boolean search operators (and, or) were used to link search terms in order to search for any of the GI infection terms, socioeconomic terms and geographical terms in conjunction. An illustration of the search terms used is reported in table 3.7.

A range of socioeconomic terms and measures were considered in order to capture the variety of methods used to define individual- or area-based socioeconomic status, as identified by the pilot review (Appendix 2.3). The infection terms selected for this search were based upon gastrointestinal infections which are known to cause the greatest burden of disease in high income countries. Whilst not exhaustive, the list is intended to provide a broad spectrum of bacterial, viral and protozoal infections. The decision to include studies which looked at non-laboratory confirmed disease, as defined by symptoms experienced, was taken in order to be as inclusive as possible and to determine potential inequalities in healthcare interaction or reporting for gastrointestinal symptoms. Studies which reported on chronic or long-term gastrointestinal symptoms were not included.

Table 3.7: Illustration of search terms used

Infection terms	AND	**Socioeconomic terms**	AND	**Geographical terms**
Acute gastroenteritis		Depriv*		Australia
Bacillus cereus		Disadvantag*		Austria
Caliciviridae		Education*		Belgium
Campylobacter		Employment		Canada
Clostridium perfringens		Income*		Chile
Cryptosporidiidae		Inequalit*		Czech Republic
Diarrh*		Occupation*		Denmark
Diarrhea		Poorest		Estonia
Dysentery, Bacillary		Poverty		Finland
Enteric infection*		Salary		France
Enterobacteriaceae Infection*		Social Class		Germany
Escherichia coli		Social determinant*		Greece
Escherichia coli Infections		Social factor*		Hungary
Food poisoning*		Socio*		Iceland
Foodborne Diseases		Socioeconomic*		Ireland
Gastric flu		Socioeconomic Factors		Israel
Gastrointestinal bacteria		Underprivileged		Italy
Gastrointestinal infection*				Japan
Gastrointestinal pathogen*				Korea
Giardia				Luxembourg
Hepatitis A				Mexico
Hepatitis E				Netherlands
Infectious intestinal disease*				New Zealand
Listeria				Norway
Norovirus				Poland
Rotavirus				Portugal
Salmonella				Slovak Republic
Salmonella Infections				Slovenia
Sapovirus				Spain
Scombro*				Sweden
Shigella				Switzerland
Small round structured virus*				Turkey
STEC				United Kingdom
Stomach bug*				United States
Stomach flu				
Stomach virus*				
VTEC				
Winter vomiting disease*				
Yersinia enterocolitica				

*indicates truncated term to allow for different suffixes

Secondly, reference lists of studies selected for inclusion in the review were searched to identify potentially relevant articles that may have been missed by the electronic database searching. References deemed potentially relevant were sought and screened for inclusion using the pre-defined inclusion and exclusion criteria.

Finally, grey literature searching was conducted using a sub-set of the search terms ("gastrointestinal infection", "gastroenteritis", "diarrhoea", "diarrhea", "socioeconomic", "social class", "income", and "deprivation"), which were entered

into the Google internet search engine and Google Scholar search application and the first 100 results returned from each were screened for inclusion against the inclusion and exclusion criteria. References within the grey literature selected for inclusion were also reviewed against the same criteria. The results from the three searches were exported into reference managing software (Endnote version X8) where they were combined and duplicates removed.

Study selection

Studies were assessed against the inclusion and exclusion criteria (Table 3.6) based on titles and abstracts. To ensure consistency of the application of the criteria, studies were screened independently by two researchers and discrepancies reviewed and resolved through discussion. Studies deemed eligible for inclusion on this basis were then reviewed using the full-text in the same way. Where full-texts were not available online, they were sought via institutional library sharing agreements. Where this was not possible, authors of the studies were contacted directly via email. All full-text studies were screened independently by the same two reviewers to ensure that they conformed to the inclusion and exclusion criteria.

Details of included studies are presented in Appendix 2.4. Studies which reported quantitative results or reported quantitative results in-text without data or results tables to corroborate the findings were selected for inclusion in the synthesis using Harvest Plots. Where results were reported in-text but without quantitative results or where clarifications on data items were required, study authors were contacted and asked to provide further details. Where it was not possible to gain quantitative results, either from the full-text or through contact with authors, these studies were excluded from the meta-analyses. Full details of the studies excluded from the meta-analyses are provided in Appendix 2.4. Details of the studies excluded following full text review are provided in Appendix 2.5.

Data collection process

A record of databases, terms and dates of searches was kept and Endnote (version X8) was used to record and store references generated by the search. Data were extracted into a standardised Excel spreadsheet by one reviewer and were checked for accuracy by the second reviewer. This standardised data extraction form was

piloted prior to use. Where information was missing or unclear, authors were contacted to obtain the relevant details.

Data items

For each study, information on the aim/hypothesis; study design; level of analysis; country; sample size; age of individuals in the sample under analysis; measurement of GI infection; measurement of SES; covariates and results of each study (Appendix 2.6) were recorded. Where studies did not present quantitative results via results tables and it was not possible to obtain further details from authors, these results were recorded as non-significant findings. Studies which presented adjusted results were recorded and noted, in addition to unadjusted results. Results from multivariate and univariate analyses were recorded separately. Where more than one study design was reported, the study design which yielded the results was recorded. Where multiple SES and GI measures were used, all were reported separately.

Risk of bias in individual studies

Risk of bias and quality assessment of the identified studies were conducted by each reviewer independently and then reconciled. The Liverpool University Quality Assessment Tool (LQAT) was used for this review, which allowed for the methodological quality of the studies to be assessed using a tool specific to each study design (Pope, 2015). The LQAT incorporates a star rating system to assess and quantify absence of bias, misclassification and confounding and the number of stars available differed by study design (Appendix 2.7). Stars were awarded for absence of bias and the combined total number of stars obtained after reviewing the study against all the relevant criteria was then converted to a percentage of the total stars available for each study design. Tertiles were created for each study design based on the quality score, with the highest tertile equating to high quality and the lowest tertile equating to low quality. Scores in the middle tertile were regarded as medium quality. The LQAT has been used in a number of other reviews (Puzzolo et al., 2013, Rehfuess et al., 2014) and has been independently evaluated against other quality assessment tools and compared favourably (Voss and Rehfuess, 2013). Any discrepancies between reviewers in the quality assessment of the studies were discussed, the studies were re-examined, and an agreement reached. The quality rating was reflected in the data synthesis through the height of bars in the harvest plot

and the impact of quality across studies was tested through a sensitivity meta-analysis excluding studies with a low quality score.

Summary measures

The direction of the association was recorded for summary in the harvest plots. The principal summary measures extracted from the studies were odds ratios, relative risks, hazard ratios and rate ratios. As hazard ratios, odds ratios and rate ratios can be considered analogous (Higgins and Green, 2011) and odds ratios approximate the relative risk when the disease incidence is rare (Davies et al., 1998, Higgins and Green, 2011, Stare and Maucort-Boulch, 2016), these summary measures were combined in both the harvest plots and meta-analyses. Odds ratios may be slightly more extreme than rate ratios depending on how rare the outcome is; in this review, rates were generally low and, given the large heterogeneity across studies, this was not considered to alter the results. Adjusted results were selected over unadjusted results. Where necessary, standard methods were used to calculate the risk ratios and confidence intervals (Higgins and Green, 2011).

Several studies measured socioeconomic status as a continuous variable and/or presented results per unit increase. For these studies the corresponding increase for a high compared to low comparison was estimated based on the literature and this was used to calculate the relevant measure of effect for this change. Where studies provided information on the distribution of the data and the range of the exposure, it was possible to validate this assumption. Where studies presented the relationship between high socioeconomic status compared to low socioeconomic status, the inverse of the results were taken in order to present the results as high compared to low socioeconomic status. One study (Beale et al., 2010), presented results as a cumulative incidence. For this study, the ratio of low socioeconomic status to high socioeconomic status was taken as a proxy for the measure of overall risk. For three studies (Beale et al., 2010, Bemis et al., 2014, Fein et al., 1995), only the total sample size was known. For Bemis et al. (2014), the total number of cases and the incidence per 100,000 population in the high and low socioeconomic groups were the only known total and therefore it was assumed that the denominators in the exposed and unexposed groups were the same in order to calculate the correct measure of effect. For Fein et al. (1995), only the total sample size and the total number of cases were

known, along with the proportion of cases in the two socioeconomic measure groups. It was therefore assumed that the two socioeconomic groups were equal in size to calculate the number of cases in each group and to calculate the correct measure of effect. For Beale et al. (2010), only the total sample size and the proportion of cases in each group was known. The original group sizes, prior to exclusions, were also known therefore it was assumed that the proportion of exclusions in each group was equal and the group sizes recalculated on this basis. The number of cases in each of these new group sizes was calculated using the known proportions in order to calculate the measure of effect. The reformatting of these results was conducted in Stata 13.1 (Statacorp, Texas).

Synthesis of results

Harvest plots were used to display and summarise the results of the studies. Harvest plots allow also for the inclusion of studies which do not provide the quantitative results of analyses (Ogilvie et al., 2008) – such studies would be excluded from statistical analytical approaches such as meta-analysis. In the Harvest Plots, each study was represented by a single bar. The height of each bar was used to demonstrate the quality score of each study, with the lowest bar representing low quality studies and the highest bar representing high quality studies, in order to determine the strength of the evidence. The position of the bars detailed whether the results showed lower risk with lower SES, no association or greater risk with lower SES. Children were defined as participants of studies aged less than 18 years; adults were defined as participants of studies aged 18 years or older. Studies were grouped according to the level at which they measured SES, with area-level SES referring to studies which grouped individuals and assigned SES based on characteristics of the geographical area in which they resided, and individual-level SES referring to studies in which individuals were directly asked to report their SES. The harvest plots classify studies according to three potential categories of results; GI infection risk was lower in more disadvantaged groups; GI infection risk was higher in more disadvantaged groups; no significant association found between GI infection and SES. The latter category of no significant association was determined either where statistical significance could not be verified or where no significant association was demonstrated in the quantitative results presented. Harvest Plots were stratified by age (adults compared to children); level of SES measure (individual compared to

aggregate); type of GI infection measure (population based surveys, GP presentation, hospitalisation and laboratory records); and predominant mode of transmission for studies using laboratory confirmed GI infections (foodborne, person-to-person, waterborne and environmental). Trends identified within the harvest plots were used to inform the subsequent meta-analyses.

Studies which presented statistical summary measures for the association between SES and GI infection, or for which these data could be obtained, were included in the meta-analysis. Meta-analyses were conducted in R (version 3.3.1). Where studies analysed the same cases and could not be considered independent samples or provided multiple estimates, only one estimate was retained in the meta-analysis to avoid double counting of cases. Where results of analyses using multiple SES measures were presented, education level was chosen as the SES measure (where available) as this was the most commonly used measure across all studies. Estimates which adjusted for the potential confounders, such as age and sex, were preferentially chosen for inclusion over univariate results. A potential issue with including the same study multiple times, where the study presented results of independent samples (for example adults and children, or multi-pathogen studies), in random-effects meta-analyses is that within-study variability of the different estimates would be treated as between-study variability. Studies with multiple estimates would therefore have disproportionately high weight in the pooled estimate. To overcome this, fixed-effects meta-analyses were first used to combine estimates from the same study. Fixed-effect meta-analysis assumes that the observed differences in study results are not due to statistical heterogeneity but rather due to chance and are assumed to estimate a common (fixed) effect; any differences are therefore the result of sampling variability (Higgins and Green, 2011).

The combined estimates generated by the fixed-effects meta-analyses were then pooled with the remaining studies using an inverse variance random-effects model which estimates the amount of variation between studies and adjusts the study weightings according to the amount of heterogeneity among the effects (Higgins and Green, 2011). Due to the high levels of heterogeneity between the studies in terms of population, design, GI infection measure, SES measure and measure of effect, and the potential for unexplained heterogeneity, random-effects meta-analysis allows these differences to be considered random (Higgins and Green, 2011) and the

corresponding confidence intervals and therefore interpretation of results will be more conservative. The I^2 statistic was used to assess the level of statistical heterogeneity in the meta-analyses, with values of 30 to 60%, 50 to 90% and 75 to 100% used to denote moderate, substantial and considerable levels of heterogeneity, respectively (Higgins and Green, 2011).

Risk of bias across studies

Small study effects, as an indicator of publication bias, were assessed using a funnel plot. Studies were also summarised by quality rating and the robustness of the meta-analysis was tested through sensitivity analysis on the basis of study quality.

Sensitivity analyses

Potential moderating factors were identified a priori (Rose et al., 2016) and subgroup analyses were performed on this basis, with separate tables created for each subgroup. These subgroups included pathogen type; age; country (based on climate and level of development); GI infection measure used; SES measure used; and level of analysis (individual or aggregate). The Human Development Index (United Nations Development Programme, 2016) was used to classify the countries by relative level of development, and climate zones were assigned based on the Köppen Climate Classification System (Met Office, 2015). The harvest plots detailed the number of studies finding a positive, negative and no association, across the categories of the subgroup.

Random-effects meta-regression (Berkey et al., 1995, Thompson and Sharp, 1999) and subgroup meta-analyses were conducted using the potential moderating factors described above. Sensitivity analyses were performed on the forest plots to see whether the pooled estimates were affected by restricting inclusion to high and medium quality studies only, adjusted estimates only, studies which did not control for or match on SES only and studies which provided results only as opposed to results calculated from raw data provided by studies.

Study 2: Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 Study) (Chapter 5)

In Chapter 5, I present results from a longitudinal analysis of data collected through the IID2 Study to explore the association between socioeconomic status (SES) and

risk of IID in the community in the UK. The methods used in this study chapter are detailed below.

Rationale

A large proportion of the burden of GI infection remains hidden; it is estimated that there are 147 cases in the community for every one case reported to national surveillance (Tam et al., 2011a); many individuals do not present to healthcare as most infections are self-limiting. Additionally, it is unclear whether socioeconomic patterns reported in hospital and laboratory-based surveillance reflect differences in risk of infection or reporting and healthcare interaction (Dunlop et al., 2000). Longitudinal population-based survey data may provide better estimates of differences in risk of infection, particularly in the community, which may not be captured through routine surveillance.

Objectives

This study aimed to explore whether different socioeconomic groups experience different risk of GI infection in the UK, through the secondary analysis of a large prospective population cohort, to improve understanding of the role of SES in IID in the community and to inform policies to reduce health inequalities. This study provides an up-to-date assessment of the association between IID and SES for all ages in the UK.

Information source

The data for this study were taken from the prospective cohort study element of the IID2 Study. This study is described in more detail in the description of data sources section above and in Chapter 2. The study estimated that approximately 25% of people in the UK experience an episode of IID each year and that the age-sex standardised incidence of IID in the community was 274 per 1,000 person-years (Tam et al., 2011a). Although several analyses have been conducted on this dataset, the role of SES in the risk of IID has not been the focus.

Ethics

Ethical approval and informed consent were originally obtained for the main study (07/MRE08/5). This included the provision to use the data for future research.

Approval for this secondary analysis of the fully anonymised datasets was not required (Tam et al., 2011a).

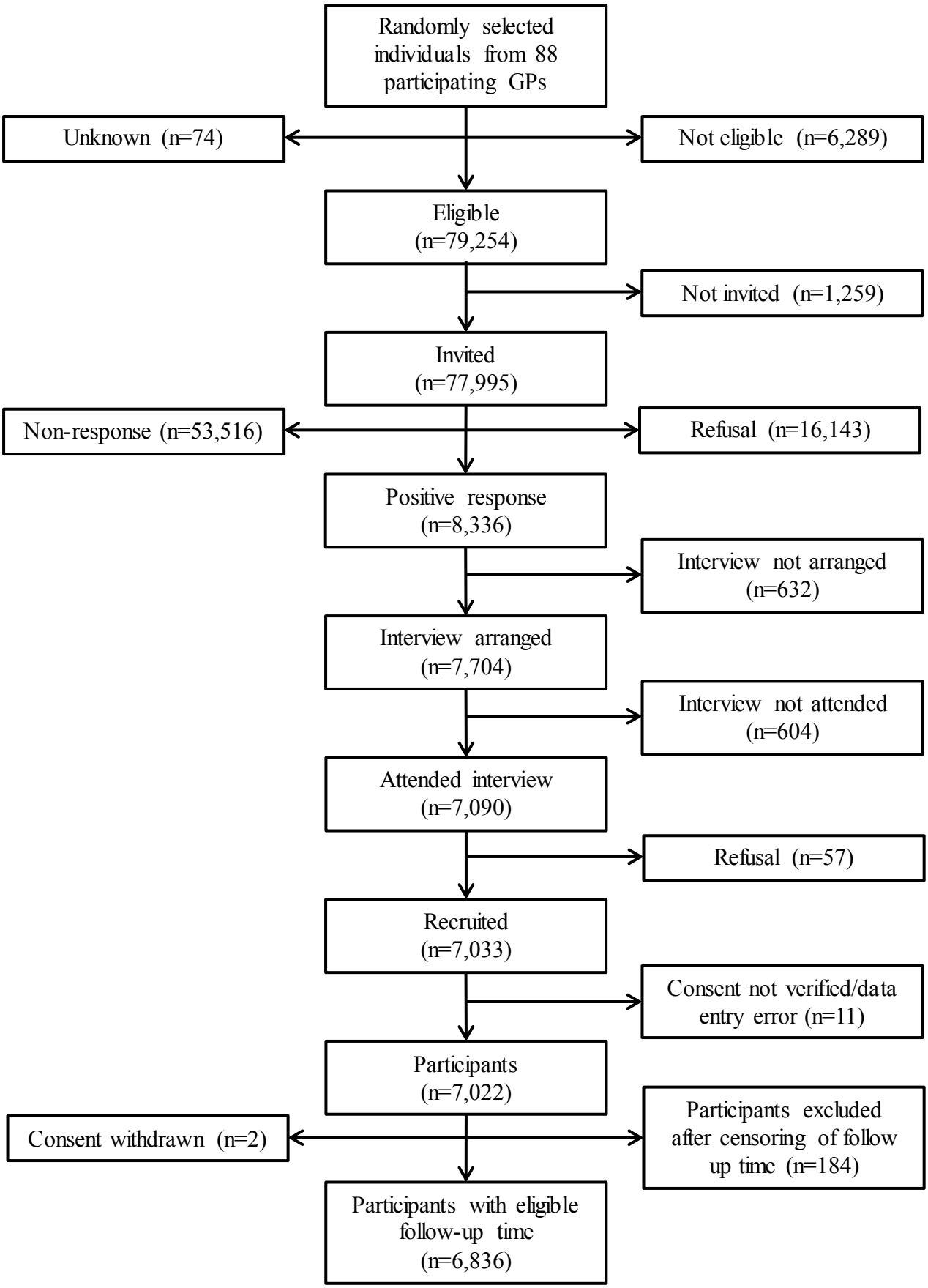
Eligibility/entry criteria

After assessment for eligible follow-up time, a cohort of 6,836 randomly-selected participants was recruited from 88 representative general practices in the UK (Figure 3.4). Sociodemographic information including age, gender and occupation were obtained through a baseline survey upon entry to the cohort and details of IID symptoms were recorded on a weekly basis for up to one year, from October 2007 to August 2009, through the return of an email or postcard indicating whether symptoms of diarrhoea and/or vomiting had been experienced in the previous week. Individuals who reported symptoms completed a more in-depth questionnaire through which details of illness and healthcare contact were recorded.

Infectious intestinal disease was defined as loose stools or clinically significant vomiting (vomiting occurring more than once in 24-hours and if it incapacitated the case or was accompanied by other symptoms such as cramps or fever (Tam et al., 2011a)) lasting less than two weeks, in the absence of a known non-infectious cause, preceded by a symptom-free period of three weeks (Tam et al., 2011a). Cases experiencing illness considered to be travel-related were excluded.

Figure 3.4: Recruitment of participants into the cohort study (Study 2)

Source: Tam et al. (2011a; p.113)



Extraction of data

Follow-up data for all 6,836 participants with eligible follow-up time in the cohort study, including details of demographics and, for definite cases, information on symptoms and healthcare contact were available. These data were held across several Microsoft Excel datasets; therefore data were linked in Microsoft Access to provide one dataset with the relevant socio-demographic, clinical and follow-up data. Small numbers necessitated the recoding of ethnicity into two categories (White and non-White). The NS-SEC variable was recoded in to the 3-class version from the 5-class version in order to provide a hierarchy of socioeconomic status in the analysis.

Individuals were able to re-enter the cohort following a period of three weeks to allow for a period of recovery to ensure any subsequent episodes captured were likely to be the result of a new episode rather than the continuation or recurrence of the same episode. For the purpose of the primary analysis only the first event was retained and subsequent episodes were coded as such for use in robustness testing.

Primary outcomes and covariates

The primary outcome of interest was the development of IID. See case definitions above for further details on how the outcome was defined. The primary exposure of interest was an individual-level measure of SES, self-reported occupation, with each individual assigned an NS-SEC category using the five-class self-coded version (Office for National Statistics). For participants aged less than 16 years, NS-SEC was assigned based on the occupation of the head of the household. Routine and manual occupations were assumed to be broadly equivalent to low SES and professional and managerial occupations to high SES (Office for National Statistics).

A second measure of SES was also available; the Index of Multiple Deprivation (Department for Communities and Local Government, 2011), which was assigned to individuals based on their postcode. This measure of SES was used to conduct robustness testing to assess whether using an area-based measure of SES, instead of an individual measure, modified the results.

Data on the initial age at the start of the follow-up period, sex, ethnicity, type of follow-up (postcard or email) and rural urban classification based on individual postcode were collected for participants. Foreign travel could not be used as a

covariate in this study as this information on foreign travel was collected for cases only.

Statistical methods

Analyses were conducted in Stata 13.1 (Statacorp, Texas). Rates of IID within the study population and by SES were calculated accounting for follow-up time, to produce rates of IID per 1,000 person-years with associated 95% confidence intervals. This was calculated by summing the number of new cases of IID during the follow-up period and dividing this by the total person-time at risk during the follow-up period. Comparisons of incidence rate between exposed (low SES; NS-SEC routine/manual) and unexposed (high SES; NS-SEC professional/managerial) groups were made using incidence rate ratios.

Descriptive statistics were undertaken to understand the differences between the SES groups as well as the sample sizes for each stratified variable. Data on IID episode, age, sex, ethnicity, follow-up type and rural urban classification were considered for differences between SES groups. Associations between the primary exposure of interest (SES) and the explanatory variables were assessed using the Chi Square test.

The main analysis investigated the relationship between SES, as measured by NS-SEC, and time to first IID episode for each participant using Cox proportional hazard regression modelling, with subsequent episodes of IID for an individual being dropped. Univariate relationships between SES and the covariates of interest; rurality and employment status (employed/not working); were explored before fitting a multivariate Cox proportional hazard regression model. Ethnicity and follow-up type were excluded as these were not considered to be confounders although the model was re-run to include ethnicity as a sensitivity analysis. Adjusted hazard ratios and 95% confidence intervals were produced. Interaction terms between the socioeconomic variable NS-SEC and each variable in turn were tested for inclusion to investigate whether the strength of any relationship was moderated by the inclusion of another variable. The baseline hazard was stratified on age group and sex to allow for a direct comparison between individuals of the same age group and sex when assessing the estimate of interest whilst still allowing for the inclusion of age and gender as potential effect modifiers. The assumptions of Cox proportional hazards modelling were tested using log-log plots to test the proportional hazards

assumption for each predictor and a global test to assess whether the model overall met the proportional hazards assumption. A Kaplan-Meier survival curve was constructed showing time to occurrence of first episode of IID by NS-SEC.

Sensitivity analyses

A number of robustness tests were conducted. In this cohort study multiple outcomes were possible, whereby a participant could be studied for the same outcome more than once due to the outcome being reversible and the individual recovering to go on and experience a new, and likely unrelated, outcome later in the follow-up period. This means that the observations may not be truly independent and there may be less variation than anticipated. The analysis was therefore repeated allowing individuals with multiple episodes of IID to re-enter the cohort following a period of censoring (due to symptoms meeting the case definition and requiring a censored period of three weeks after cessation of symptoms; non-response; or symptoms not meeting the case definition), accounting for clustering within individuals by using a robust estimate of variance allowing for inter-person correlation. This was achieved through the use of a frailty model, an extension of the Cox proportional hazards, which is a random effect model for time to event data where the random effect (frailty) has a multiplicative effect on the baseline hazard function (Wienke, 2003).

The analysis was repeated using a less sensitive case definition, whereby individuals reporting symptoms which could not be verified against the case definition (due to a lack of further details about foreign travel or symptom duration) were also included as cases in the analysis. The analysis also was repeated including those unclassifiable within NS-SEC to investigate whether this had an impact on the results. This NS-SEC group comprised individuals for whom it was not possible to classify their occupation or who did not respond to occupation questions.

Additionally, as the NS-SEC data for this group could be classed as missing, Multiple Imputation using Chained Equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was used to statistically simulate NS-SEC categories for these cases. As the outcome for this analysis was binary (IID symptoms experienced or not) this method was the most appropriate as it uses a separate conditional distribution for each imputed variable (UCLA Institute for Digital Research and Education, 2017a). The analysis was then repeated.

The analysis was also stratified by age group to determine whether there were differences in the rate of IID by SES for children, adults and older participants by repeating the analysis on the relevant subset of data by age group. The main analysis was re-run to include ethnicity, as this was excluded from the main analysis. Finally, the analysis was repeated using an area-level measure of SES, the Index of Multiple Deprivation (IMD) (Department for Communities and Local Government, 2011), assigned to each individual based on their postcode.

Study 3: Investigating socioeconomic inequalities in diarrhoea and vomiting in the community accessing telephone-based advice (Chapter 6)

In chapter 6, I present results from an observational analysis of calls to national telephone helplines for health advice (NHS Direct and NHS 111). The methods used in this study chapter are detailed below.

Rationale

National studies of IID in the UK (Tam et al., 2011a) have estimated that a quarter of all individuals in the UK suffer from IID each year but a large proportion of this burden is hidden, with 147 cases in the community for every one case reported to national surveillance (Tam et al., 2011a). It is not clear whether this hidden burden is distributed equally across society. Building on analysis exploring the role of SES in the community seeking no healthcare advice, this analysis explores the role of SES in the community seeking health advice from two national telephone helplines. Crucially, this does not require a caller to attend a healthcare setting which is important if the decision to seek care is related to SES; and allows for a comparison between those not seeking care and those seeking telephone-based advice but not attending a healthcare setting. Furthermore, this is the lowest level of healthcare interaction for which there is a high level of ascertainment.

Objectives

This study aimed to investigate the relationship between socioeconomic status and calls to the national telephone helplines for health advice with symptoms of diarrhoea and vomiting; defined as GI calls; contributing to the understanding of socioeconomic and socio-demographic inequalities of GI infections in the UK.

Information sources

The data for this study were extracted from the HPA/PHE NHS Direct/111 syndromic surveillance systems. These datasets are described in more detail in the description of data sources section above.

Ethics

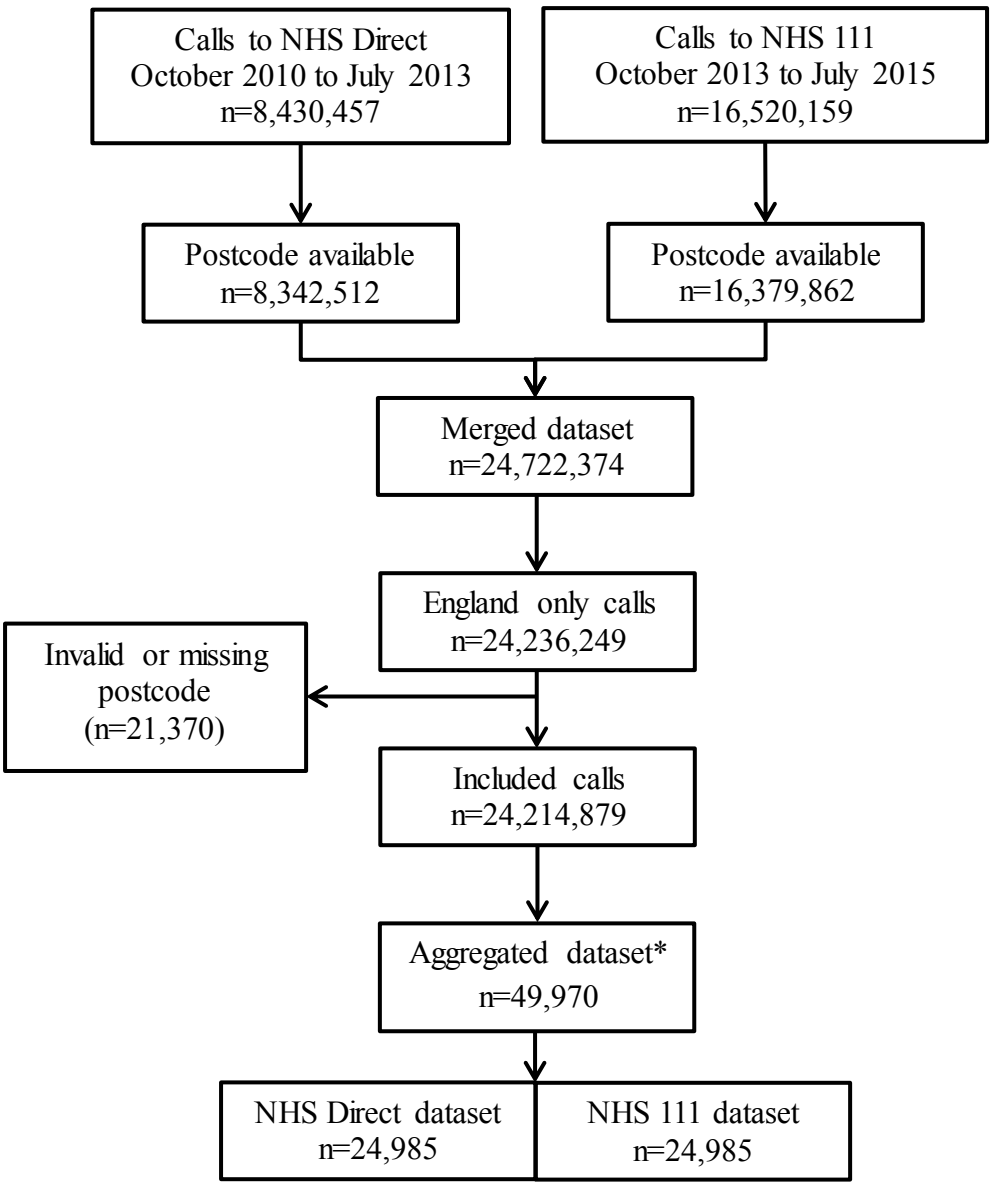
This study falls under the existing HPA (now PHE) permissions under Section 251 of the NHS Act 2006. No identifiable data were used in this study therefore specific ethical approval was not required. The data used were routine syndromic surveillance data collected by HPA/PHE to undertake the real-time syndromic surveillance service; permission to use the syndromic surveillance datasets for this research was obtained from the PHE/NHS 111 steering group.

Eligibility/entry criteria

All calls made to NHS Direct (October 2010 to July 2013) or NHS 111 (October 2013 to July 2015) and recorded in the HPA/PHE NHS Direct/111 syndromic surveillance systems for which a valid English postcode was available were included in this study. Only calls from postcodes in England were selected in order to ensure consistency between the two systems as NHS 111 is a service for England only. Figure 3.5 describes the flow of participants included in this study. Due to the changeover between systems, no data were extracted between August 2013 and September 2013 to allow for potential drop-off and uptake of reporting across the two systems.

Data were aggregated at call-level; by number of calls about diarrhoea, vomiting or other non-diarrhoea or vomiting symptoms for the given date, postcode district, age and gender.

Figure 3.5: Selection of participants to national telephone helpline for health advice study (Study 3)



*Aggregated to postcode district, age and gender

Extraction of data

Data were extracted from the NHS Direct and NHS 111 systems, which are held securely by PHE, into two separate Microsoft Excel spreadsheets. These were combined and cleaned in R (version 3.3.1).

Primary outcomes and covariates

The primary outcome of interest was calls for which symptoms of diarrhoea and/or vomiting was recorded as the reason for the call; these outcomes were used as proxies for acute gastroenteritis (GI-calls). Calls for reasons other than diarrhoea/vomiting were coded as non GI-calls and used as a comparator group.

For this analysis, only postcode district, the first part of a postcode, was available. The primary exposure of interest was SES which was determined using a small area deprivation measure, the Index of Multiple Deprivation 2010 (IMD) (Department for Communities and Local Government, 2011). In order to link the IMD to each postcode district, it was necessary to create population-weighted IMD scores to account for the potential for multiple IMD scores to relate to a single postcode district due to differing geographical areas. This was then divided into population-level quintiles, with the first quintile representing the least disadvantaged and the fifth quintile representing the most disadvantaged; see description of data sources section above which provides further details on how the exposure was derived and defined.

Other covariates of interest included age (coded as age groups 0-4, 5-9, 10-15, 16-19, 20-59 and 60 years and over); sex (male/female); date of call (used to assign calls to pre- and post-introduction of NHS111 to assess for any shifts in the socio-demographic composition of callers based on differences between the two systems); and proportion of the population within the postcode district classed as urban (operationalised as deciles) to account for potential differences in call volume from rural compared to urban areas were included as covariates in this study in order to understand the association between GI infection and SES by these variables and to assess whether any of them confounded or mediated the relationship. Population by postcode district, age and sex were merged with the call data in order to account for the underlying population. Due to no individuals of a specific age group and gender being resident in some postcode districts, postcode districts with a population of zero (n=1,357, 0.1%) were dropped from the analysis.

Statistical methods

Analyses were conducted in R (version 3.3.1). An observational study design was used to assess the relationship between SES and GI-calls to telephone-based

healthcare advice systems in England. Firstly, descriptive analysis to assess the distribution of all calls by IMD quintile and each covariate in turn was undertaken. In order to take account of the underlying population at risk (including those who did not call NHS 111 or NHS Direct), rates per 10,000 person-months were calculated for GI-calls and non GI-calls. Age- and sex-adjusted incidence rates were also calculated for both GI- and non GI-calls separately for comparison.

The main analysis explored the relationship between SES, as measured by IMD quintiles, and GI-calls using a GLM with a Poisson family and log-link function with the log population in each postcode district, age group and gender as an offset. The Poisson family was selected because the outcome is relatively rare. Poisson is a form of regression which allows count data to be modelled; in this model it was used to model rate; i.e. the count of events divided by the population. The event (calls) was an integer. Calls were thought to be independent and the average frequency for the time period of the study was known as it was possible to count how many calls had been made and the underlying population was accounted for. The log-link function is used for variables which follow the Poisson distribution and exponentiates the linear predictors. The log population was used as an offset because the underlying population may have had a proportional effect; that is, to account for the fact that areas with a higher population may have a higher call volume.

Univariate relationships were first explored between the outcome of interest, GI-calls, and each of the covariates of interest; IMD; age; sex; and urban decile; before fitting a multivariable model. All variables were retained in the final model as potential confounders or mediators. Adjusted risk ratios and 95% confidence intervals were produced. Interactions terms between age and sex, and age and IMD Quintile were included.

Sensitivity analyses

Several sensitivity analyses were undertaken. Firstly, to explore whether there was a linear relationship between GI calls and SES or proportion of the population classed as urban, the main model was re-run with raw IMD Score and urban proportion included as continuous variables. Secondly, in the main model, postcode districts with a population of zero were dropped from the analysis. To assess whether this impacted on the results, these populations were recoded to one in order to be

included in the analysis; the model was then re-run. Due to a change in protocol for NHS Direct in November 2011 which resulted in symptom data in syndromic surveillance being unavailable for calls regarding infants under 1, the analyses were repeated excluding calls regarding infants under 1 to explore whether this impacted on the results.

Finally, as the rotavirus vaccine was introduced in July 2013 and NHS 111 was introduced in October 2013, to assess whether the shift from NHS Direct to NHS 111 or the introduction of the rotavirus vaccine had any impact on GI-calls, a vaccine eligible cohort was assumed to be 0-4 year olds and incidence rate ratios and 95% confidence intervals were produced, by year, for this age group and for those aged 5 years old and over for both GI- and non GI-calls in order to compare trends.

Study 4: Social patterning of risk factors and development of severe complications in a diagnosed GI infection (Chapter 7)

In chapter 7, I present results from a two-part study; a descriptive analysis of the social patterning of risk factors for infection with a specific and severe GI pathogen, (STEC); and an analysis of the socio-demographic risk factors in development of haemolytic uraemic syndrome (HUS), a severe complication of STEC infection, among paediatric STEC cases.

In this section I will describe the methods related to the descriptive analysis of the social patterning of risk factors for STEC.

Rationale

In comparison to other GI infections, infection with STEC is rare but potentially very serious. Risk factors for STEC infection have been well-documented in the literature (Gillespie et al., 2005, Locking et al., 2001, Parry et al., 1998). Despite differences in overall risk of STEC by age and gender (Adams et al., 2016b, Byrne et al., 2015, Lynn et al., 2005, Vally et al., 2012) it is not clear how risk and exposure to these risk factors vary by socio-demographic characteristics. In contrast with the previous studies in this thesis, this study uses a laboratory confirmed infection with in-depth follow-up of cases to assess socio-demographic patterning of risk factors for STEC.

Objectives

This study aimed to explore whether there are differences in clinical symptoms, healthcare contact or risk factors for STEC by SES, to assess the direction of any associations identified and suggest hypotheses for testing in future studies.

Information source

The data for this study were taken from NESSS. This dataset is described in more detail in the description of data sources section above.

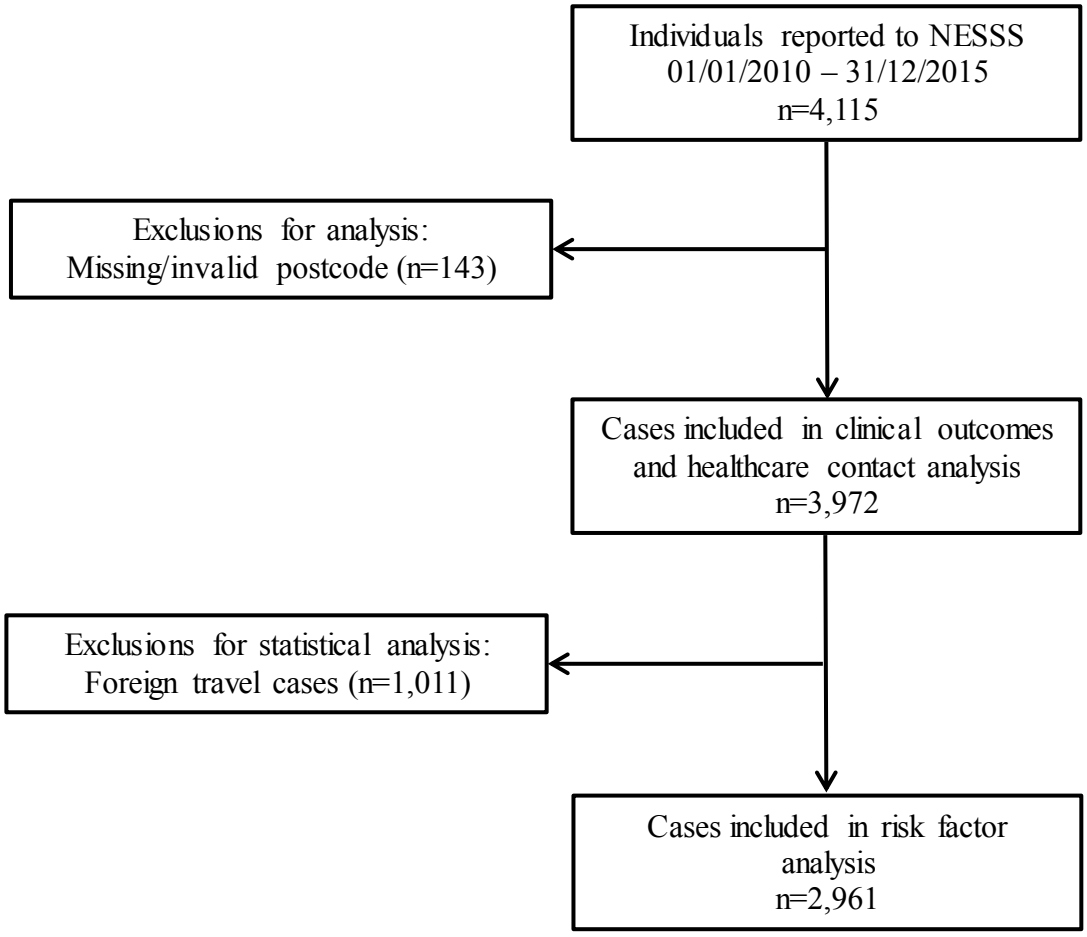
Ethics

This study falls under the existing Health Protection Agency (HPA, now PHE) permissions under Section 251 of the NHS Act 2006. In addition, a favourable ethical opinion was received from the South East Coast - Surrey Research Ethics Committee (15/LO/2138) on 1st December 2015 covering the use of this dataset (Appendix 1.1).

Eligibility/entry criteria

For inclusion in this study, cases of STEC were defined as all primary, symptomatic Shiga-toxin (*stx*) positive individuals recorded in NESSS between 1st January 2010 and 31st December 2015 (inclusive). Individuals recorded as secondary or asymptomatic, those who were only microbiologically suspect STEC cases and those subsequently identified as *stx*-negative were excluded as follow-up information is not routinely obtained for these groups. Individuals for whom no questionnaire was available were also excluded. Both sporadic and outbreak-associated cases were included. Cases believed to have contracted STEC infection through foreign travel were included in models assessing the relationship between SES and clinical symptoms or healthcare contact but were subsequently excluded from analyses assessing the relationship between SES and risk factors. Figure 3.6 describes the flow of participants included in this study.

Figure 3.6: Selection of participants to STEC risk factor study



NESSS – National Enhanced Surveillance System for STEC

Extraction of data

Data were extracted from NESSS using the inclusion criteria described above. Data from NESSS are held in a secure web-based system, which is cleaned and stored in a Microsoft Access database. Demographic, clinical, microbiological and exposure data were extracted. Small numbers necessitated the recoding of ethnicity into two categories (White and non-White).

Primary outcomes and covariates

The primary exposure of interest was SES. This was determined using a small area deprivation measure, the Index of Multiple Deprivation 2010 (IMD) (Department for Communities and Local Government, 2011), assigned to each individual based on their full postcode as a proxy for SES and divided into population-level quintiles,

with the first quintile representing the least disadvantaged and the fifth quintile representing the most disadvantaged; see description of data sources section above which provides further details on how the exposure was derived and defined.

Other covariates of interest included in the analysis were age group (coded as <1, 1-4, 5-9, 10-15, 16-19, 20-59 and 60 years and over); sex (male/female); ethnicity (White/non-White); rurality (rural/urban); whether the case was associated with an outbreak (yes/no); clinical symptoms (diarrhoea, bloody diarrhoea, nausea, vomiting, abdominal pain, fever; operationalised as yes/no); healthcare contact (yes/no); antibiotic and anti-diarrhoeal use (yes/no); microbiology (s*tx*); whether the case reported foreign travel (yes/no) and whether the case reported exposure to a range of foodborne, waterborne or environmental risk factors. The presence of s*tx* was used as the main microbiological variable in order to assess the relative role of s*tx* genes in the development of HUS.

Missing data

Where symptoms, travel status, healthcare contact or exposure variables were blank or unknown, these were recoded as a negative response. This is a reasonable assumption based on known questionnaire completion practices. There were no other missing data.

Statistical methods

Analyses were conducted in Stata 13.1 (Statacorp, Texas). An observational study design was used to assess the relationship between SES and a variety of clinical presentation variables and risk factors for STEC infection in England. This was first assessed through a descriptive analysis of the patterning of risk factors by IMD quintiles. The distribution of the population by IMD quintile overall and by age and sex were then used to produce crude incidence rates accounting for the population 'at risk' to explore whether the socio-demographic patterns identified in the descriptive analysis could be explained as following the expected population-level distribution.

As with Study 2, Multiple Imputation using Chained Equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was used to statistically simulate ethnicity categories where ethnicity was missing. As the outcomes for this analysis were binary (experiencing clinical symptoms, contact with healthcare and

exposure to risk factors or not) this method was the most appropriate as it uses a separate conditional distribution for each imputed variable (UCLA Institute for Digital Research and Education, 2017a). As the level of missingness was relatively high, and there were a limited number of variables and cases, fifty imputed datasets were generated. The MAR assumption was tested by assessing the distribution of the missing ethnicity variable by age and sex.

A series of univariate and multivariable logistic regression models were estimated to assess the relationship between SES and each outcome variable in turn, which were selected based on whether each had a statistically significant relationship with IMD in the descriptive analysis ($p=<0.05$). Firstly, univariate logistic regression models were fitted with IMD quintile and each clinical (outbreak-associated, diarrhoea, vomiting) and healthcare contact (GP, A&E, hospitalisation) variable in turn. Multivariable logistic regression models were then fitted with IMD quintile and each of these variables, adjusting for potentially confounding variables (age group, sex, ethnicity, rurality and s*tx* gene which was used as a marker of severity). The univariate and multivariable relationship between SES and foreign travel was estimated in the same way. Cases thought to have acquired their infection through foreign travel were then excluded from further analysis as risk factor data for these cases are not routinely collected.

Univariate and multivariable associations between SES and each of the known risk factors for STEC were then estimated based on those which had a statistically significant relationship with SES in the descriptive analysis. Firstly, univariate logistic regression models were fitted with IMD quintile and each of the risk factor variables in turn before fitting a multivariable model to adjust for potentially confounding variables (age, sex, ethnicity, rurality). As s*tx* genes are markers of severity, this was not included as a potential confounder in this model as this does not feature on the causal pathway between SES and exposure to risk factors for STEC.

Sensitivity analyses

Multiple sensitivity analyses were performed to assess the robustness of the main findings. The analysis was repeated excluding outbreak-related cases in order to assess the relationship between SES and clinical/healthcare contact and risk factors

sporadic infections. The analysis was also repeated for cases aged less than 16 years to determine whether there were differences in clinical/healthcare contact or risk factors by SES for children and for use as supplementary material for the subsequent analysis of paediatric HUS cases following STEC infection.

In the next section I will describe the methods related to the analysis of socio-demographic risk factors in development of haemolytic uraemic syndrome (HUS) among paediatric STEC cases.

Rationale

Despite being is a relatively rare but potentially very serious complication which can occur following infection with STEC, HUS is recognised as the most common cause of acute renal failure in children in the UK (Lynn et al., 2005). There is evidence to suggest that the consequences of GI infections are greater for those who are more disadvantaged (Olowokure et al., 1999, Pockett et al., 2011, Rose et al., 2017), but it is not known whether the burden of HUS in children is equally distributed across society; previous studies have suggested that risk of HUS is greater amongst the least disadvantaged (Rowe et al., 1991, Whitney et al., 2015) but this has never been explored in England.

Objectives

This study aimed to investigate the relationship between childhood socio-economic circumstances (SEC), STEC infection and subsequent development of HUS in a paediatric population in England.

Information sources

The data for this study were taken from the National Enhanced Surveillance System for STEC (NESSS) and the British Paediatric Surveillance Unit (BPSU) Haemolytic Uraemic Syndrome (HUS) Study. These datasets are described in more detail in the description of data sources section above.

Ethics

Ethical approval was obtained for the original study (Ref: 11/LO/1412). As of October 2010, HUS is a statutory reportable condition and this study falls under the existing HPA (now PHE) permissions under Section 251 of the NHS Act 2006. In

addition, a favourable ethical opinion was received from the South East Coast - Surrey Research Ethics Committee (15/LO/2138) on 1st December 2015 covering the use of this dataset (Appendix 1.1).

Eligibility/entry criteria

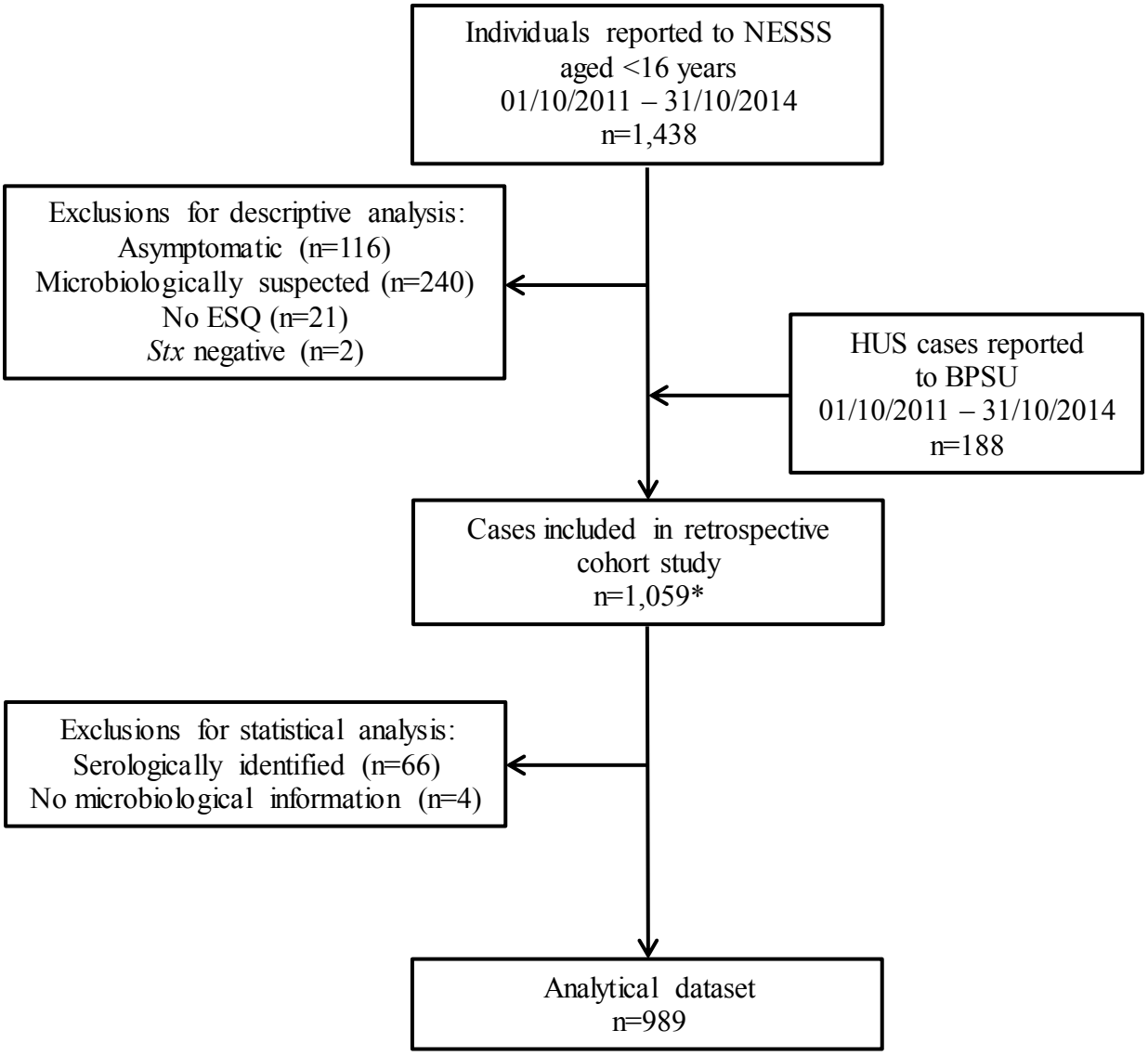
Paediatric HUS cases, conforming to a set of clinical criteria (Byrne, 2017) and the case definition detailed previously, were identified as individuals aged less than 16 years recorded in the BPSU HUS Study which ran from 1st October 2011 to 31st October 2014, and/or through being recorded as a HUS case in the ESQ and therefore in NESSS.

Paediatric STEC cases were identified as individuals aged less than 16 years recorded in NESSS who reported experiencing symptoms consistent with STEC infection between 1st October 2011 and 31st October 2014. Unlike the previous analysis whereby secondary cases were excluded due to the absence of risk factor information, in this analysis secondary cases were included as it is possible for secondary STEC cases to develop HUS and the socio-demographic characteristics of these individuals were known. Cases which were part of known outbreaks were also included as exclusion on this basis may have reduced the ascertainment of HUS. Cases identified in the absence of microbiological confirmation of STEC were included as it is not always possible to obtain a stool sample in HUS cases and excluding these individuals may have led to an under ascertainment of HUS.

Individuals recorded as asymptomatic, those who were only microbiologically suspect STEC cases and those identified as s*tx*-negative were excluded as follow-up information is not routinely obtained for these groups. Individuals for whom no questionnaire was available were also excluded as details of illness for these individuals and their HUS status could be established. Figure 3.7 describes the flow of participants included in the cohort. Cases of HUS were matched to their record in NESSS to create a defined cohort of cases of HUS thought to have developed as a result of STEC infection (known as diarrhoeal, or typical, HUS).

For the purposes of the main analysis, individuals with a serological result and those for whom there was no microbiological confirmation of STEC were excluded in order to assess the role of s*tx* genes specifically.

Figure 3.7: Selection of participants to HUS Cohort Study (Study 4)



*An additional 19 HUS cases not reported to BPSU were identified in NESSS; NESSS – National Enhanced Surveillance System for STEC; HUS – haemolytic uraemic syndrome; BPSU – British Paediatric Surveillance Unit; ESQ – enhanced surveillance questionnaire

Extraction of data

Data were extracted from NESSS and the BPSU HUS Study using the inclusion criteria described above. Both datasets were held in separate secure Microsoft Access databases and cases in the BPSU HUS Study dataset were linked using NHS Number to those in the NESSS dataset to create a retrospective cohort of STEC cases. Small numbers necessitated the recoding of ethnicity into two categories (White and non-White).

Primary outcomes and covariates

The primary outcome of interest was the development of HUS. See description of data sources section above for HUS case definition which provides further details on how the outcome was defined. The primary exposure of interest was an area-level measure of childhood socioeconomic conditions (SEC), the Index of Multiple Deprivation 2010 (IMD) (Department for Communities and Local Government, 2011). This measure was used as a proxy for childhood SEC. The IMD score was assigned to each individual based on their full postcode and divided into population-level quintiles, with the first quintile representing the least disadvantaged and the fifth quintile representing the most disadvantaged.

Other covariates of interest included in the analysis were age group (coded as <1, 1-4, 5-9 and 10-15 years); sex (male/female); ethnicity (White/non-White); travel (yes/no); rurality (rural/urban); microbiology (s*tx*); antibiotic use (yes/no); and clinical symptoms (diarrhoea, bloody diarrhoea, nausea, vomiting, abdominal pain, fever). The presence of Shiga-toxin (s*tx*) was used as the main microbiological variable in order to assess the relative role of s*tx* genes in the development of HUS.

Missing data

Where symptoms, travel status and healthcare contact variables were blank or unknown, these were recoded as a negative response. This is a reasonable assumption based on known questionnaire completion practices. The s*tx* gene was unknown for four individuals but, as previously mentioned, these cases were excluded from the analysis to assess the role of s*tx* genes. There were no other missing data.

Statistical methods

Analyses were conducted in Stata 13.1 (Statacorp, Texas). Descriptive statistics were undertaken to understand the differences in distribution by SEC, as well as the sample sizes, for each stratified variable. Data on HUS development, age group, sex, ethnicity, travel, rurality, region, s*tx* gene, symptoms and healthcare contact were considered for differences by SEC. Associations between the primary exposure of interest (SEC) and the explanatory variables were assessed using the Chi Square test.

The main analysis explored the relationship between SEC, measured by IMD quintiles, and progression to HUS among paediatric STEC cases using binary logistic regression with multiple imputation. Univariate relationships between HUS and the covariates of interest; IMD; age; sex; ethnicity; travel; rurality; region; *stx* gene; symptoms; and healthcare contact; were explored before fitting a multivariable logistic regression model. This model is appropriate as the outcome of interest is binary. All variables were retained in the final model as potential confounders or mediators. Adjusted odds ratios and 95% confidence intervals were produced. Interaction terms between the socioeconomic variable IMD and each variable in turn were tested for inclusion to investigate whether the strength of any relationship was moderated by the inclusion of another variable.

As with Study 2 and the earlier risk factor analysis, Multiple Imputation using Chained Equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was used impute ethnicity. Again, as the outcome for this analysis was binary (development of HUS or not) this method was the most appropriate. As the level of missingness was relatively high, and there were a limited number of variables and cases, fifty imputed datasets were generated. The MAR assumption was tested by assessing the distribution of the missing ethnicity variable by age, sex and region.

Sensitivity analyses

A number of sensitivity analyses were undertaken to explore the robustness of the results. Fractional polynomial prediction plots for age and sex by HUS status were produced to explore the relationship between these variables as it was considered to be a non-linear relationship. Fractional polynomial prediction plots can detect the best functional form of variables.

The analysis was repeated excluding cases that were thought to have acquired their STEC infection through travel outside of the UK to determine if there were differences in progression to HUS by deprivation amongst children who did and did not travel.

Due to the proportion of missing data for ethnicity and the use of multiple imputation to account for this, the analysis was repeated excluding the ethnicity variable to explore whether the inclusion of the multiply imputed ethnicity variable modified the

relationship between SEC and development of HUS. To assess for potential collinearity between IMD and ethnicity a post-hoc matched analysis on ethnicity was conducted. To achieve this, HUS and non-HUS cases were separated into two datasets before being merged using a 1:1 match based on ethnicity. Conditional multivariable logistic regression was then used and the results compared to that of the main model. This approach removes potential confounding by ethnicity through matching on this variable and therefore allows for the assessment of the role of SEC independently of ethnicity without the need for dropping this variable. To assess collinearity using a different approach, penalised logistic regression on the multiply imputed dataset was also conducted. Penalised logistic regression applies a penalty to the covariates to control for potential collinearity. The output from this model was also compared with that of the main model.

3.6 Summary

The methods used within the four empirical studies comprising this thesis have been selected in order to answer the overarching question of whether there are socioeconomic inequalities in risk of and exposure to GI infections in the UK, as well as the specific research objectives of each chapter. Multiple data sources and sensitivity analysis have been used to cross-validate and assess the robustness of the findings wherever possible and together provide a clearer picture of the role of SES in risk of and exposure to GI infections. The results of these studies are presented in the chapters that follow.

Chapter 4 – Study 1 (Objective 1)
Relationship between socioeconomic status and risk of gastrointestinal infections in high income countries: A systematic review and meta-analysis

Abstract

Background

The association between socioeconomic status (SES) and ill-health is well-documented but limited up-to-date evidence on the relationship between SES and risk of gastrointestinal (GI) infections exists, with individual published studies producing apparently conflicting results. A systematic review was needed to clarify these discrepancies in the literature.

Review question

What is the extent and nature of the association between SES and risk of GI infection in high income countries?

Methods

Three systematic methods were used to identify relevant literature published from 1980 until October 2015. Firstly, searches were conducted using three electronic databases: MEDLINE, Scopus and Web of Science Core Collection. Secondly, reference lists of identified studies were interrogated for additional relevant literature. Finally the grey literature was searched. Screening of the results was performed by two reviewers using pre-defined inclusion and exclusion criteria. The reference lists of included studies were then searched and Google was used to search for grey literature.

Observational studies were included if they were conducted between 1980 and October 2015 and reported a quantitative association between symptomatic GI infections and SES in a representative population sample from a member-country of the Organisation for Economic Co-operation and Development (OECD).

Data were extracted using a standardised, validated form. Study quality was assessed using the Liverpool University Quality Assessment Tools (LQAT). Harvest plots were created for comparison stratified by age; level of SES variable; GI ascertainment method; and predominant mode of transmission. Meta-analysis was performed on a subset of data for which quantitative results were available.

Results

In total, 6021 studies were identified; 102 met the inclusion criteria. Overall risk of GI infection for low versus high SES was 1.06 (95% CI 0.95–1.19). For children, risk of GI infection was higher for those of low SES versus high (RR 1.51, 95% CI 1.26–1.83), but there was no association for adults (RR 0.83, 95% CI 0.61–1.14).

Limitations

There was a large amount of unexplained heterogeneity across studies, which is likely due to the variations in the primary aims and the measures used to assess the association between SES and risk of GI infection as well as the variables used to provide adjusted estimates. Non-English language studies were excluded and countries in transition between low/middle and high income were included, which could limit interpretation of the results.

Conclusions and implications of key findings

Disadvantaged children, but not adults, have greater risk of GI infection compared to their more advantaged counterparts. Increased risk may relate to different exposures, vulnerability or healthcare-seeking behaviours across socioeconomic groups. Gaining further insight into this relationship will help inform policies to reduce inequalities in GI illness in children.

Systematic review registration number

The systematic review was registered on PROSPERO: CRD42015027231.

4.1 Introduction

This study was designed to address gaps in the literature identified in the literature review (Chapter 2). The overall objectives of the research within this thesis have been detailed in Chapter 1. This chapter addresses Objective 1: To conduct a systematic review of existing evidence of socioeconomic inequalities in risk of GI infections in high income countries, by answering the research question: 'For individuals from high income countries, is there evidence that lower compared to higher SES is associated with higher or lower incidence/prevalence of GI infection?'.

Gastrointestinal infections, caused by organisms such as bacteria, viruses or protozoa (Adak et al., 2002), are common, leading to diarrhoea and vomiting as well as other more serious health problems, such as haemolytic uraemic syndrome (HUS) (Byrne et al., 2015), Guillain-Barré syndrome (McCarthy and Giesecke, 2001), irritable bowel syndrome (Neal et al., 2002) and reactive arthritis (Dworkin et al., 2001); and can result in interference with normal day-to-day activities. Published estimates suggest that around one in four people in the UK suffer an episode of IID per year, and foodborne illness (a proportion of IID) in England and Wales results in costs of around £1.5 billion per year to the NHS, the economy and individuals (Tam et al., 2011a). Whilst associations between low SES and higher risk of a variety of other diseases, such as cancer and coronary heart disease (Graham, 2009), human immunodeficiency virus and tuberculosis (Biering-Sørensen et al., 2012, Hughes and Gorton, 2015, Semenza, 2010) have been established in the literature, there is no conclusive evidence to indicate whether the burden of GI infection is felt equally across the whole socioeconomic spectrum of society or whether certain population groups are at greater risk.

As outlined in Chapter 2 (Literature Review), multiple risk factors for GI infection have been investigated in the literature, including environmental risk factors such as population density and rurality, and individual-level risk factors such as sex, ethnicity, host-factors and foodborne, waterborne and environmental exposures (Bessell et al., 2010, Davis et al., 2013, Tam et al., 2013, Zappe Pasturel et al., 2013). The youngest age groups and the elderly are most at risk (Majowicz et al., 2007, Olowokure et al., 1999, Tam et al., 2013). As established in Chapter 2 (Literature Review), the role of SES in the risk of GI infection is far less clear and inconsistent results have been observed in the published literature, with some studies reporting

higher risk of GI infections among lower socioeconomic groups (Biering-Sørensen et al., 2012, Olowokure et al., 1999, Pockett et al., 2011) and others observing lower risk of GI infections among lower socioeconomic groups (de Wit et al., 2001a, Pollard et al., 2014). This was echoed in a previously published systematic review exploring the impact of SES on laboratory confirmed foodborne illness in high income countries, which identified 16 studies across four pathogens with mixed results, differing by pathogen (Newman et al., 2015). That review was limited to foodborne pathogens and did not include syndromic definitions of GI infections, which are important if the pathway to laboratory diagnosis is thought to vary by SES. Due to the small number of studies identified by the (Newman et al., 2015) review, the authors were unable to conduct a meta-analysis to statistically synthesise their findings.

Given the conflicting results of the systematic review conducted by Newman et al. (2015) and as emphasised in the illustrative literature review conducted in Chapter 2, a systematic review was warranted to clarify these contradictory results using a statistical approach. The systematic review reported here aimed to explore the relationship between SES and a full range of GI infections, including those defined syndromically, to assess the magnitude and direction of the association, and explore some of the possible explanations for any observed differences across studies.

Specifically, this study explores whether different socioeconomic groups experience different risk of GI infection, through a systematic review and meta-analysis of the literature, with the aims to improve understanding of the role of SES in risk of GI infection; to generate hypotheses for testing both within and out-with this thesis; and inform policies to reduce health inequalities. It provides the first comprehensive assessment of the association between SES and risk of GI infection for a range of GI infections including those defined syndromically, stratified by age, in high income countries. In this chapter I will present the results from this systematic review and meta-analysis, including details of sensitivity and robustness analyses, as well as the implications of the findings for future work, and the public health implications of this improved understanding of the relationship between SES and GI infection.

4.2 Methods

The methods adopted in this study are reported in detail in Chapter 3. An overview is provided below.

Protocol and registration

The systematic review was registered in the PROSPERO International prospective register of systematic reviews (PROSPERO 2015:CRD42015027231 (Adams et al., 2015); Appendix 2.1). The published protocol for this systematic review can be found in Appendix 2.2 (Rose et al., 2016).

Eligibility criteria

Full details of the inclusion and exclusion criteria were reported in Chapter 3, Table 3.6. Studies reporting on the incidence or prevalence of any symptomatic GI infection were included. Studies using syndromic definitions of GI infections without a laboratory diagnosis were also included under the hypothesis that various socioeconomic factors might influence whether an individual is diagnosed with a GI infection. Studies from high income countries, as defined by membership of the Organisation for Economic Co-operation and Development (OECD), conducted after 1980 and published up to and including the search date (13th October 2015) were included. The focus only on high income countries was to ensure that selected studies were as comparable as possible to the UK population and to present day socioeconomic conditions. Studies that referred to the same sample of individuals were included if they analysed different exposures or outcomes. Where more than one study analysed the same sample of individuals using the same outcomes and exposures, only one study was included, namely the study with the greatest focus and amount of information on the relationship between SES and GI infections.

Information sources and search strategies

Firstly, electronic searching of MEDLINE (Ovid), Scopus and Web of Science Core Collection was performed. Search terms were piloted prior to selection and comprised specific GI infections and symptom-based terms, socioeconomic and inequality terms and high income countries of interest. These search terms are detailed for each of the three databases in Appendix 2.3.

Secondly, the reference lists of studies selected for inclusion in the review were screened to identify relevant articles that were not captured via electronic searching. The grey literature was searched using the Google internet search engine and Google Scholar search application. The first 100 results from each search were screened for inclusion. References of these results which were selected for inclusion were also reviewed. Finally, experts in the field of GI infections were consulted to identify any extra studies which also met the date parameters. Search results were exported to Endnote X8.

Study selection

Titles and abstracts of all the publications were screened independently by two reviewers to ensure consistency in the application of inclusion and exclusion criteria. Discrepancies were discussed and resolved through consensus between the two reviewers. The full text for studies deemed relevant after title and abstract screening, were sought and screened in the same way. Inter-library loans and direct contact with authors were used where full texts for studies were not available.

Studies which reported quantitative results or reported results in-text but without results tables or data were selected for inclusion in the synthesis using Harvest Plots. For studies where quantitative data were reported in text form only, authors were contacted to obtain the relevant data. Studies included in the systematic review are detailed in Appendix 2.4. Where it was not possible to obtain quantitative results, either from the full-text or through contact with authors, these studies were excluded from the meta-analyses (Appendix 2.4). Studies which did not meet the inclusion criteria following full text review are detailed in Appendix 2.5 with the rationale for exclusion.

Data collection process and data items

Data were extracted into a standardised Excel spreadsheet one reviewer and were checked independently by the second reviewer. The data extraction form was piloted prior to use in May 2015. Data items extracted included: aim/hypothesis; study design; level of analysis; country; sample size; age; measurement of GI infection; measurement of SES; covariates and results (Appendix 2.6).

Where results were reported in-text but without quantitative results, or where clarifications on data items were required, study authors were contacted and asked to provide further details.

Where studies did not present quantitative results and it was not possible to obtain further details from authors, these results were recorded as non-significant findings but retained for inclusion in the harvest plots. Adjusted results were recorded where provided, in addition to unadjusted results. Results from univariate and multivariate analyses were recorded and noted. Where more than one study design was reported to have been used within a study, the design which yielded the results was recorded. Where multiple SES and GI measures were used within a single study, all were recorded.

Risk of bias in individual studies

Risk of bias and a quality assessment of each included study were conducted by two reviewers independently, outcomes were compared and, in case of discrepancies, agreement was reached through discussion. The Liverpool University Quality Assessment Tool (LQAT) was the adopted instrument. Studies were classified as low, medium or high quality depending on the star rating achieved. The quality rating was reflected in the data synthesis through the height of bars in the harvest plot and through sensitivity meta-analysis.

Summary measures extracted from studies

The direction of the association between SES and GI infection was summarised in the harvest plots. The principal summary measures extracted from the studies were odds ratios, relative risks, hazard ratios and rate ratios which were combined as risk ratios in meta-analyses and meta-regression.

Synthesis of results

Both harvest plots and meta-analysis were used to synthesise the data. Harvest plots were used to demonstrate the association between risk of GI infection and SES, by displaying and summarising the results of the included studies and the subgrouping graphically (Ogilvie et al., 2008). Factors identified as potentially important in the literature review (Chapter 2) were used to guide the subgroup analysis; this included age, predominant mode of transmission, measure of SES used and GI ascertainment

method. An inclusive strategy was used for the harvest plots, allowing all included studies to be captured graphically, irrespective of whether quantitative data were provided. Each reported association between SES and GI infection was represented by a single bar. The process for assigning quality is described in detail in Chapter 3. The height of each bar was used to indicate low, medium or high quality studies to visually determine the strength of evidence, and greater weight given to conclusions drawn from the most methodologically robust and reliable studies. The findings from the harvest plots were used to inform the methods used in the meta-analysis and to identify potential explanations for the contrasting findings observed in the literature.

Meta-analyses were conducted in R (version 3.3.1) using an inverse variance random-effects model on combined results. Where necessary, standard methods were used to calculate the risk ratios and confidence intervals (Higgins and Green, 2011). Where studies analysed the same cases, or provided numerous estimates for the relationship between SES and GI infection, only one estimate was retained in the meta-analysis to avoid the double counting of cases Eleven studies provided more than one estimate but the cases used for each estimate were considered independent of each other, so all estimates were included in the meta-analysis. Fixed-effect meta-analyses were used to combine estimates from the same study, allowing these pooled estimates to be combined with the remaining studies then using random-effects meta-analysis (Higgins and Green, 2011).

Statistical heterogeneity was assessed by applying the I^2 statistic with values of 30 to 60%, 50 to 90% and 75 to 100% used to denote moderate, substantial and considerable levels of heterogeneity, respectively (Higgins and Green, 2011).

Sensitivity analyses

A funnel plot was used to assess publication bias. Random-effects meta-regression (Berkey et al., 1995, Thompson and Sharp, 1999) and subgroup meta-analyses were conducted to investigate potential moderating factors of the relationship between SES and GI infections, guided by the harvest plot findings. Subgroup analyses were performed on study design factors and potential moderating factors of the relationship identified a priori (Rose et al., 2016).

Using forest plots, sensitivity analyses on the basis of study design, measure, adjustment, quality and whether the results used were provided by the study or were calculated from the data provided in the studies were conducted to explore the robustness of the meta-analysis.

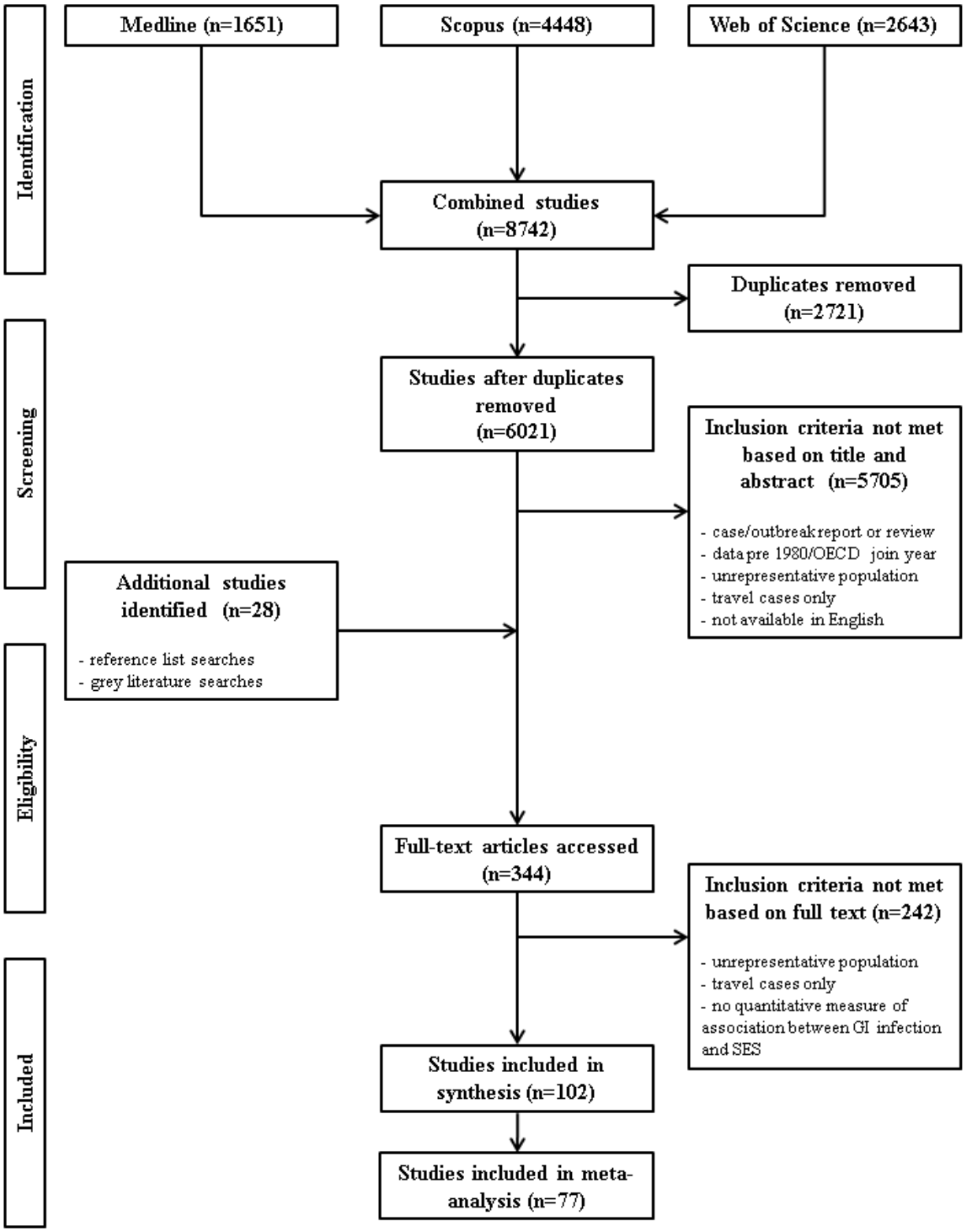
4.3 Results

Study selection

Following duplicate removal, the database search identified 6021 citations, and 344 were full-text screened. Of these, 102 were regarded as eligible for inclusion in the review and 77 were eligible for inclusion in the meta-analysis (Figure 4.1).

Figure 4.1: PRISMA flow diagram of studies included in the systematic review and meta-analysis

Source: Adapted from Moher et al. (2009)



Study characteristics

Table 4.1 shows the summary characteristics of the included studies. The majority of included studies were conducted in Europe, had ecological study designs, used laboratory records to identify GI infection cases, and did not stratify by age. Education level was identified as the most commonly used measure of SES across the studies but many studies used more than one measure. Full details of the included studies can be found in Appendix 2.4.

Table 4.1: Summary characteristics of included studies

Study characteristics	Studies (n)
Total	102
Year of publication	
Before 2000	17
2000-2005	15
2006-2010	38
After 2010	32
Level of analysis	
Individual	59
Area	43
Region	
Asia	3
Europe	49
North America	34
Oceania	16
Sample size	
<200	3
200-1000	25
1001-5000	15
5001-10000	9
10001-100000	5
>100000	45
Age category	
Children (<18 years old)	27
Adults	8
All ages combined	61
Not stated	6
Gastrointestinal infection outcome	
Acute GI infection (syndromic)	41
Campylobacteriosis	20
Cryptosporidiosis	4
Giardiasis	3
Hepatitis A	3
Listeriosis	1
Norovirus	1
Rotavirus	3
Salmonellosis	8
Shigellosis	3
Shiga toxin-producing *E. coli* infection	4
Yersinia enterocolitica	1
Multiple pathogens	10
Gastrointestinal infection measure	
Population-based survey	30
General practice (GP) presentation	5
Hospital admission	13
Laboratory records	52
Multiple measures	2

Study characteristics	Studies* (number)
Socioeconomic status measure	
Deprivation	17
Education	22
Employment	7
Income	10
Occupation	8
Social class	10
Multiple measures	28
Study design	
Case-control	25
Cohort	16
Cross-sectional	18
Ecological	43
Quality	
High	19
Medium	27
Low	56

Risk of bias within studies

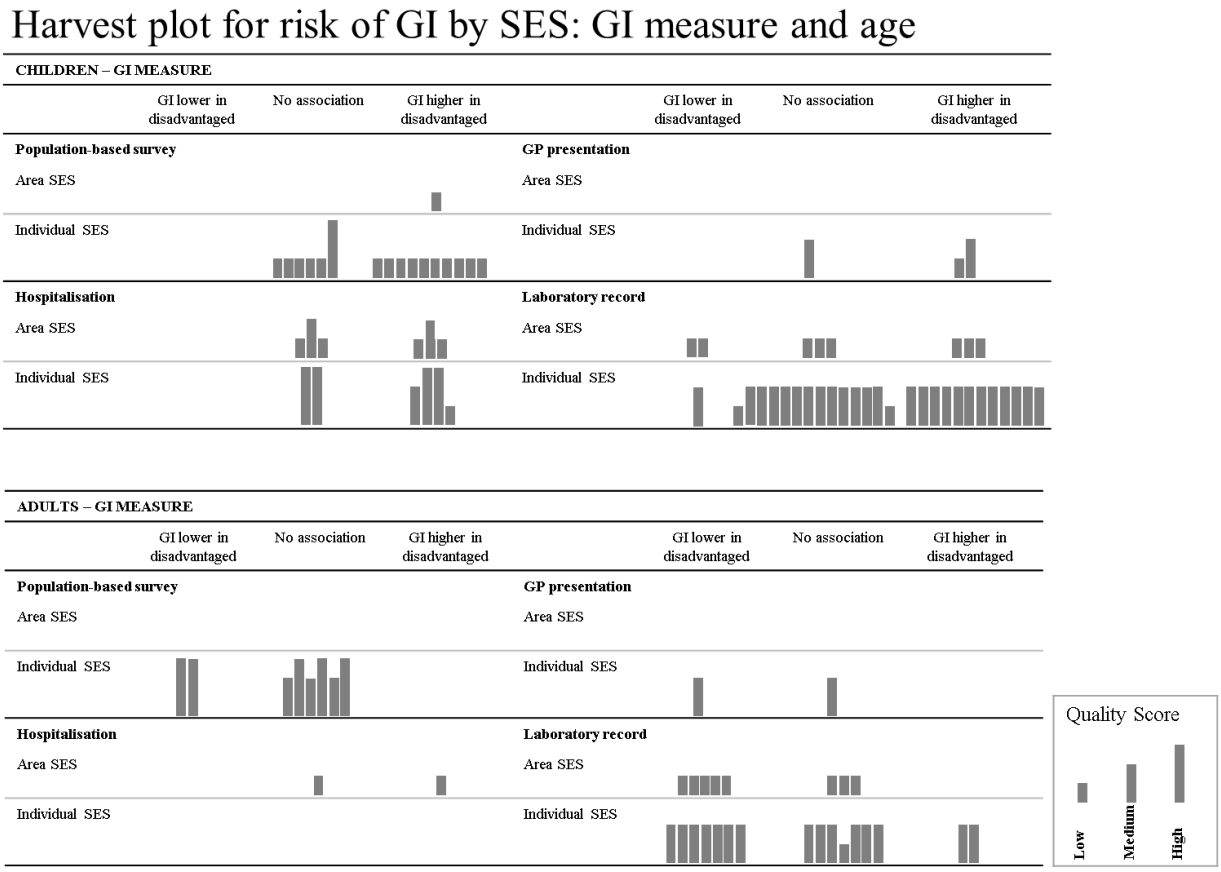
The majority of the included studies were graded as low quality (n=56). Of these there were four cross-sectional, 35 ecological, eight cohort and nine case-control studies. Twenty-seven studies were graded as being of medium quality, including seven cross-sectional, four ecological, four cohort and 12 case-control studies. Finally, 19 studies were graded as high quality, seven cross-sectional, four ecological, four cohort and four case-control studies. Details of the study quality rating for each of the included studies can be found in Appendix 2.4.

Synthesis of results

Of the 102 studies included, there were 103 point estimates for the association between SES and GI infection risk for adults or children specifically, and these point estimates were represented graphically as bars in the harvest plots. In the harvest plots, each bar represents one study. The height of the bar represents the quality of the study. Studies are classed into those showing lower risk in disadvantaged individuals/areas, no association or higher risk in disadvantaged individuals/areas.

Figure 4.2 shows the harvest plot for GI infection by SES measure, stratified by age and method of identifying GI infection cases. The harvest plot (Figure 4.2) illustrates that the relationship between SES and GI infection varied with age.

Figure 4.2: Harvest plot for risk of GI infection by SES, stratified by age, GI infection measure and SES measure



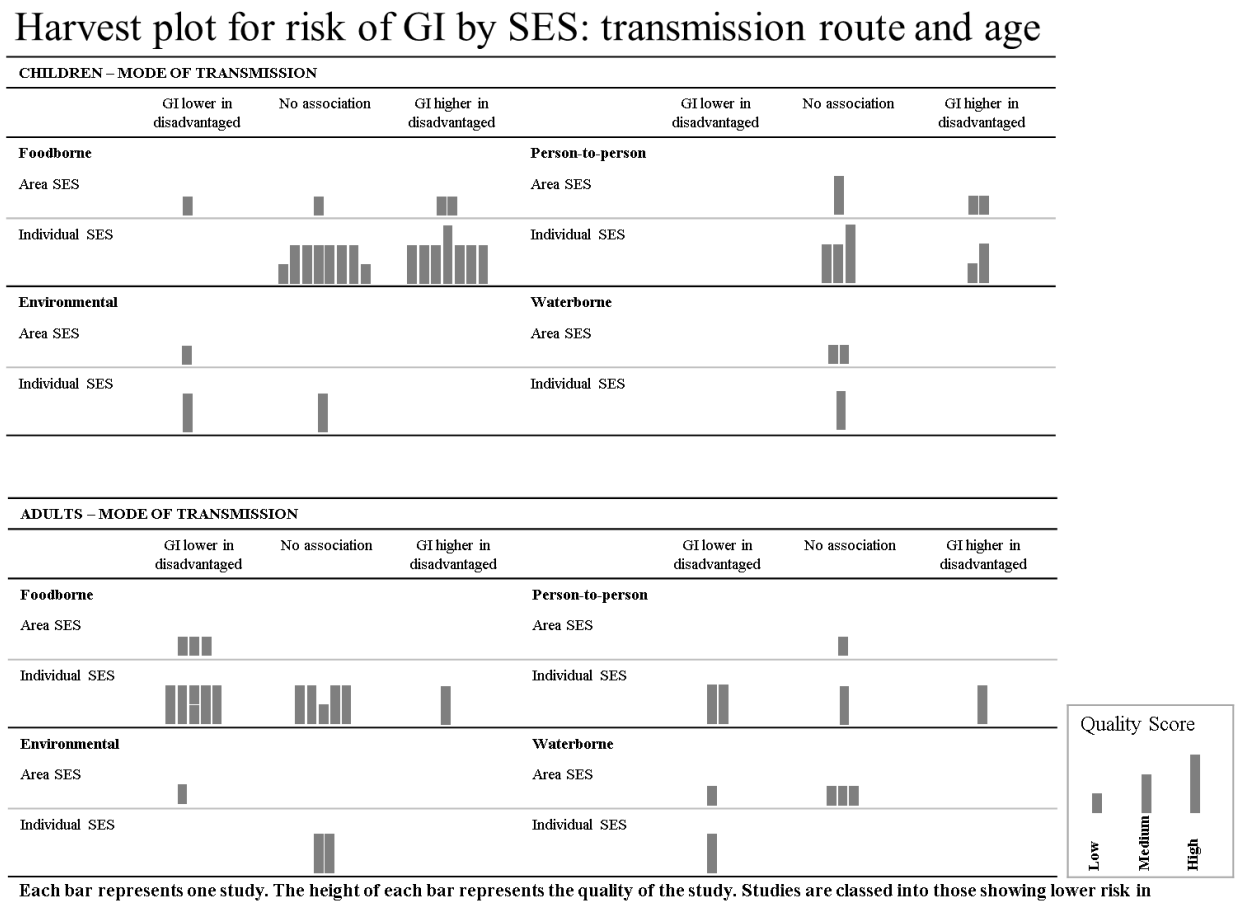
Each bar represents one study. The height of each bar represents the quality of the study. Studies are classed into those showing lower risk in disadvantaged individuals/areas, no association or higher risk in disadvantaged individuals/areas and are stratified by individual and area based SES measures.

The results for children by method of GI data collection are presented in the upper half of Figure 4.2. There was a clear social patterning for children in the reviewed studies, showing higher risk of GI infection in disadvantaged children or no association between GI infection and SES; although most studies were of low quality. With the exception of a small number of laboratory record studies, none of the studies found a lower risk of GI infection in disadvantaged children. There were also gaps in the literature using GP presentation to explore the relationship between GI infection and SES.

The results for adults by method of GI data collection are presented in the lower half of Figure 4.2. The pattern for adults was different to that for children, most studies weighted towards lower risk of GI infection in disadvantaged adults or no association. There were far fewer studies exploring the association between GI

infection and SES in adults, notable gaps in studies exploring the association using hospitalisation or GP presentation data and only low quality studies using hospitalisation data.

Figure 4.3: Harvest plot for risk of GI infection by SES, stratified by age, transmission route and SES measure



Each bar represents one study. The height of each bar represents the quality of the study. Studies are classed into those showing lower risk in disadvantaged individuals/areas, no association or higher risk in disadvantaged individuals/areas and are stratified by individual and area based SES measures.

The harvest plot stratified by age, pathogen transmission route and SES measure is presented in Figure 4.3. There were fewer studies which could be included in this harvest plot. For children (upper half of Figure 4.3), as for the previous harvest plot, the results were socially patterned towards higher risk of foodborne (*Campylobacter, Salmonella, Yersinia enterocolitica*) and person-to-person (viral GI infections, *Shigella*) GI infection in disadvantaged children and no association for waterborne infections (*Giardia, Cryptosporidium*). Only three studies explored the relationship between predominantly environmental GI infections (STEC) and SES and none was of high quality.

For adults (lower half of Figure 4.3), there were also notable gaps in studies exploring the relationship in environmental or waterborne GI infections. There was a clear pattern with studies reporting lower risk in more disadvantaged adults or no association for studies exploring the relationship between predominantly foodborne GI infections and SES, and these studies were generally of medium quality.

No clear difference was observed in the relationship between SES and GI infection when comparing point estimates based on area and individual SES measures, or when comparing point estimates from different countries (based on level of development or climate) (data not shown).

Meta-analysis results

Of the 102 studies included in this systematic review, 77 studies were included in the meta-analysis. These 77 studies contributed 83 effect estimates. Summary data including effect estimates and confidence intervals for these studies can be found in Appendix 2.8. Of the 25 studies that could not be included in a meta-analysis, 15 did not provide sufficient quantitative data, six did not use a dichotomous outcome and four analysed the same cases as other studies (Appendix 2.4). Since age was highlighted as a potentially important effect modifier in the harvest plots, estimates from the same study stratified by age were retained individually in the meta-analysis to allow for the investigation of this variable.

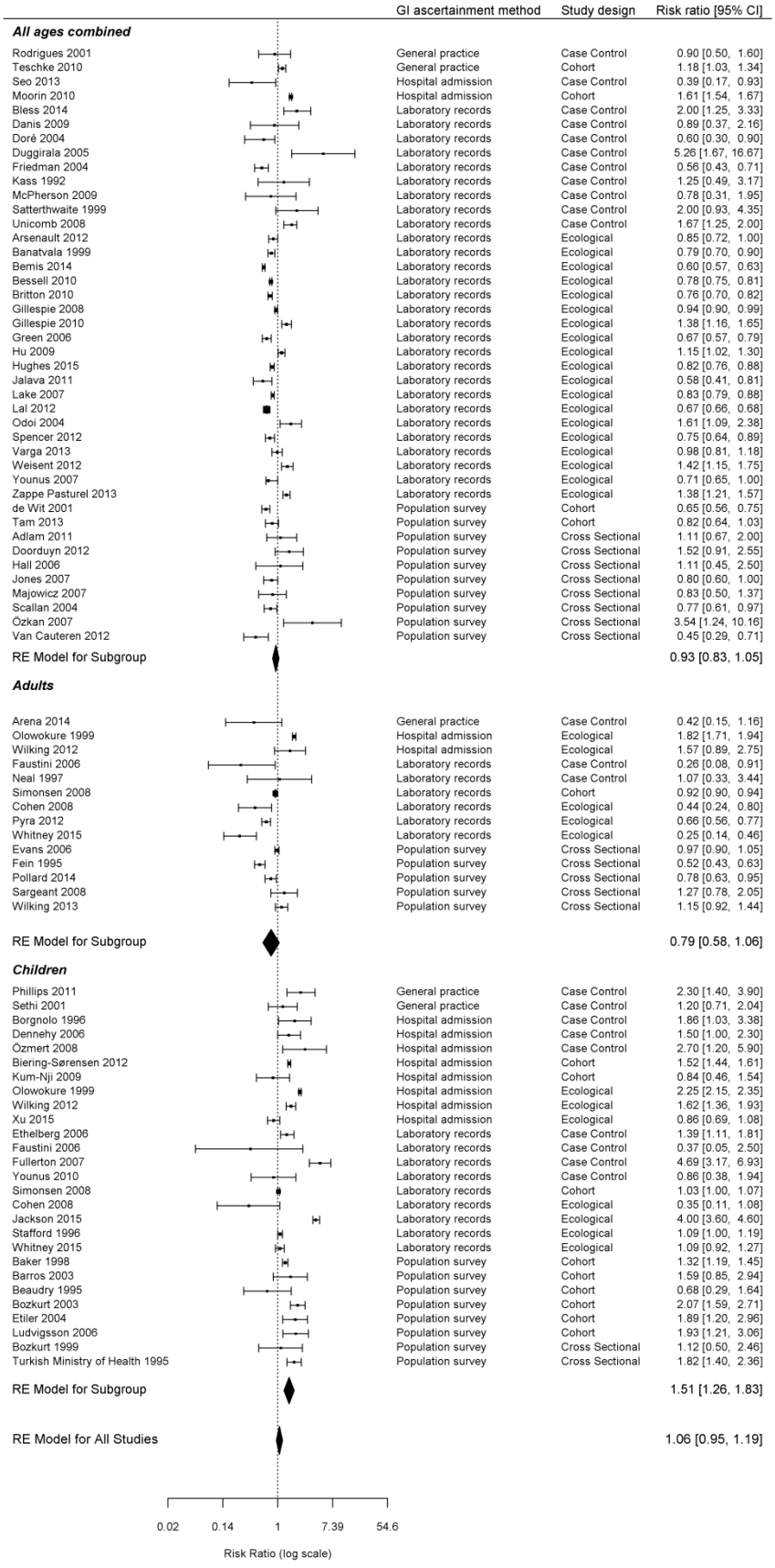
Three main outliers were identified in the forest plot (Figure 4.4; Özkan et al. (2007), Jackson et al. (2015), Fullerton et al. (2007)). Two of these studies (Fullerton et al., 2007, Jackson et al., 2015) were conducted in the United States using laboratory records, and Özkan et al. (2007) was conducted in Turkey.

The pooled risk ratio for GI infection comparing low verses high SES for all studies combined was 1.06 (95% CI 0.95–1.19), with considerable statistical heterogeneity (I^2 99.08%). A forest plot for the studies stratified by age is shown in Figure 4.4. For children, the pooled risk ratio was 1.51 (95% CI 1.26–1.83) with I^2 97.87%. For adults, the pooled risk ratio was 0.79 (95% CI 0.58–1.06) with I^2 98.64%.

Stratifying by GI ascertainment method, to investigate whether stratifying by one likely source of variability produced subsets with lower heterogeneity, did find slightly lower levels of heterogeneity for cases identified in population surveys and

GP presentation studies although there was considerable statistical heterogeneity remaining for all four GI ascertainment methods - population survey: I^2 91.97%; GP presentation: I^2 82.93%; laboratory records: I^2 99.45% ; hospital admission: I^2 98.03%).

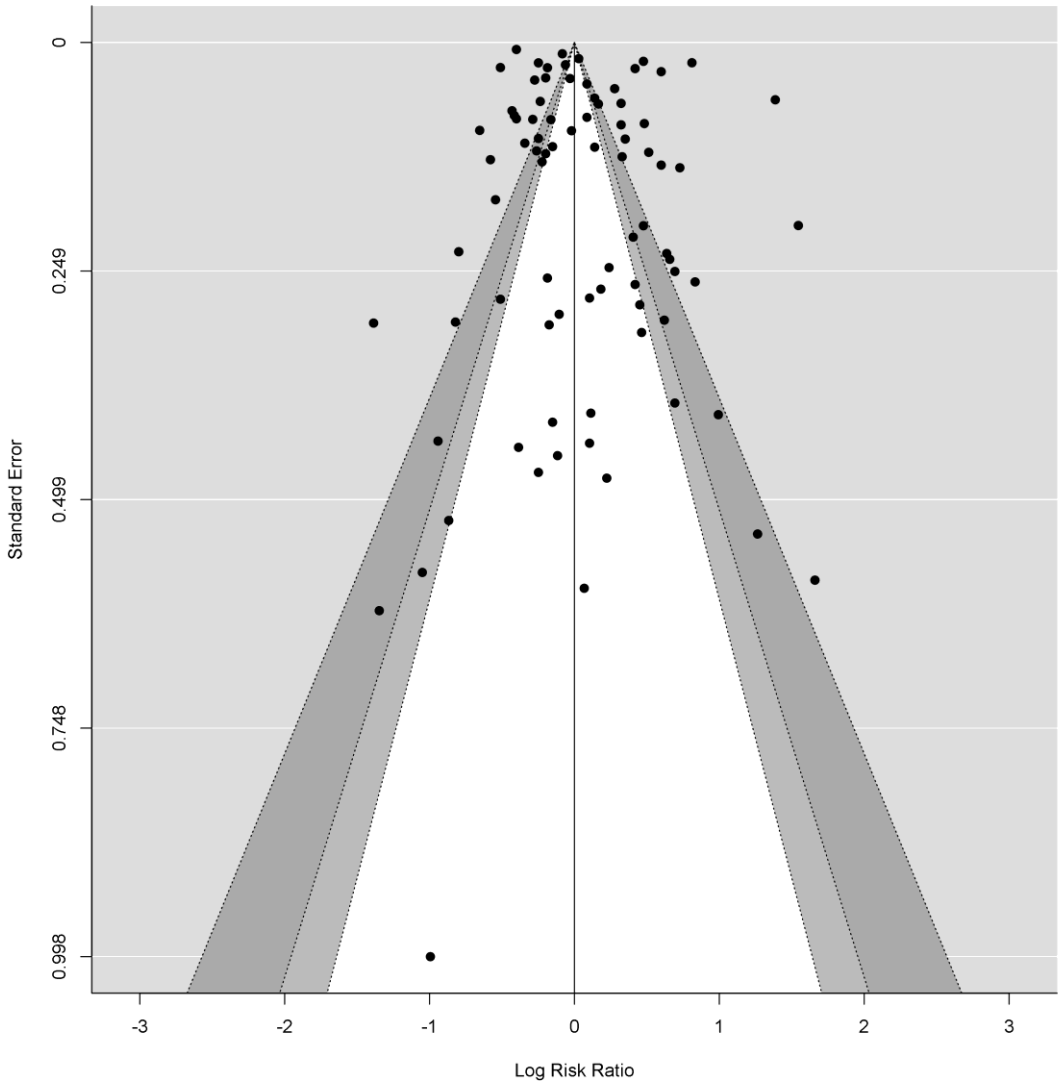
Figure 4.4: Forest plot for all studies by age, GI ascertainment method and study design (n=83 effect estimates; 77 studies)



Risk of bias across studies

A contour-enhanced funnel plot (Peters et al., 2008, Sterne and Egger, 2001, The Metafor Package: a meta-analysis package for R, 2004) was produced to assess publication bias (Figure 4.5). Points within the plot appeared largely symmetrical, indicating that publication bias was unlikely.

Figure 4.5: Contour enhanced funnel plot



White = P-value > 0.10
Mid-grey = P-value between 0.10 and 0.05
Dark-grey = P-value between 0.05 and 0.01
Region outside of funnel = P-value < 0.01

Potential effect modifiers and sources of heterogeneity were further explored in a multivariable random-effects meta-regression in an attempt to quantitatively explain some of the observed heterogeneity. In univariate meta-regression the risk of GI infection for low compared to high SES was on average significantly higher among studies that analysed hospitalised cases, and non-significantly higher among studies that analysed cases identified via population-based surveys and general practices, compared to studies that analysed laboratory recorded cases (Table 4.2).

Amongst studies using laboratory records, the risk of GI infection for low compared to high SES was significantly lower among studies that analysed environmental pathogens, and significantly higher among studies that analysed person-to-person pathogens, compared to studies that analysed foodborne pathogens. There was no statistically significant difference in risk of GI infection by SES between studies conducted in countries with different climates and levels of development. Additionally, the risk of GI infection for low versus high SES was non-significantly lower among studies that used area-level compared to individual-level SES measures.

In multivariate meta-regression (excluding pathogen type since not all studies analysed specific pathogens), age was identified as the only statistically significant potential effect modifier. The risk ratios for GI infection between low and high SES groups observed by studies that analysed children were on average 1.87 times higher than the risk ratios observed by studies that focused on adults, after controlling for other study differences included in the model (Table 4.2).

Table 4.2: Univariate and multivariate random-effects meta-regression for GI infection

Variable	Category	n	Univariate	Multivariable
			RR (95% CI)	**RR (95% CI)**
GI ascertainment method	Laboratory records	43	1.00 (reference)	1.00 (reference)
	Population-based survey	23	1.11 (0.85–1.44)	1.04 (0.75–1.43)
	GP presentation	5	1.18 (0.71–1.94)	1.02 (0.62–1.69)
	Hospital admissions	12	1.49 (1.08–2.07)*	1.24 (0.81–1.73)
SES	Individual level	50	1.00 (reference)	1.00 (reference)
	Area level	33	0.87 (0.69–1.09)	0.92 (0.70–1.22)
Age	Adult	15	1.00 (reference)	1.00 (reference)
	All ages	41	1.17 (0.88–1.54)	1.22 (0.90–1.66)
	Child	27	1.89 (1.40–2.55)***	1.87 (1.35–2.59)***
Country Human Development Index†	Upper tertile	39	1.00 (reference)	1.00 (reference)
	Middle tertile	30	0.98 (0.76–1.25)	1.09 (0.84–1.41)
	Lower tertile	14	1.04 (0.73–1.49)	0.88 (0.62–1.25)
Country climate	Temperate/Mediterranean	62	1.00 (reference)	1.00 (reference)
	Arid	7	1.05 (0.69–1.61)	1.01 (0.67–1.52)
	Snow	14	0.81 (0.60–1.10)	0.89 (0.67–1.19)
Pathogen type‡	Foodborne	28	1.00 (reference)	-
	Waterborne	8	0.73 (0.46–1.14)	-
	Environmental	3	0.46 (0.23–0.91)*	-
	Person to person	7	1.65 (1.05–2.59)*	-

CI - confidence interval; GI - gastrointestinal; RR - ratio of risk ratios; SES - socioeconomic status
† Higher values indicate higher level of human development
‡ Not all studies analysed specific pathogens and therefore this variable was not included in the multivariate model
*P-value < 0.05; ***P-value < 0.001

Sensitivity analyses

Sensitivity analyses of stratified forest plots restricted to high and medium quality studies only; adjusted estimates only; studies which did not control for or match on SES only; and studies which provided results rather than raw data through which estimates were calculated are summarised in Table 4.3. These analyses did not differ from the findings of the main forest plot.

Table 4.3: Risk ratios and 95% confidence intervals for sensitivity analyses

Restriction		Risk Ratio	95% CI
High and medium quality studies only	All ages combined	0.95	0.80-1.15
	Adults	0.82	0.64-1.05
	Children	1.47	1.03-2.09
Adjusted estimates only	All ages combined	0.94	0.82-1.08
	Adults	0.82	0.66-1.03
	Children	1.47	1.12-1.94
Studies which did not control for or match on SES only	All ages combined	0.93	0.81-1.06
	Adults	0.71	0.43-1.19
	Children	1.45	1.16-1.82
Studies which provided results only	All ages combined	0.94	0.81-1.09
	Adults	0.79	0.63-0.97
	Children	1.54	1.22-1.95

4.4 Discussion

This systematic review and meta-analysis of observational studies from high income countries provides evidence of an association between lower SES and a higher risk of GI infections for children, but not for adults. Overall, age explained a small proportion of the heterogeneity observed across the studies as a whole.

No previous studies conducting a meta-analysis on this topic have been identified through an accurate search of the literature and therefore, to the best of my knowledge, this study represents the first meta-analysis on the topic. A broad range of study designs and data sources were included in this systematic review, as well as a wide range of definitions of GI infections. One of the main strengths of the harvest plot is the ability to include all studies, not exclusively studies with a quantitative measure (Ogilvie et al., 2008). This allowed a more comprehensive exploration of heterogeneity and provided important insights to inform the subsequent meta-analysis. The harvest plots showed a clear pattern towards higher risk in more disadvantaged children or no association although the association was less clear for adults and notable gaps in studies exploring GP presentation in children and hospitalisation in adults were observed. Selection bias was mitigated by double screening throughout.

The inclusion of syndromic definitions of GI infections in the absence of laboratory confirmation was a particular strength of this analysis as the decision to seek healthcare may in fact be related to SES and attempting to increase our understanding of the role of SES in GI infection is a key aim of this thesis. As such, it was possible to identify and include literature on the burden of GI symptoms by SES and capture population groups who may not usually seek healthcare for their illness and that, consequently, may not be included in studies relying on laboratory data to identify cases.

The potential for publication bias was explored, and this was not evident in the funnel plot. Subgroup analyses were defined a priori (Rose et al., 2016) which minimised the potential bias which can be introduced through performing multiple analyses of the data. Furthermore, these results reflect trends in inequalities of GI infections across numerous high income countries and should therefore be generalisable.

To explore sources of heterogeneity, stratified meta-analyses and meta-regression were performed on the basis of factors mentioned in the literature. Despite this, a large amount of heterogeneity remained unexplained. As seen in the forest plot (Figure 4.4), effect estimates were similar across studies although there were some outliers with wide confidence intervals combined with studies with narrow confidence intervals, which may provide some explanation for the extreme statistical heterogeneity observed. It is possible that factors that could not be adjusted for may explain the high residual heterogeneity. The studies covered a broad range of healthcare and social systems such as countries without a national health service or countries in transition between low/middle and high income. As such, different countries, such as the USA, Sweden and Turkey are likely to have different incidences and proportions of various GI pathogens, as well as wide differences in absolute and relative poverty. No differences were identified when comparing estimates from different countries (based on level of development or climate).

SES was measured in numerous ways, and categorisation of low and high SES may have differed considerably between studies. Where studies reported estimates for multiple SES measures, education was chosen. This may have introduced bias if different SES measures reveal different social patterning of risk of GI infection in adults. This study did not aim to explore the differential effects of different measures of SES on GI infection, but rather focused on the overall impact of SES on GI infections, and in fact SES measures are likely to be highly correlated.

The primary aims of the separate studies varied, as did the variables used to statistically adjust the associations between SES and GI infection. Further, non-English language studies were excluded due to time limitations and costs of translating studies, and countries that have been in transition between low/middle and high income (such as Turkey) were included, both of which could potentially limit the interpretation of the results. Lastly, the large amount of heterogeneity may have negatively affected the power to detect statistically significant modifiers in the meta-regression (Table 4.2), and therefore non-significance should not necessarily be interpreted as evidence that a potential modifier had no effect on the relationship between SES and risk of GI infection (Hempel et al., 2013).

There was a large number of ecological studies and studies assessed as generally low quality with potential for bias due to study design; such as case-control studies, several of which selected controls based on the geographical residence of cases or through case-nomination, thereby potentially biasing the relationship between SES and GI infections towards the null. The results were similar when sensitivity analyses were conducted excluding studies which controlled for or matched by SES.

Despite the remaining heterogeneity, this study suggests that the relationship between SES and GI illness varies with stage in the life course, with disadvantaged children at greater risk of GI infection compared to more advantaged children. This is the first systematic review and meta-analysis conducted in-depth on this topic. Newman et al. (2015) undertook a systematic review of the association between SES and foodborne illness, a subset of GI infections; but they did not look at differences across the life course or different levels of healthcare reporting such as hospitalisation. Our results corroborate those of Newman et al. (2015) for foodborne and laboratory confirmed pathogen-specific results, in that there were no consistent trends across all studies or pathogens for a single SES measure. Exploring this more broadly by predominant mode of transmission, risk of GI infection was significantly higher in lower SES groups for pathogens spread via person-to-person transmission and significantly lower in lower SES groups for pathogens spread via environmental contamination compared to foodborne transmission. This may indicate differential effects of SES by pathogen type (Simonsen et al., 2008).

It could be speculated that children may be more likely to be taken to seek medical help regardless of SES, so the higher risk of GI infection seen in children might reflect real differences by SES, rather than bias due to differential healthcare seeking behaviour. Our findings may reflect differential exposures or immunity by SES in children. A study conducted in the UK found that seropositivity to *Helicobacter pylori* in adults was associated with lower SES and adverse housing conditions in childhood (Pearce et al., 2013). Within low income countries, C*ampylobacter* is almost exclusively seen in disadvantaged children (Fernández et al., 2008, Kakai et al., 1995, Lloyd-Evans et al., 1983, Quetz et al., 2010) while adults are rarely infected or identified (Coker et al., 2002), potentially due to differential healthcare interaction or exposure risk due to poor sanitation. This pattern is also seen for other bacteria and intestinal parasites, such as *Shigella* and *Giardia* (Nematian et al., 2004,

O'Ryan et al., 2005). As such, it may be that disadvantaged children are more exposed to these GI infections in childhood and that re-exposure may lead to improved immunity and subsequent asymptomatic infection later in life.

It is important to note that many of the studies analysing hospital admission cases also analysed child cases only, therefore it is challenging to separate out the potential modifying effects of these variables. Future research investigating the relationship between SES and GI infection should provide results stratified by adult and child age groups wherever possible.

To conclude, in high income countries disadvantaged children, but not adults, are at greater risk of GI infection compared to their more advantaged counterparts. Strategies to improve childhood socioeconomic conditions have potential to reduce the burden of GI illness. Gaining greater insight into this relationship will help to inform policies to reduce the health inequalities identified.

4.5 Interpretation

In this chapter I have provided novel evidence to suggest that there is differential risk of GI infection by SES for children in high income countries, with higher risk of GI infections amongst disadvantaged children, but that this inequality is not as strong for adults. Furthermore, the results of this study have also highlighted the necessity to provide an update of the role of SES in GI infection risk in the UK, with the majority of UK-based studies included in this review published prior to 2010, with many using data collected in the 1990s.

Studies combining all ages may mask differential risk in children. Subsequent studies in this thesis will stratify by age, and inequalities in risk of GI infection across the life course will be explored throughout the remainder of this thesis.

The findings of this chapter could suggest differential symptom recognition or healthcare interaction across individuals of different SES, although this is likely to be minimised for children as they may be more likely to be taken to seek care regardless of their SES. This hypothesis will be tested further in Studies 2, 3 and 4 (Chapters 5-7 respectively) through the analysis of a longitudinal community cohort assessing risk of IID, calls to telephone-based health advice services and through the analysis of social patterning of risk factors and exposures for a severe GI infection.

Additionally, the findings of this chapter may reflect differential severity (Baker et al., 2012, Olowokure et al., 1999, Phillips et al., 2011, Pockett et al., 2011, Rose et al., 2017); many of the studies using hospital admission as the GI ascertainment method explored the relationship between GI infection and SES in children and therefore may have ascertained more severe illness requiring hospitalisation. Children who are more disadvantaged may have a genuinely higher risk of GI infections. This may be related to social patterning of risk factors and exposures, which could influence vulnerability to GI infections in childhood, a hypothesis that will be further tested in Study 4 (Chapter 7). This may also explain the lack of a significant differential risk in adults; any increased exposure in disadvantaged adults may then be masked by immunity gained in childhood.

Chapter 5 – Study 2 (Objective 2)
Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 Study)

Abstract

Background

Infectious intestinal diseases (IID) are common, affecting around 25% of people in the UK each year at an estimated annual cost to the economy, individuals and the NHS of £1.5 billion. Whilst there is evidence of higher IID hospital admissions in more disadvantaged groups, the association between socioeconomic status (SES) and risk of IID infection remains unclear. This study aims to investigate the relationship between SES and IID in a large community cohort.

Methods

Longitudinal analysis of a prospective community cohort in the UK following 6,836 participants of all ages in 2008-2009 was undertaken. Hazard ratios for IID by SES were estimated using Cox proportional hazard, adjusting for follow-up time and potential confounding factors.

Results

In the fully adjusted analysis, hazard rate of IID was significantly lower amongst routine/manual occupations compared to managerial/professional occupations (HR 0.75, 95% CI 0.61-0.91).

Conclusion

In this large community cohort, lower SES was associated with lower IID risk. This could be explained by the low response rate which varied by SES, with disadvantaged individuals being underrepresented and which may underestimate their risk. It may be related to differences in exposure or in recognition of IID symptoms by SES. Higher hospital admissions associated with lower SES observed in some studies could relate to more severe consequences, rather than increased infection risk.

5.1 Introduction

This study was designed to address gaps in the literature identified in the literature review (Chapter 2) and Study 1. The objectives of the research within this thesis are detailed in Chapter 1. This chapter seeks to meet Objective 2: To investigate the extent and nature of socioeconomic inequalities in risk of GI infections in the community in the UK, with estimates derived from the most up-to-date population-based household survey.

Gastrointestinal (GI) infections are common, leading to diarrhoea, vomiting and, occasionally, more serious complications such as renal failure. Previous estimates suggest around 25% of people in the UK suffer an episode of infectious intestinal disease (IID) per year (Tam et al., 2011a) and that foodborne illness, a subset of IID, in England and Wales costs individuals, the economy and NHS around £1.5 billion annually (Tam et al., 2011a). Many infections are socially patterned, but the role of socioeconomic status (SES) in risk of GI infection in high income countries, such as the UK, is not well understood (Newman et al., 2015).

Socioeconomic inequalities in GI infections could result from differences in risk of, or vulnerability to, infection between socioeconomic groups or from differences in the consequences of infection, with some groups potentially having more severe infection, requiring more healthcare, and experiencing greater disruption in daily activities (Diderichsen et al., 2001). A large proportion of the burden of GI infection remains hidden; it is estimated that there are 147 cases in the community for every one case that is reported to national surveillance (Tam et al., 2011a); many individuals do not present to healthcare as most infections are self-limiting. Additionally, it is unclear whether socioeconomic patterns reported in hospital and laboratory-based surveillance reflect differences in risk of infection or in reporting and healthcare interaction on the basis of SES more generally (Dunlop et al., 2000). Longitudinal population-based survey data can provide a better estimate of differences in risk of infection that may not be captured through routine surveillance.

This study aims to explore whether different socioeconomic groups experience different risk of GI infection in the UK, through the analysis of a large prospective population cohort, to improve understanding of the role of SES in IID in the community and to inform policies to reduce health inequalities. This study provides

an up-to-date assessment of the association between IID and SES for all ages in the UK. In this chapter I will present the results from the published analysis of this community cohort (Adams et al., 2016a, Adams et al., 2017), including details of sensitivity and robustness analyses and the implications of the findings for future work and the impact on our understanding for public health.

5.2 Methods

The methods adopted in this study are reported in detail in Chapter 3. A brief overview is provided below.

Design, setting and data source

A longitudinal analysis of a large prospective community cohort in the UK was conducted using data collected through the IID2 Study (Tam et al., 2011a). Greater detail on the IID2 Study is provided in Chapters 2 and 3. A cohort of 6,836 randomly-selected participants was recruited from 88 representative general practices in the UK. Sociodemographic information including age, gender and occupation were obtained through a baseline survey upon entry to the cohort and details of IID symptoms were recorded on a weekly basis for up to one year, from October 2007 to August 2009, through the return of an email or postcard indicating whether symptoms of diarrhoea and/or vomiting had been experienced in the previous week. Individuals who reported symptoms completed a more in-depth questionnaire through which details of illness and healthcare contact were recorded.

Participants

The 6,836 participants contributed 4,658 person-years of follow-up; median follow-up duration was 39 weeks (Tam et al., 2011a). Overall participation rate was low, only 9% of those initially invited to take part, and individuals who declined to participate were younger, more disadvantaged, living in urban rather than rural areas and employed in lower supervisory and technical occupations (Tam et al., 2011a). Average follow-up time was similar for those who experienced an episode of IID and those who did not. Managerial/professional occupations were over-represented in the study, while intermediate, and semi-routine and routine occupations were under-represented, in comparison to the UK population. Those of White ethnicity were slightly over-represented (Tam et al., 2011a). These differences may limit the

generalisability of the results, but, amongst participants, no differences in follow-up were identified by sex, socioeconomic status or rural-urban classification and the internal associations are likely to be valid.

Outcome and covariates

The primary outcome, infectious intestinal disease, was defined as loose stools or clinically significant vomiting (vomiting occurring more than once in 24-hours and if it incapacitated the case or was accompanied by other symptoms such as cramps or fever (Tam et al., 2011a)) lasting less than two weeks, in the absence of a known non-infectious cause, preceded by a symptom-free period of three weeks (Tam et al., 2011a). Cases experiencing illness considered to be travel-related were excluded.

The primary exposure of interest was an individual-level measure of SES, self-reported occupation, with each individual assigned a National-Statistics Socioeconomic Classification (NS-SEC) using the five-class self-coded version (Office for National Statistics). For participants aged less than 16 years, NS-SEC was assigned based on the occupation of the head of the household. For the purposes of this study, the NS-SEC variable was recoded into the three-class version to provide a hierarchy of socioeconomic status, with routine and manual occupations assumed equivalent to low SES and managerial/professional occupations to high SES (Office for National Statistics).

Analysis strategy

Analyses were conducted in Stata 13.1 (Statacorp, Texas). Rates of IID within the study population and by SES were calculated accounting for follow-up time, to produce rates of IID per 1,000 person-years with associated 95% confidence intervals. The main analysis investigated the relationship between SES, as measured by NS-SEC, and time to first IID episode for each participant using Cox proportional hazard regression modelling, with subsequent episodes of IID for an individual being dropped. Univariate relationships were explored between SES and the covariates of interest (rurality and employment status (employed/not working)) before fitting a multivariable Cox proportional hazard regression model, adjusting for the potentially confounding covariates and stratifying the baseline hazard on age and sex. Proportionality was tested using a log-log plot and Kaplan-Meier survival curves were estimated to check the proportional hazards assumption. Interaction terms

between the socioeconomic variable NS-SEC and each variable in turn were tested for inclusion to investigate whether the strength of any relationship was moderated by the inclusion of another variable.

A number of robustness tests were conducted. Firstly, allowing individuals with multiple episodes of IID to re-enter the cohort following a period of censoring (due to symptoms meeting the case definition and requiring a censored period of three weeks after cessation of symptoms; non-response; or symptoms not meeting the case definition), accounting for clustering within individuals by using a robust estimate of variance allowing for inter-person correlation. Secondly, the analysis was repeated using a less sensitive case definition, whereby individuals reporting symptoms which could not be verified against the case definition (due to a lack of further details about foreign travel or symptom duration) were also included as cases in the analysis.

To assess the role of the not-classifiable NS-SEC groups, the analysis was repeated including those unclassifiable within NS-SEC to investigate whether this had an impact on the results. This NS-SEC group comprised individuals for whom it was not possible to classify their occupation or who did not respond to occupation questions. As there were missing NS-SEC data for a group of participants for whom it was not possible to classify their occupation or who did not respond to the occupation question, Multiple Imputation using chained equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was also used in order to include these cases. Finally, the analysis was repeated using an area-level measure of SES, the Index of Multiple Deprivation (IMD) (Department for Communities and Local Government, 2011), assigned to each individual based on their postcode.

Stratification by age groups were also conducted as sensitivity analyses to determine whether there were differences in the rate of IID by SES for children, adults and older participants.

5.3 Results

Characteristics of participants

A total of 6,836 participants were included in the study. NS-SEC was not classifiable for 1,112 individuals (16.3%), employment status was missing for five individuals

and rural-urban classification was missing for three individuals. After excluding records with missing data, 5,716 individuals were included in the main analysis.

Of the 6,836 participants in the cohort, 998 individuals reported an episode of IID during 4583.5 person-years of follow-up. Fifty-two percent (n=3,557) were from managerial/professional occupations, 15% (n=1,002) were from intermediate occupations and 17% (n=1,165) were from routine/manual occupations, compared to 31%, 22% and 33% respectively in the general population (Appendix 3.6) (Office for National Statistics [dataset], 2011). Socioeconomic status was associated with IID, age group, sex, rurality, employment status and the method of follow-up that participants elected to use (email or postcard) (Table 5.1). Mean follow-up time was similar between NS-SEC groups.

Table 5.1: Characteristics of cohort participants by NS-SEC (n=6,836)

		Managerial/professional n (%)	Intermediate n (%)	Routine/ manual n (%)	Not classifiable n (%)	p value[a]
Total		3,557 (52.0)	1,002 (14.7)	1,165 (17.0)	1,112 (16.3)	<0.001
Incidence rate /1,000PYs[¥]		235.4	243.9	166.3	194.0	
Follow-up time (mean days)		242.1	240.6	245.2	257.4	
Age (mean)		47.2	48.5	49.3	53.0	
Case	Yes	555 (15.6)	161 (16.1)	130 (11.2)	152 (13.7)	<0.01
	No	3,002 (84.4)	841 (83.9)	1,035 (88.8)	960 (86.3)	
Age Group	<18	605 (17.0)	152 (15.2)	178 (15.3)	156 (14.0)	<0.001
	18-64	2,095 (58.9)	593 (59.2)	627 (53.8)	525 (47.2)	
	65+	857 (24.1)	257 (25.7)	360 (30.9)	431 (38.8)	
Sex	Female	2,175 (61.2)	669 (66.8)	640 (54.9)	676 (60.8)	<0.001
	Male	1,382 (38.9)	333 (33.2)	525 (45.1)	436 (39.2)	
Ethnicity	White	3,464 (97.4)	981 (97.9)	1,145 (98.3)	1,077 (96.9)	0.13
	Non-White	93 (2.6)	21 (2.1)	20 (1.7)	35 (3.2)	
Rurality	Urban	2,522 (70.9)	694 (69.3)	958 (82.3)	789 (71.0)	<0.001
	Rural	1,034 (29.1)	307 (30.7)	206 (17.7)	323 (29.1)	
Follow-up	Email	2,564 (72.1)	622 (62.1)	529 (45.4)	539 (48.5)	<0.001
	Postcard	993 (27.9)	380 (37.9)	636 (54.6)	573 (51.5)	
Employment status	Employed	2,493 (70.2)	713 (71.3)	769 (66.0)	423 (38.9)	<0.001
	Not working	1,061 (29.9)	287 (28.7)	396 (34.0)	664 (61.1)	

[¥]Person-Years

[a]Statistical significance of relationship between NS-SEC and each variable, tested using χ^2 test

Missing data: Employment status was missing for 30 individuals. Rural-urban classification was missing for three individuals.

As shown in Tables 5.2 and 5.3, incidence was significantly lower among routine/manual occupations compared to managerial/professional occupations (166 per 1,000 person-years, 95% CI 140-197 per 1,000 person-years compared to 235 per 1,000 person-years, 95% CI 217-256 per 1,000 person-years); IRR 0.71, 95% CI 0.58-0.86).

Table 5.2: Rates of IID by NS-SEC and explanatory variables

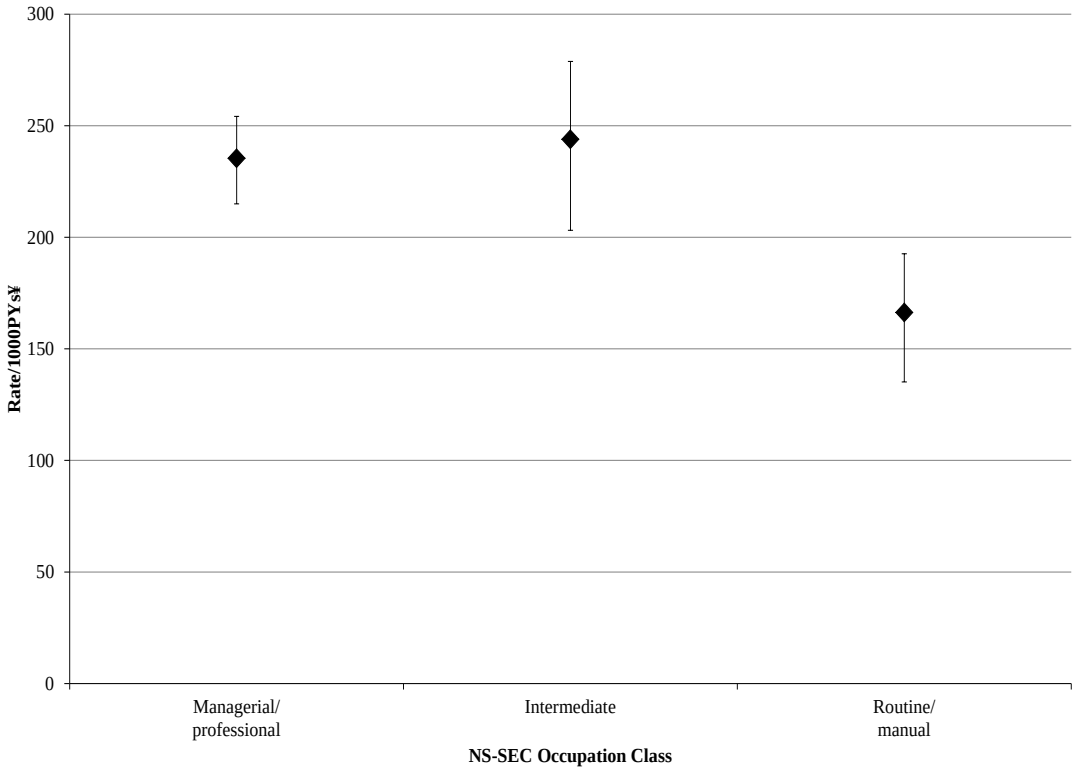
		Cases	PY[¥]	Rate*	(95% CI)
NS-SEC	Managerial/professional	555	2357.9	235.4	(216.6-255.8)
	Intermediate	161	660.0	243.9	(209.0-284.7)
	Routine/manual	130	781.9	166.3	(140.0-197.4)
	Not classifiable	152	783.6	194.0	(165.5-227.4)
Age Group	<18	237	678.7	349.2	(307.5-396.6)
	18-64	537	2534.6	211.9	(194.7-230.6)
	65+	224	1370.2	163.5	(143.4-186.4)
Sex	Female	654	2759.4	237.0	(219.5-255.9)
	Male	344	1824.0	188.6	(169.7-209.6)
Ethnicity	White	979	4474.8	218.8	(205.5-232.9)
	Non-White	19	108.7	174.8	(111.5-274.0)
Rurality	Urban	688	3310.7	207.8	(192.8-223.9)
	Rural	310	1271.0	243.9	(218.2-272.6)
Follow-up	Email	663	2781.6	238.4	(220.9-257.2)
	Postcard	335	1801.9	185.9	(167.0-206.9)
Employment status	Employed	690	2857.8	241.4	(224.1-260.1)
	Not working	301	1705.6	176.5	(157.6-197.6)

[¥]Person-Years
*Rate per 1,000 person-years
NS-SEC: National Statistics- Socioeconomic Classification; CI: confidence interval
Missing data: Employment status was missing for 30 individuals. Rural-urban classification was missing for three individuals.

Table 5.3: Incidence rate ratio for exposed (routine/manual occupations) compared to unexposed (professional/managerial occupations)

Rate in exposed	Rate in unexposed	Incidence rate difference	Incidence rate ratio	95% CI
166.3	235.4	-69.1	0.71	0.58-0.86

Figure 5.1: Incidence rates per 1,000 person-years by NS-SEC classification



¥Person-Years; NS-SEC: National Statistics-Socioeconomic

Proportionality was tested using a log-log plot (Figure 5.2) and Kaplan-Meier plot (Figure 5.3) showing time to occurrence of first episode of IID by NS-SEC. The parallel curves indicated that differences are approximately constant and therefore the assumption of proportionality was upheld. Tests of the proportional hazards assumption of each predictor as well as a global test were performed. Neither the individual predictors nor the global test were significant (global test p=0.44).

Figure 5.2: Log-log plot of time to occurrence of first episode of IID by NS-SEC

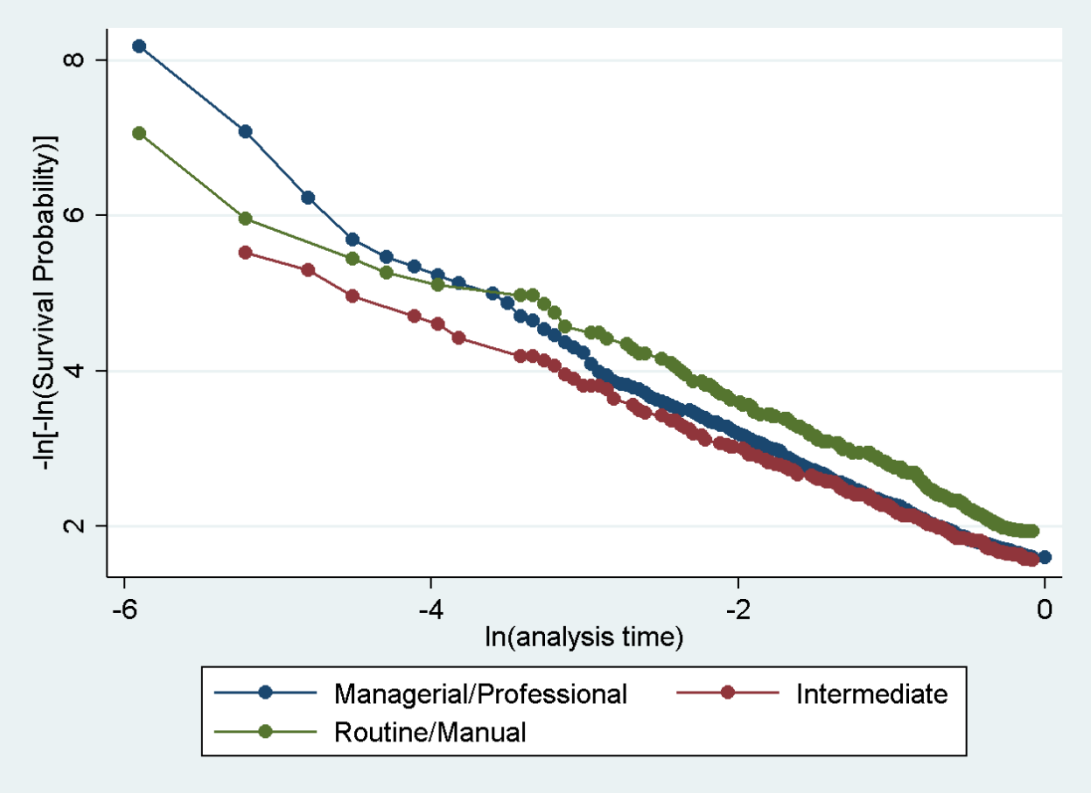
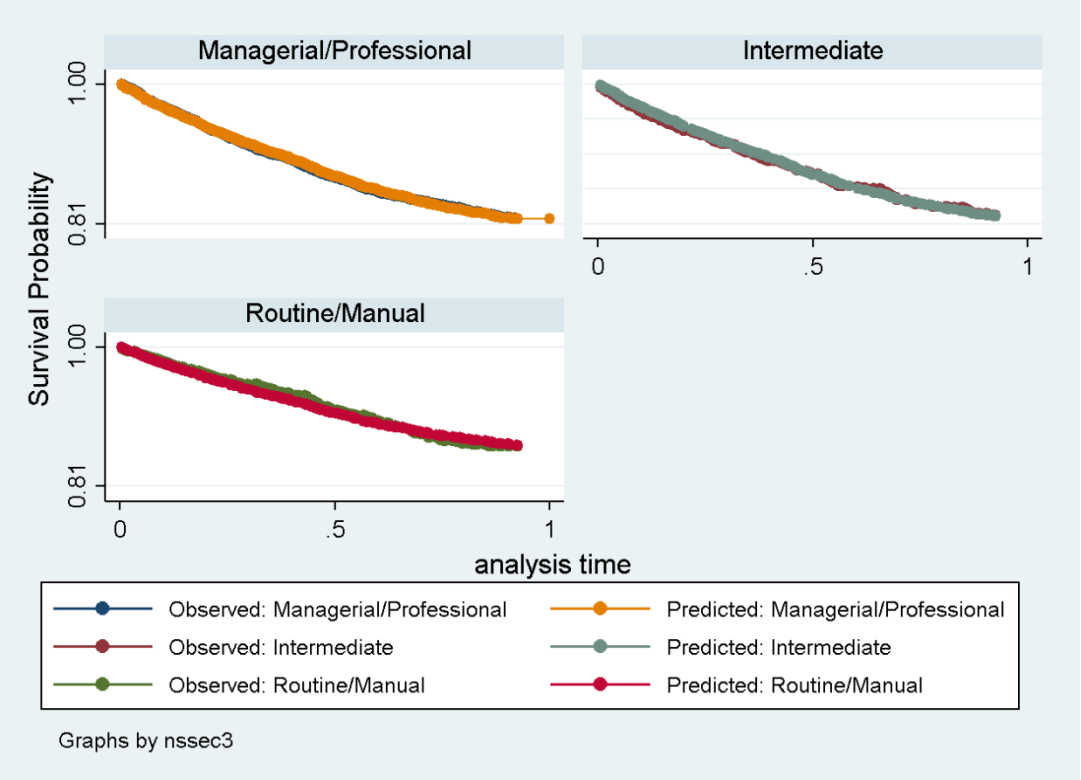


Figure 5.3 Kaplan-Meier plot of time to occurrence of first episode of IID by NS-SEC



Main analysis

All potentially confounding variables were retained in the fully adjusted model. Ethnicity and follow-up type were excluded as these were not considered to be true confounders (Table 5.4). Hazard of IID was significantly lower in routine/manual occupations compared to managerial/professional occupations (HR 0.74, 95% CI 0.61-0.90). No significant interactions were identified.

Table 5.4: Univariate and multivariable Cox regression analysis (n subjects=5716; n failures=845)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.04	(8.87-1.23)	1.03	(0.86-1.23)	0.74
	Routine/manual	0.71	(0.58-0.86)	**0.74**	**(0.61-0.90)**	**0.002**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.17	(1.01-1.36)	1.13	(0.98-1.31)	0.09
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.78	(0.67-0.91)	1.00	(0.82-1.22)	1.00

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval

[a] Adjusted for all other covariates in the model

Baseline hazard stratified by age group and sex

Missing data: NS-SEC was not classifiable for 1,112 individuals. Employment status was missing for five individuals. Rural-urban classification was missing for three individuals.

Sensitivity analyses

Multiple sensitivity analyses were conducted to assess the robustness of the findings; in particular accounting for multiple spells of follow-up, to explore the potential impact of using a less sensitive case definition, excluding individuals with non-classifiable NS-SEC; using multiple imputation to include these individuals, stratifying by age and using an area-level measure of SES (IMD). These analyses in the main did not alter the overall conclusions of this research; lower hazard in routine/manual occupations compared to managerial/professional occupations was a consistent finding across the multiple sensitivity analyses that were conducted to assess the robustness of the main results (Tables 5.5-5.8, Appendix 3.5); including possible cases (HR 0.76, 95% CI 0.64-0.89); including participants who were not-classifiable by NS-SEC (HR 0.76, 95% CI 0.65-0.89); using multiple imputation to assign NS-SEC categories for those who were not-classifiable (HR 0.78, 95% CI 0.64-0.94) and using 10-year age groupings (HR 0.73, 95% CI 0.60-0.89) and were similar to the hazard presented in the main analysis.

In the adjusted age-stratified models (Tables 5.9-5.11), the Hazard Ratio for routine/manual occupations compared to managerial/professional occupations tended to decrease with increasing age (65 and over: 0.60, 95% CI 0.40-0.89 compared to 0-17 years: 0.89 (95% CI 0.61-1.29) although these differences were non-significant. There was a non-significant lower hazard in routine/manual occupations compared to professional/managerial occupations for participants aged less than 18 years, using head of household occupation as proxy for SES (HR 0.89, 95% CI 0.61-1.29). Among participants aged 18-64 years and participants aged 65 and over, those with routine/manual occupations had significantly lower rates of IID compared to professional/managerial occupations (HR 0.76, 95% CI 0.58-0.99 and HR 0.60, 95% CI 0.40-0.89 respectively). Sensitivity analysis including ethnicity in the model can be found in Table 5.12 and the inclusion of ethnicity does not alter the results of the main analysis. No significant interactions were identified in any of the sensitivity analyses.

Using the area-level IMD as a measure of SES, the most disadvantaged (IMD quintile 5) had lower incidence compared to the least disadvantaged (IMD quintile 1) (171.9/1,000 person-years, 95% CI 132.6-222.8; 234.1, 95% CI 206.9-264.8) (Table

5.13, Figure 5.4) in accordance with the main analysis results. No statistically significant relationship between hazard of IID and SES was identified in the adjusted analysis (Table 5.14). The distribution of SES by IMD differed compared to the general population, with those in the most disadvantaged quintile being underrepresented (7% versus 20%) and in the least disadvantaged quintile (24% versus 20%) compared to the distribution in the general population (Tam et al., 2011a).

Table 5.5: Sensitivity analysis – all cases (including possible cases) (n subjects=5716; n failures=1152)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.02	(0.88-1.19)	1.02	(0.88-1.19)	0.77
	Routine/manual	0.74	(0.63-0.87)	**0.76**	**(0.64-0.89)**	**0.001**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.07	(0.94-1.21)	1.03	(0.91-1.17)	0.66
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.71	(0.62-0.81)	0.91	(0.76-1.07)	0.26

NS-SEC: National Statistics-Socioeconomic Classification; CI: confidence interval
a Adjusted for all other covariates in the model
Missing data: Employment status was missing for five individuals. Rural-urban classification was missing for three individuals.
Baseline hazard stratified by age group and sex

Table 5.6: Sensitivity analysis – Including individuals not classifiable by NS-SEC (n subjects=6803; n failures=1355)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.02	(0.88-1.19)	1.03	(0.88-1.19)	0.75
	Routine/manual	0.74	(0.63-0.87)	**0.76**	**(0.65-0.89)**	**0.001**
	Not classifiable*	0.83	(0.71-0.97)	0.90	(0.77-1.06)	0.20
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.08	(0.96-1.21)	1.05	(0.93 -1.18)	0.43
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.69	(0.62-0.78)	0.88	(0.76-1.02)	0.10

NS-SEC: National Statistics- Socioeconomic Classification; CI: confidence interval
[a] Adjusted for all other covariates in the model
* Not classifiable for other reasons.
Baseline hazard stratified by age group and sex
Missing data: Employment status was missing for 30 individuals. Rural-urban classification was missing for three individuals.

Table 5.7: Sensitivity analysis – Multiple Imputation of NS-SEC not classifiable group (n subjects=6803)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.02	(0.86-1.21)	1.01	(0.86-1.20)	0.87
	Routine/manual	0.74	(0.61-0.90)	**0.78**	**(0.64-0.94)**	**0.01**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.18	(1.03-1.35)	1.14	(1.00-1.31)	0.05
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.74	(0.65-0.85)	0.94	(0.79-1.11)	0.48

NS-SEC: National Statistics- Socioeconomic Classification; CI: confidence interval
[a] Adjusted for all other covariates in the model
Baseline hazard stratified by age group and sex
Missing data: Employment status was missing for 30 individuals. Rural-urban classification was missing for three individuals.

Table 5.8: Sensitivity analysis – ten-year age groupings (n subjects=5716; n failures=845)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.04	(0.87-1.23)	1.03	(0.86-1.23)	0.77
	Routine/manual	0.71	(0.58-0.86)	**0.73**	**(0.60-0.89)**	**0.001**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.17	(1.01-1.36)	**1.17**	**(1.01 -1.36)**	**0.03**
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.78	(0.67-0.91)	1.06	(0.86-1.30)	0.61

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval
a Adjusted for all other covariates in the model
Baseline hazard stratified by age group and sex
Missing data: Employment status was missing for five individuals. Rural-urban classification was missing for three individuals.

Table 5.9: Sensitivity analysis – age stratified <18 years (n subjects=935; n failures=201)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	0.99	(0.67-1.45)	0.99	(0.68-1.46)	0.98
	Routine/manual	0.87	(0.60-1.26)	0.89	(0.61-1.29)	0.54
Sex	Male	1.0 (reference)		1.0 (reference)		
	Female	0.98	(0.74-1.29)	0.97	(0.74-1.28)	0.84
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.20	(0.89-1.62)	1.18	(0.88-1.59)	0.84
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.68	(0.25-1.82)	0.71	(0.26-1.91)	0.50

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval
[a] Adjusted for all other covariates in the model

Table 5.10: Sensitivity analysis – age stratified 18-64 years (n subjects=3310; n failures=463)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.12	(0.89-1.41)	1.09	(0.87-1.38)	0.45
	Routine/manual	0.73	(0.56-0.95)	**0.76**	**(0.58-0.99)**	**0.04**
Sex	Male	1.0 (reference)		1.0 (reference)		
	Female	1.44	(1.17-1.77)	**1.42**	**(1.15-1.74)**	**0.001**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.12	(0.92-1.37)	1.09	(0.90-1.34)	0.38
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.94	(0.74-1.20)	0.97	(0.76-1.24)	0.81

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval
a Adjusted for all other covariates in the model
Missing data: Employment status was missing for two individuals. Rural-urban classification was missing for three individuals.

Table 5.11: Sensitivity analysis – age stratified 65+ years (n subjects=1471; n failures=181)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC[b]	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	0.95	(0.65-1.39)	0.92	(0.63-1.35)	0.67
	Routine/manual	0.58	(0.39-0.87)	**0.60**	**(0.40-0.89)**	**0.012**
Sex	Male	1.0 (reference)		1.0 (reference)		
	Female	1.44	(1.07-1.94)	**1.45**	**(1.08-1.96)**	**0.014**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.22	(0.89-1.68)	1.18	(0.86-1.63)	0.31
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	1.13	(0.78-1.64)	1.12	(0.77-1.63)	0.56

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval
a Adjusted for all other covariates in the model
Missing data: Employment status was missing for three individuals.

Table 5.12: Univariate and multivariable Cox regression analysis including ethnicity (n subjects=5716; n failures=845)

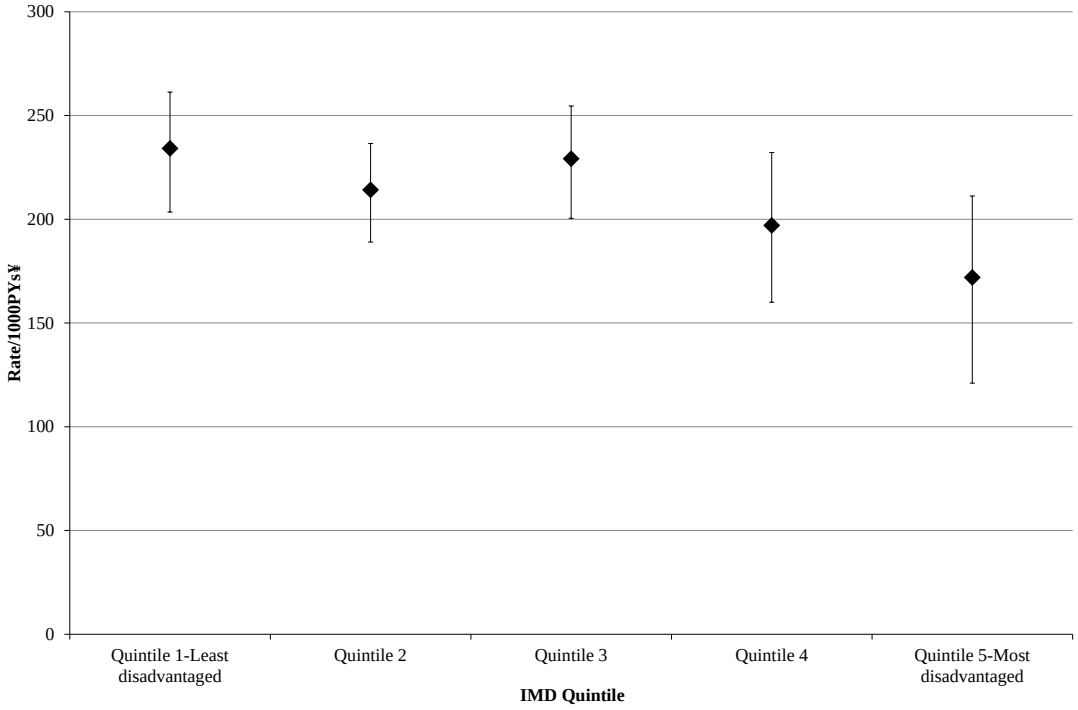
Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.04	(8.87-1.23)	1.03	(0.86-1.23)	0.74
	Routine/manual	0.71	(0.58-0.86)	**0.74**	**(0.61-0.90)**	**0.002**
Ethnicity	White	1.0 (reference)				
	Non-White	0.89	(0.55-1.43)	0.78	(0.48-1.27)	0.32
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.17	(1.01-1.36)	1.13	(0.97-1.31)	0.11
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.78	(0.67-0.91)	1.00	(0.82-1.21)	1.00

NS-SEC: National Statistics- Socioeconomic Classification; CI: confidence interval
[a] Adjusted for all other covariates in the model
Baseline hazard stratified by age group and sex
Missing data: NS-SEC was not classifiable for 1,112 individuals. Employment status was missing for five individuals. Rural-urban classification was missing for three individuals.

Table 5.13: Incidence rate ratio for exposed (IMD Quintile 5 – most disadvantaged) compared to unexposed (IMD Quintile 1– least disadvantaged)

Rate in exposed	Rate in unexposed	Incidence rate difference	Incidence rate ratio	95% CI
171.9	234.1	-62.2	0.75	0.56-0.99

Figure 5.4: Incidence rates per 1,000 person-years by IMD Quintile



¥Person-Years; IMD: Index of Multiple Deprivation

Table 5.14: Sensitivity analysis – Index of Multiple Deprivation (IMD) (n subjects=6803; n failures=991)

Variable	Category	Univariate		Multivariable[a]		p value
		Hazard ratio	(95% CI)	Hazard ratio	(95% CI)	
IMD Quintile[b]	1 (Least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.93	(0.78-1.09)	0.89	(0.75-1.05)	0.17
	3	0.98	(0.83-1.17)	0.98	(0.82-1.17)	0.83
	4	0.85	(0.67-1.07)	0.85	(0.67-1.07)	0.17
	5 (Most disadvantaged)	0.74	(0.56-0.99)	0.76	(0.60-1.02)	0.06
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.18	(1.03-1.35)	1.15	(0.99-1.32)	0.06
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.74	(0.65-0.85)	0.94	(0.80-1.12)	0.50

NS-SEC: National Statistics - Socioeconomic Classification; CI: confidence interval
[a] Adjusted for all other covariates in the model
Baseline hazard stratified by age group and sex
Missing data: Employment status was missing for 30 individuals. Rural-urban classification was missing for three individuals.

5.4 Discussion

In this analysis of a large UK sample following a prospective community cohort to monitor the development of IID symptoms, the relationship between IID and SES using occupation as an individual-level measure of SES was investigated. Lower SES was associated with significantly lower risk of IID in adults but no significant difference in children. There were no significant age-stratified differences in the relationship between IID and SES.

This is a novel analysis of an existing population-based community cohort assessing the association of both individual and area-based measures of SES with IID. Survival analysis explored the relationship between IID and SES accounting for censored observations and different time to event for participants. Multiple sensitivity analyses were conducted to assess the robustness of the main results. A key strength of this study is that it does not require an individual to seek care or have a specimen taken in order to be included in the study, thus reducing potential bias if healthcare interaction differs by SES.

There are some limitations to be acknowledged. A major limitation is that participation in the cohort study was low; only around 9% of the original number recruited and screened for participation, lower than the first IID study (35%) (Tam et al., 2011a). Furthermore, the response varied by SES, with lower response among lower SES groups. Participation bias within cohort studies, particularly by SES, is a recognised limitation. The characteristics of the cohort population differed from the UK population, as those who were most disadvantaged were underrepresented compared to the UK population, whilst those who were advantaged were overrepresented (Office for National Statistics [dataset], 2011), and a large proportion of participants (16.3%, n=1,112) could not be classified by NS-SEC. It is possible that those who agreed to participate had a different risk of IID compared to those who refused, which may limit the generalisability of results. In addition, the lack of a significant difference in risk by SES for children could be related to small numbers in the stratified groups which means the study may lack power for detecting a difference, although the trend was of a lower risk for lower SES participants.

Despite these limitations, this study represents an important analysis of a large prospective community cohort in the UK which suggests differences in risk of IID by

SES amongst the population within this study. To the best of my knowledge, this is the most comprehensive analysis of IID by SES conducted in the UK. This study differs from two earlier analyses of the IID2 data. Tam et al. (2013) used data from the IID2 Study and found no significant difference in risk of multiple-spells of IID in disadvantaged compared to advantaged individuals, while Tam et al. (2011b) found no significant difference in incidence by socioeconomic groups. The different findings between these papers could relate to differences in research questions which were answered using different and question-specific methods, as well as differences in the outcome; as the outcome for my study was time to event and used survival analysis to account for differing follow-up times.

Despite potential issues with participation bias by SES, cohort studies are generally a more robust method of assessing individual-level exposures although few population-based cohort studies have been conducted in high income countries to investigate differences in IID risk by SES. Studies investigating this relationship between age groups are particularly limited.

In a Dutch cohort study, individuals with a low level of education had significantly lower odds of gastroenteritis compared to those with a high level of education (OR 0.65, 95% CI 0.56-0.75) (de Wit et al., 2001a); comparable with our adjusted estimate. Another cohort study (Simonsen et al., 2008), in Denmark, which looked at specific bacterial pathogens as opposed to IID, found an increased risk in adults in higher SES groups for most pathogens (*Campylobacter, Salmonella* Enteritidis and *Shigella*) although the pattern was less clear in children, with no association between risk and SES for most pathogens; these findings also concur with our results.

By contrast, a Canadian cohort study (Teschke et al., 2010) found that individuals in neighbourhoods with low and medium household incomes had higher rates of IID compared to those living in neighbourhoods with high household incomes. In contrast to the other cohort studies above, including our study, Teschke et al. (2010) used physician visits rather than self-reported symptoms to define IID; when hospitalisation was used to define IID, the authors found no significant difference in rates by SES. Further, the Teschke et al. (2010) study was designed to assess the association between environmental factors and IID incidence rather than SES specifically.

Cohort studies which have focussed on children have found higher risk in more disadvantaged groups (Baker et al., 1998, Beale et al., 2010, Eaton-Evans and Dugdale, 1987, Ludvigsson, 2006), in contrast with our findings. Two of these studies (Baker et al., 1998, Beale et al., 2010) were from the same survey, although used different SES measures to investigate the relationship, and specifically sampled very young children. Studies assessing SES specifically in children may be better powered or designed to investigate this relationship than studies looking at all ages.

Many studies assessing the relationship between IID and SES in high income countries have used study designs other than population-based cohorts, such as cross-sectional population surveys, which have produced conflicting results. Some support our finding that lower SES is associated with lower risk of IID (Fein et al., 1995, Herikstad et al., 2002, Majowicz et al., 2007, Pollard et al., 2014, Scallan et al., 2004, Van Cauteren et al., 2012). These studies looked at adults specifically or all ages combined and used mainly education as a measure of SES, with the exception of one study which used occupation (Scallan et al., 2004). Most cross-sectional population surveys found no significant association (Adlam et al., 2011, Doorduyn et al., 2012, Evans et al., 2006, Hall et al., 2006, Herikstad et al., 2002, Jones et al., 2007, Majowicz et al., 2004, Majowicz et al., 2007, McAteer et al., 2011, Sargeant et al., 2008, Van Cauteren et al., 2012, Wilking et al., 2013), including three studies which found significant associations with education but not with income and occupation (Herikstad et al., 2002, Majowicz et al., 2007, Van Cauteren et al., 2012), suggesting that the association may vary with different measures of SES. The variability in these results also suggests that cross-sectional study designs may not provide the most robust estimates of the relationship between SES and IID.

Studies which have used hospitalisation as their outcome have found higher rates amongst those of a lower SES (Baker et al., 2012, Olowokure et al., 1999, Pockett et al., 2011). While our findings suggest that the risk of IID is lower in more disadvantaged population groups, the consequences of having IID or the likelihood of medical referral if disadvantaged may actually be greater.

There are several possible explanations for the finding of lower IID rates amongst individuals of lower SES. It may be artefactual and related to low response rate. The underrepresentation of disadvantaged individuals and the overrepresentation of

advantaged individuals, or differential reporting by SES, may have resulted in a biased population. Conversely, differences in the recognition or reporting of symptoms by SES or by healthcare interaction may partially explain the results. It may also represent a real lower risk of IID amongst those who are disadvantaged through differences in exposures by SES (such as the consumption of less risky foods, reduced opportunity to eat meals outside of the home, reduced exposure to animal attractions, such as open farms, and reduced levels of foreign travel amongst those of a lower SES) (Pollard et al., 2014, Scallan et al., 2006).

There is some evidence from this study and others to suggest the existence of a relationship between IID and SES, with lower SES associated with lower rates of IID. Evidence from the literature suggests that the consequences of IID are more severe for more disadvantaged population groups, with higher hospital admission rates for those of lower SES (Baker et al., 2012, Olowokure et al., 1999, Pockett et al., 2011), and that disadvantaged children may be at higher risk of IID infections (Baker et al., 1998, Beale et al., 2010, Eaton-Evans and Dugdale, 1987, Ludvigsson, 2006). Our results may underestimate the risk in disadvantaged groups and in children. Whilst more disadvantaged individuals may be at a lower risk of, or vulnerability to, GI infections, the possibility of more severe consequences amongst these groups has implications for the clinical management of IID and for healthcare utilisation.

Further research is required to explore the role of symptom recognition, perception, healthcare interaction and other potentially mediating exposures to complement these results and help to explain the relationship between SES and GI infection. Focussing on children may clarify the inconsistent results seen across the literature, as would further research on the most appropriate SES measure to use to produce the most robust estimates of the association between IID and SES. Finally, a greater understanding of the individual behaviours and environmental risk factors by SES is crucial to understanding differential risk, vulnerability, and consequences of IID. These results contribute to the evidence on community-level risk of and vulnerability to GI infections. Alongside the other analyses presented within this thesis and in conjunction with future planned qualitative and mixed-methods approaches, this could ultimately be used to provide evidence to inform policies to address inequalities in risk, vulnerability and consequences of IID.

5.5 Interpretation

In this chapter I have demonstrated evidence to suggest that there is differential risk of IID symptoms in the community in the UK, with lower risk of IID symptoms amongst those in the community who are more disadvantaged. This may appear counterintuitive, with much evidence to suggest that other diseases, infectious and non-infectious, are associated with higher risk amongst more disadvantaged individuals (Biering-Sørensen et al., 2012, Graham, 2009, Hughes and Gorton, 2015) and greater consequences of infection (Baker et al., 2012, Olowokure et al., 1999, Phillips et al., 2011, Pockett et al., 2011, Rose et al., 2017) and given the findings from Chapter 4 which suggested that more disadvantaged children experience a greater burden of GI infections. As described above, this finding could be artefactual and related to underrepresentation of those of a lower SES. This hypothesis will be tested further through the analysis of a large dataset of calls to telephone-based health advice in Chapter 6, which may not suffer to such an extent from low response by lower SES groups.

In theory, this finding could also be due to differential symptom recognition or healthcare-seeking behaviour. This hypothesis will be tested further in Chapters 6 and 7 through the analysis of calls to telephone-based health advice services and through the analysis of social patterning of risk factors and exposures for a severe GI infection.

Finally, this finding could suggest that those who are more disadvantaged do have a lower risk of GI infections. This may be related to social patterning of risk factors and exposures which could influence vulnerability to GI infections. This hypothesis will be further tested in Chapter 7.

Chapter 6 – Study 3 (Objective 3)
Social patterning of telephone advice for diarrhoea and vomiting in the community: analysis of 24 million calls to NHS Direct/NHS 111 in England

Abstract

Background

Gastrointestinal (GI) infections, leading to diarrhoea and vomiting as well as more serious health problems, are common, however many of these infections are hidden with an estimated 147 cases in the community for every one case recorded in national surveillance systems. There is evidence to suggest that the consequences of these infections vary by socioeconomic status (SES), but the impact of SES on the risk of GI infections is less well-understood. This study uses data from calls to the national telephone helplines for health advice (NHS 111 and NHS Direct) in England to assess the association between SES and callers with symptoms suggestive of GI infections.

Methods

Data from over 24 million calls to NHS Direct/NHS 111 in England were extracted from PHE syndromic surveillance systems. Calls about 'diarrhoea' and/or 'vomiting' were collectively defined as GI calls; calls made about other symptoms were defined as non-GI calls. Call data were linked to area-level SES and population by age and sex. The relationship between SES and GI calls compared to non-GI calls was assessed using a generalised linear model (GLM), adjusting for potentially confounding variables.

Results

A total of 24,214,879 calls were included in the study (NHS Direct n=7,874,257; NHS 111 n=16,340,622), of which 6% (n=1,450,843) were classed as GI calls. Risk of calling for GI symptoms was significantly higher in the most disadvantaged compared to the least disadvantaged areas in both NHS Direct and NHS 111 for all ages (NHS 111; 0-4 years RR 1.27, 95% CI 1.25-1.29; 5-9 years RR 1.43, 95% CI 1.36-1.51; 10-14 years RR 1.36, 95% CI 1.26-1.41; 15-19 years RR 1.59, 95% CI 1.52-1.67; 20-59 years RR 1.50, 95% CI 1.47-1.53, 60 years and over (RR 1.12, 95% CI 1.09-1.14).

Conclusion

In this exceptionally large sample of calls (over 24 million) made to the NHS telephone helplines for health advice, lower SES was associated with higher risk of GI calls. This may be related to greater exposure or vulnerability to GI infections for more disadvantaged groups.

6.1 Introduction

This study was designed to address gaps in the literature identified in Study 1, and to further investigate socioeconomic inequalities in GI infections in the community. Due to the low and biased response rate identified in Study 2 of this thesis, Study 3 uses data from two national telephone helplines for health advice; the lowest level of healthcare interaction available, which brings us as close to the true community incidence as we can get from routinely collected data. The objectives of the research within this thesis are detailed in Chapter 1. This chapter seeks to meet Objective 3: To analyse the extent of, and mechanisms underlying, socioeconomic inequalities in risk of GI infections in the community, with estimates derived from routine data on members of the public seeking telephone-based healthcare advice in England.

Gastrointestinal (GI) infections are common, leading to diarrhoea and vomiting as well as more serious health problems. Previous estimates suggest that around 25% of people in the UK suffer an episode of infectious intestinal disease (IID) per year and that foodborne illness in England and Wales costs around £1.5 billion annually (Tam et al., 2011a). Many infections are known to vary by social group but the role of socioeconomic inequalities in risk of GI infection in high income countries, such as in the UK, is not well understood, with studies presenting conflicting findings (Newman et al., 2015). Many individuals do not present to healthcare as most GI infections are self-limiting; it is estimated that there are 147 cases in the community for every one case that is reported to national surveillance, such as via laboratory reports (Tam et al., 2011a). This level of underreporting presents a challenge to understanding the relationship between infection and socioeconomic status (SES) due to the potential bias in healthcare interaction within certain groups of society. It is therefore important to attempt to capture potential inequalities in GI infections particularly amongst individuals who would not be captured in formal surveillance systems either within the community, as in the previous chapter, or in those accessing healthcare advice through telephone based systems. Telephone helplines are underutilised for GI surveillance and potentially give a closer reflection of true community incidence than other routine measures.

This study aims to investigate the relationship between SES and calls to the national telephone helplines for health advice with symptoms of diarrhoea and vomiting; defined as GI calls. This study will contribute to the understanding of socioeconomic

and socio-demographic inequalities of GI infections in the UK. In this chapter I will present the results from the analysis of this observational study, including details of sensitivity and robustness analyses, the implications of the findings for future work and the impact of the results on our understanding for public health.

6.2 Methods

The methods for this analysis are found in greater detail in Chapter 3. A brief overview is provided below.

Design, setting and data source

An observational study design was used to assess socioeconomic inequalities in calls about 'diarrhoea' and/or 'vomiting', which were collectively defined as GI calls with calls made about other symptoms, which were defined as non-GI calls. An analysis of calls made to the two national telephone helplines for health advice, NHS Direct and its successor NHS111, was undertaken to explore the role of socioeconomic status on reporting of GI symptoms. These systems provide advice delivered to individuals over the telephone as opposed to face-to-face consultation. Both telephone helpline systems are described in greater detail in Chapter 3. Data from NHS Direct and 111 were extracted from the HPA/PHE NHS Direct/111 syndromic surveillance systems, based upon anonymised data routinely collected and used by PHE for routine public health surveillance from October 2010 to July 2015. For comparability to NHS 111, the NHS Direct dataset was restricted to calls from England only. Due to the changeover between systems, no data were extracted in August or September 2013 to allow for potential drop-off and uptake of reporting across the two systems.

Participants

All calls made to either NHS Direct (October 2010 to July 2013) or NHS 111 (October 2013 to July 2015) with a valid postcode district, the first part of a postcode, in England and reported to the HPA/PHE syndromic surveillance systems were included. Reason for call was coded as diarrhoea/vomiting (GI calls) or non-diarrhoea/vomiting (non-GI calls). Data on the number of calls were aggregated by postcode district, age group and gender. Population by age group and gender for each

postcode district were merged with the call data to allow for population-level comparisons.

Outcome and covariates

The primary outcome of interest for this study was rate of calls for which diarrhoea and/or vomiting were the symptoms recorded, per 10,000 population. The comparator group was identified by grouping all calls without diarrhoea or vomiting as the symptoms (non-GI calls).

The primary exposure of interest was SES, determined using the area-level measure of SES, IMD (Department for Communities and Local Government, 2011) generated using the population-weighted mean IMD score for each postcode district which was assigned to each call. IMD quintiles were generated as detailed in Chapter 3. The Office for National Statistics Rural Urban Classification (Department for Environment, Food and Rural Affairs, 2016) was used to assign the proportion of the population classified as urban for each postcode district.

Other covariates of interest included in the analysis were age (coded as 0-4, 5-9, 10-15, 16-19, 20-59 and 60 years and over); sex (male/female) and urban decile (proportion of population classed as urban, operationalised as deciles).

Analysis strategy

Analyses were conducted in R (version 3.3.1). A descriptive comparison of GI call rate and non-GI call rate by SES was undertaken. Crude incidence rates, incidence differences and incidence ratios by SES were calculated, stratified by gender and age group.

The main analysis explored the relationship between GI calls and SES using a GLM with a Poisson family and log-link function. To model the call rate, the log of the population in each postcode district, age group and gender was included in the model as an offset. Due to some age groups within postcode districts having a population of zero, these were excluded from the main analysis (n=1,357, 0.1%). Separate analyses were undertaken for NHS Direct and NHS 111 as it was not justified to pool the results due to differences in rates between two systems as a result of NHS 111 also acting as an out of hours GP service which increased call rates. The multivariable model described above was then fitted with SES (IMD quintile) as the

exposure variable and calls as the outcome variable, adjusting for the potential confounders (age group, sex and urban decile and interactions between age and sex, and age and IMD quintile). As the previous literature and study 1 of this thesis suggested that the relationship between SES and GI risk may vary across the life course, an interaction term between IMD quintile and age group was included. Risk ratios and 95% confidence intervals were estimated.

Sensitivity analyses

Firstly, to assess whether the exclusion of postcode districts with no population affected the results, postcode districts with a population of zero were recoded to one and therefore included in the analysis, again stratified by system. Secondly, to test whether there was a significant trend across levels of deprivation and rurality, the analysis was repeated using IMD Score and the proportion of the population classed as urban as continuous variables. Thirdly, due to changing protocols in NHS Direct which meant symptom information was unavailable for infants <1 year of age after November 2011, sensitivity analysis excluding calls regarding infants <1 was conducted.

Additionally, to explore any change in the trend of the call rate particularly in relation to the introduction of the rotavirus vaccine (July 2013) and NHS 111 (October 2013), incidence rates and incidence rate ratios by SES with associated 95% confidence intervals were calculated by year for 0-4 year olds (assumed to be the vaccine-eligible cohort) compared to GI calls in all other ages, non-GI calls in 0-4 year olds and non-GI calls in all other ages.

6.3 Results

Characteristics of participants

A total of 24,214,879 calls were included in the study (NHS Direct n=7,874,257; NHS 111 n=16,340,622). Of these, 6% (n=1,450,843) were classed as GI calls. Age was missing for 431,239 records (1.8%); sex for 314,982 records (1.3%) and; whether the caller reported diarrhoea was unknown for 2 records. After excluding records with missing data, 23,762,217 calls remained.

Incidence rate ratios for GI calls were significantly higher among the most disadvantaged compared to the least disadvantaged (Table 6.1); this pattern was also reflected in all the age- and sex-adjusted incidence rates for the larger NHS 111 dataset, and for both men and woman and for adults but not for children, in the NHS Direct dataset. Non-GI calls were significantly higher compared to the least disadvantaged in both systems, although in NHS 111 the incidence rate ratio was lower in comparison to the trend in GI calls.

Table 6.1: Incidence rate ratio for exposed (most disadvantaged) compared to unexposed (least disadvantaged)

	Rate/10,000 person-months in most disadvantaged	Rate/10,000 person-months in least disadvantaged	Incidence rate difference	Incidence rate ratio	95% CI
		NHS Direct			
Overall GI calls	2.9	2.7	0.2	1.07	1.05-1.08
Female	3.3	3.0	0.3	1.10	1.08-1.12
Male	2.5	2.4	0.1	1.03	1.01-1.05
0-4	14.1	17.7	-3.7	0.79	0.78-0.81
5-9	2.0	2.1	-0.1	0.94	0.88-1.00
10-14	0.9	0.9	0.0	1.02	0.94-1.12
15-19	2.3	2.0	0.3	1.17	1.10-1.24
20-59	2.1	1.9	0.2	1.11	1.09-1.13
60+	2.0	1.8	0.2	1.10	1.06-1.14
Overall non-GI calls	43.3	36.1	7.2	1.20	1.19-1.20
Female	50.7	41.3	9.4	1.23	1.22-1.23
Male	35.7	30.7	5.1	1.17	1.16-1.17
0-4	112.4	146.3	-33.9	0.77	0.76-0.77
5-9	23.7	26.7	-3.1	0.89	0.87-0.90
10-14	16.7	17.5	-0.8	0.95	0.93-0.98
15-19	46.0	34.1	11.9	1.35	1.33-1.37
20-59	42.7	32.4	10.4	1.32	1.31-1.33
60+	29.7	25.7	4.0	1.16	1.15-1.17
		NHS 111			
Overall GI calls	8.1	5.4	2.7	1.50	1.49-1.52
Female	9.3	6.2	3.1	1.49	1.47-1.51
Male	6.9	4.5	2.4	1.52	1.50-1.55
0-4	46.7	34.9	11.7	1.34	1.31-1.36
5-9	5.6	3.7	1.9	1.51	1.44-1.59
10-14	2.4	1.7	0.7	1.43	1.33-1.54
15-19	6.3	3.7	2.6	1.70	1.52-1.78
20-59	4.5	2.9	1.7	1.57	1.54-1.60
60+	6.4	5.4	1.0	1.18	1.16-1.21
Overall non-GI calls	125.0	118.3	6.7	1.06	1.05-1.06
Female	143.0	133.6	9.4	1.07	1.07-1.07
Male	106.7	102.4	4.3	1.04	1.04-1.05
0-4	289.9	381.8	-92.0	0.76	0.76-0.76
5-9	76.0	91.5	-15.5	0.83	0.82-0.84
10-14	49.1	55.0	-5.9	0.89	0.88-0.91
15-19	131.2	104.5	26.6	1.25	1.24-1.27
20-59	113.3	91.5	21.8	1.24	1.23-1.24
60+	133.9	137.7	-3.8	0.97	0.97-0.98

Main analysis

Data were aggregated to postcode district, age group and gender for the main analysis. The aggregated data used in the regression analysis consisted of 49,970 postcode district, age and gender groups. As there was a significant interaction between age group and IMD quintile (Appendix 4.5 and 4.6), Table 6.2 presents the risk ratio for IMD in each age group for NHS Direct derived from the interaction terms, adjusting for sex and proportion of the population classed as urban (Table 6.2); there was a statistically significant lower risk of calling with GI symptoms amongst the most disadvantaged compared to the least disadvantaged children under 10 years of age but there was no significant difference in age for adults. The full model with the separate parameters is presented in Appendix 4.5.

In Table 6.3 the risk ratio for IMD in each age group for NHS Direct derived from the interaction terms, adjusting for sex and proportion of the population classed as urban is presented (Table 6.3); there was a statistically significant higher risk of calling with GI symptoms amongst the most disadvantaged compared to the least disadvantaged. The trend across quintiles was clearer in NHS 111, with the risk statistically significantly higher in the most disadvantaged compared to the least disadvantaged in all age-groups. The full model with the separate parameters is presented in Appendix 4.6.

Table 6.2: Univariate and multivariable regression analysis presenting main effect with interaction terms for GI calls in each age group – NHS Direct (n=24,985)

Age group	IMD Quintile	Univariate risk ratio (95% CI)	Multivariable risk ratio[a] (95% CI)	p value
0-4	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	0.98 (0.96-1.00)	0.98 (0.96-1.00)	0.05
	3	0.97 (0.95-0.99)	0.93 (0.92-0.95)	<0.001
	4	0.93 (0.91-0.95)	0.86 (0.84-0.87)	<0.001
	5 (Most disadvantaged)	0.79 (0.78-0.81)	0.72 (0.71-0.74)	<0.001
5-9	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	0.99 (0.93-1.04)	0.98 (0.93-1.04)	1.00
	3	1.06 (1.00-1.12)	1.01 (0.96-1.07)	1.00
	4	1.04 (0.99-1.10)	0.96 (0.91-1.01)	1.00
	5 (Most disadvantaged)	0.94 (0.88-1.00)	0.85 (0.80-0.90)	<0.001
10-14	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.09 (0.99-1.18)	1.09 (1.00-1.18)	0.05
	3	1.18 (1.07-1.26)	1.11 (1.25-1.32)	0.01
	4	1.17 (1.08-1.27)	1.08 (0.99-1.17)	0.10
	5 (Most disadvantaged)	1.02 (0.94-1.12)	0.92 (0.85-1.01)	0.10
15-19	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.07 (1.01-1.13)	1.07 (1.01-1.13)	0.05
	3	1.17 (1.11-1.24)	1.12 (1.06-1.18)	<0.001
	4	1.22 (1.15-1.29)	1.12 (1.06-1.18)	<0.001
	5 (Most disadvantaged)	1.17 (1.10-1.24)	1.05 (0.99-1.11)	1.00
20-59	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.03 (1.00-1.05)	1.02 (1.00-1.04)	0.05
	3	1.06 (1.04-1.08)	1.02 (1.00-1.04)	0.10
	4	1.13 (1.11-1.15)	1.04 (1.02-1.06)	<0.001
	5 (Most disadvantaged)	1.11 (1.09-1.13)	1.01 (0.99-1.03)	1.00
60+	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	0.99 (0.96-1.02)	1.00 (0.97-1.03)	1.00
	3	1.06 (1.03-1.09)	1.03 (1.00-1.06)	0.10
	4	1.12 (1.09-1.15)	1.03 (1.00-1.06)	0.10
	5 (Most disadvantaged)	1.10 (1.06-1.14)	0.99 (0.96-1.03)	1.00

[a]Linear combination of main effect + interaction between age and IMD quintile, adjusted for sex and % urban; [b]Reference age category

Table 6.3: Univariate and multivariable regression analysis presenting main effect with interaction terms for GI calls in each age group – NHS 111 (n=24,985)

Age group	IMD Quintile	Univariate Risk ratio (95% CI)	Multivariable Risk ratio[a] (95% CI)	p value
0-4	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.20 (1.19-1.22)	1.20 (1.18-1.22)	<0.001
	3	1.40 (1.38-1.42)	1.37 (1.35-1.39)	<0.001
	4	1.40 (1.37-1.42)	1.34 (1.31-1.36)	<0.001
	5 (Most disadvantaged)	1.34 (1.31-1.36)	1.27 (1.25-1.29)	<0.001
5-9	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.17 (1.11-1.22)	1.16 (1.11-1.22)	<0.001
	3	1.41 (1.35-1.48)	1.38 (1.32-1.45)	<0.001
	4	1.51 (1.44-1.58)	1.44 (1.37-1.51)	<0.001
	5 (Most disadvantaged)	1.51 (1.44-1.59)	1.43 (1.36-1.51)	<0.001
10-14	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.19 (1.11-1.28)	1.19 (1.11-1.28)	<0.001
	3	1.39 (1.29-1.48)	1.36 (1.27-1.46)	<0.001
	4	1.42 (1.33-1.52)	1.36 (1.27-1.46)	<0.001
	5 (Most disadvantaged)	1.43 (1.33-1.54)	1.36 (1.26-1.30)	<0.001
15-19	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.24 (1.18-1.30)	1.24 (1.18-1.30)	<0.001
	3	1.51 (1.44-1.58)	1.47 (1.41-1.54)	<0.001
	4	1.62 (1.55-1.70)	1.54 (1.47-1.61)	<0.001
	5 (Most disadvantaged)	1.70 (1.62-1.78)	1.59 (1.52-1.67)	<0.001
20-59[b]	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.24 (1.22-1.26)	1.23 (1.21-1.26)	<0.001
	3	1.45 (1.43-1.48)	1.42 (1.40-1.45)	<0.001
	4	1.53 (1.50-1.56)	1.46 (1.44-1.49)	<0.001
	5 (Most disadvantaged)	1.57 (1.54-1.60)	1.50 (1.47-1.53)	<0.001
60+	**1** (Least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.13 (1.11-1.15)	1.13 (1.11-1.15)	<0.001
	3	1.28 (1.26-1.31)	1.26 (1.23-1.28)	<0.001
	4	1.25 (1.23-1.28)	1.19 (1.17-1.22)	<0.001
	5 (Most disadvantaged)	1.18 (1.16-1.21)	1.12 (1.09-1.14)	<0.001

[a]Linear combination of main effect + interaction between age and IMD quintile, adjusted for sex and % urban;
[b]Reference age category

Sensitivity analyses

Sensitivity analyses were conducted to assess the robustness of the findings by; recoding postcodes with no population to one in order to include them in the analysis (Appendix 4.7); entering deprivation and rurality as continuous variables into the model to assess whether there were significant trends by these two variables (Appendix 4.8); and comparing results for NHS Direct including and excluding calls about children under 1 year of age (Appendix 4.9a and 4.9b). In NHS Direct there was no significant linear trend in IMD score; in NHS 111, GI calls significantly increased with increasing deprivation. In both NHS Direct and NHS 111, GI calls significantly increased with decreasing rurality. The results of analyses excluding calls regarding infants aged under 1 in NHS Direct were comparable to the results including infants under 1 year of age, therefore it was valid to include records for these infants in the main analysis. These analyses did not alter the overall conclusions of this research; namely that risk of GI calls is significantly higher in disadvantaged areas compared to less disadvantaged areas.

Additional analysis exploring the relationship between SES and GI calls by year to explore whether there was any change in reporting that could be attributed to the introduction of the rotavirus vaccine is presented in Table 6.4. There was a large increase in GI- and non GI-call rates in 2013-2014, coinciding with the introduction of NHS 111. In GI calls for aged under 5 years of age, prior to 2013, call rates were lower in the most disadvantaged areas compared to the least disadvantaged areas but from 2013, call rates were significantly higher in the most disadvantaged areas compared to the least disadvantaged areas. The incidence rate ratio increased substantially from 2012-13, substantially again from 2013-14 and by a lesser, although statistically significant, amount from 2014-15. For GI calls for 5 years and over, the rate is greater in the most disadvantaged compared to the least disadvantaged across the two systems although the incidence rate ratio widened, with a substantial increase from 2013-14 and smaller, but statistically significant, increases from 2012-13 and 2014-15.

For non-GI calls, the rates remain relatively stable across the two systems; in calls for under 5 year olds, the rate is lower in the most disadvantaged compared to the least disadvantaged areas and the ratio stays in the 0.73-0.79 range across the 6 year

period, despite the change in the system and the increase in overall rates. In callers aged 5 years and older, the rate is higher in the most disadvantaged areas compared to the least disadvantaged areas, with the ratio decreasing from 2012-13, but otherwise relatively stable.

Table 6.4: Rates per 10,000 person-months and incidence rate ratio for exposed (most disadvantaged) compared to unexposed (least disadvantaged) by age group

	2010*	**2011**	**2012**	**2013±**	**2014**	**2015±**
			GI calls			
Overall rate 0-4	31.4	23.7	11.7	19.0	51.3	39.2
0-4 Rate/10,000* in most disadvantaged	27.2	19.9	10.4	18.2	51.3	40.3
0-4 Rate/10,000* in least disadvantaged	32.7	26.1	12.4	16.1	38.9	28.9
Incidence rate ratio (95% CI)	0.83 (0.79-0.88)	0.76 (0.74-0.79)	0.84 (0.80-0.87)	1.13 (1.09-1.18)	1.32 (1.29-1.35)	1.40 (1.36-1.44)
Overall rate 5+	2.7	2.3	2.0	2.0	5.3	4.0
5+ Rate/10,000* in most disadvantaged	2.7	2.4	2.1	2.0	5.5	4.3
5+ Rate/10,000* in least disadvantaged	2.4	2.2	1.9	1.7	3.9	3.0
Incidence rate ratio (95% CI)	1.14 (1.08-1.19)	1.11 (1.08-1.14)	1.10 (1.07-1.13)	1.19 (1.15-1.22)	1.38 (1.36-1.41)	1.42 (1.38-1.45)
			Non-GI calls			
Overall rate 0-4	164.1	150.8	148.7	161.1	405.5	291.4
0-4 Rate/10,000* in most disadvantaged	139.2	127.0	127.8	129.1	327.2	237.8
0-4 Rate/10,000* in least disadvantaged	181.1	167.2	164.9	177.1	433.9	301.3
Incidence rate ratio (95% CI)	0.77 (0.75-0.79)	0.76 (0.75-0.77)	0.77 (0.77-0.78)	0.73 (0.72-0.74)	0.75 (0.75-0.76)	0.79 (0.78-0.80)
Overall rate 5+	41.5	40.1	35.2	41.6	127.3	96.7
5+ Rate/10,000* in most disadvantaged	47.0	45.3	40.4	41.7	125.4	97.5
5+ Rate/10,000* in least disadvantaged	36.2	35.3	31.5	39.0	117.3	84.4
Incidence rate ratio (95% CI)	1.30 (1.28-1.31)	1.28 (1.28-1.29)	1.29 (1.28-1.29)	1.07 (1.06-1.08)	1.07 (1.07-1.07)	1.15 (1.15-1.16)

*person-months; ±denotes that the data do not represent a full calendar year

6.4 Discussion

In this analysis of a large representative dataset of calls to the NHS telephone helplines for health advice, the relationship between GI calls and SES was investigated using a population-weighted area-based measure of SES. There was a greater risk of GI calls from more disadvantaged areas compared to less disadvantaged areas. The trend was most clear, increasing across all quintiles, in the NHS 111 dataset; in NHS Direct, there was a significantly lower risk in children but no significant difference in risk for adults. The introduction of NHS 111 greatly increased GI calls, particularly amongst more disadvantaged individuals. Although overall rates also increased for non-GI calls, the relative rate of deprivation did not change. This difference could be related to differences in the way individuals are interacting with NHS 111 compared to NHS Direct – this could be related to NHS 111 being a Freephone number which was not the case for NHS Direct and NHS 111 being a true gateway to unscheduled care as individuals will also use NHS 111 to access out of hours GP services whilst NHS Direct was a standalone health system – although it is unclear why this would have a greater impact upon calls for GI symptoms.

This is a novel use of two existing datasets to assess social patterning of GI calls in England. From a search of the literature and questioning of syndromic surveillance experts, to the best of my knowledge this is one of the largest studies conducted on this topic in England, including over 24 million calls. This dataset is the lowest level of healthcare interaction available, which brings us as close to the true community incidence as we can get from routinely collected data. In addition, there were very low levels of missing data, which is a strength of using these datasets, compared to the low and potentially biased level of coverage in the community cohort. A key strength of this study is that, as for the previous study, it does not require an individual to present to formal healthcare settings nor have a sample taken and as such potentially represents a significant proportion of the GI infections which remain hidden from national surveillance systems. This is important if the decision to seek care is related to SES.

It was possible to assess GI calls in comparison to non-GI calls, which provided a control group. Postcode districts for which no calls were received were included in order to take account of the underlying population at risk, including those who did

not call and as such it was possible to assess whether the pattern of calls by SES reflected the population-level distribution of calls or differed for GI infections. Multiple sensitivity analyses were also possible to check the robustness of the findings and to explore social patterning within GI calls. Despite being observational, this study provides evidence of the existence of socioeconomic inequalities in GI infections. It is also important to note that community-level infection is important for GI infections due to person-to-person transmission and therefore it may actually be more appropriate to consider population-level analyses for this type of infection.

Although this was a large study, it is also possible that residual confounding, bias which remains after confounding has been adjusted for as far as possible, remains. Despite being nationally representative in terms of coverage, it is possible that the two datasets may not be representative of the population in terms of use of telephone helplines by SES. Furthermore, NHS Direct was under-representative of the elderly, who preferred to speak directly to a GP (Cook et al., 2014). Previous research conducted using data from NHS Direct has suggested that demand is highest in areas where deprivation is at or just above the national average (Cooper et al., 2005), and that extreme deprivation appeared to raise adult call rates but reduce call rates in children (Cooper et al., 2005). This could suggest a baseline difference in the demographics of the population interacting with this service but we were able to include postcodes from which no calls originated as well as calculating crude incidence rates to compare callers in the context of the wider population at risk.

In addition, area-level measures of SES, as used here, may not be sensitive enough to detect socioeconomic inequalities particularly where such inequalities are potentially generated by individual factors. As postcode district was the only available geographical measure, and as these may have crossed multiple LSOA boundaries, misclassification of SES is possible; we used population-weighted IMD scores to minimise this issue. A proportion of records for which there was no match to IMD score initially were manually cleaned where it was possible to identify the postcode district due to missing spaces, or letters substituted for numbers. This introduces the possibility of misclassification of IMD or the proportion of the population classed as urban, although this affected only 0.1% (n=14,639) of total calls included and is therefore unlikely to have affected the results. Postcode districts which bordered England and Scotland or Wales may have been misclassified but this affected only a

small number of postcode districts and is unlikely to have caused a socioeconomic bias. The large size of this dataset is also likely to reduce these potential biases.

Due to the introduction of NHS 111 (replacing NHS Direct) and the contemporaneous introduction of the rotavirus vaccine, it was not possible to assess whether any changes in incidence of GI calls by SES was as the result of either intervention. To estimate the incidence rate ratios for 0-4 year olds compared to other ages, we assumed the vaccine eligible cohort to be those aged 0-4 years and that the effect of the vaccine would be minimal in all other age groups, however older non-vaccinated age groups may not be an adequate control group due to indirect protection. Furthermore, in November 2011 NHS Direct changed the assessment protocol for infants aged less than 1 year which meant that symptom information was no longer available for syndromic surveillance and resulted in a drop in reported calls for this age group, which is evident when presenting the rates by year (Table 6.4) and may also explain the different findings in comparison to NHS 111. Sensitivity analysis comparing results for NHS Direct including and excluding <1 year olds demonstrated that this did not have an impact on the results.

Syndromic surveillance systems were used to assess the likely impact of the rotavirus vaccine prior to its introduction (Bawa et al., 2015). Studies have been conducted to explore the overall impact of the rotavirus vaccine (Atchison et al., 2016) however further studies to assess the relationship between SES and potential impacts of the rotavirus vaccine using syndromic surveillance systems, GP and Hospital Episode Statistics (HES) data would be beneficial.

GI symptoms were self-reported which may have resulted in some misclassification of the outcome (diarrhoea and/or vomiting) but the use of clinical decision pathways by the call-handler to assess the presenting symptoms and determine further healthcare needs of the patient is likely to reduce the potential that non-infectious causes of diarrhoea/vomiting were recorded and included in this dataset. It is also possible that the use of the clinical decision pathways might result in the prioritisation of other presenting symptoms such as headache or fever over GI symptoms. Finally, the analyses forming this study are cross-sectional and as such, it is not possible to determine causation. Further, the data are syndromic surveillance data and are therefore not linked to specific pathogens or causes.

Despite these limitations this study represents an important analysis of a large dataset of calls to NHS telephone-based healthcare advice services in England which suggests differences in odds of calling for GI symptoms by SES amongst the population within this study. This is the only study of the relationship between SES and GI infection using telephone helpline data which is the lowest level of healthcare interaction and the nearest to the population incidence for which we can get a high level of case ascertainment.

One study by Cooper et al. (2003) exploring 150,000 GI calls to NHS Direct over a six-month period at three sites found that GI calls accounted for 10.3% of total calls; this proportion was significantly higher among children under 1 year of age (23.5%) and aged 1-4 years (21.5%). This finding is slightly higher than the 6% of calls in our study being classed as GI calls. This study did not explore socioeconomic inequalities in GI calls.

Several studies have explored the social patterning of calls to NHS Direct, but not specifically for GI calls. Cooper et al. (2005) used NHS Direct calls to assess socio-demographic patterning. As found in the study mentioned above, calls were highest in children under five years of age and were higher in women compared to men; with the highest ratio in the 15-44 year age group. They found that the effect of extreme deprivation appeared to raise adult call rates but reduce call rates for children. This is similar to the findings in our study for non-GI calls. This study was conducted on all calls, not specifically for GI calls in two regions of England only and recommends national studies are undertaken to validate the findings. In our study, higher rates of GI calls in more derived compared to less disadvantaged areas were observed overall and for children, which is a novel finding.

Burt et al. (2003) found that there was a significant non-linear relationship between deprivation score and call rates to NHS Direct, with lower rates in the most affluent and the most disadvantaged areas of London. The authors suggest that the decline at the extremes of deprivation scores may reflect barriers to accessing NHS Direct. There is a very high ethnic minority proportion in London, particularly in disadvantaged areas, and this may have impacted on the results if language is a barrier to using telephone-based services.

Shah and Cook (2008) found that NHS Direct use was lower in households with low income (OR 0.67 (95% CI 0.55-0.81); adjusting for limiting illness increased the effect of socioeconomic factors on NHS Direct use. Qualitative studies have also been used to explore the social patterning of callers to NHS Direct. Cook et al. (2014) used focus groups with users and non-users of NHS Direct to explore barriers to use. The authors found that there were a range of barriers including the cost of making a phone call to NHS Direct and that this view was expressed more often by non-users from disadvantaged communities. The NHS 111 system is free to call although the authors highlighted that this change should be clearly communicated to the general public. In our study, we found significantly higher risk of GI calls amongst the most disadvantaged compared to the least disadvantaged for NHS 111, but lower risk in disadvantaged children in NHS Direct dataset. In addition, call volume greatly increased following the introduction of NHS 111; this was particularly evident for GI calls in the most disadvantaged areas which may reflect NHS 111 being a true out of hours service.

There are several possible explanations for the finding of higher odds of calls regarding GI symptoms amongst more disadvantaged individuals in this study. The finding may be artefactual; the study population may not be representative of the general population and may differ from the population not using NHS telephone-based healthcare advice services. Despite this, the sample was large, and the internal associations, which were the targets of inference within the sample population, are likely to be valid. Moreover the inclusion of postcode districts from which no calls were received enabled us to take account of the underlying population at risk. On the other hand, it could also be that more disadvantaged individuals have a genuinely higher risk of GI symptoms compare to less disadvantaged individuals. This may relate to differential exposure, differential vulnerability to disease, or reflect differences in the recognition or reporting of symptoms or differential healthcare seeking behaviour.

In summary, there is evidence from this study and others to suggest the existence of a relationship between GI infections and SES. Amongst people calling NHS telephone-based healthcare advice services, people from more disadvantaged areas were more likely to call for GI symptoms compared to people calling from less disadvantaged areas, and this relationship is stronger than for non-GI calls. This finding has

implications for service providers and the NHS in terms of resource allocation. Further research is required to explore the role of symptom recognition, perception, healthcare interaction and other potentially mediating exposures to complement these results and help to explain the relationship between SES and GI infection. A greater understanding of the individual behaviours and risk factors by SES is crucial to understanding the differential risk, vulnerability, and consequences of GI infections. These results contribute to the evidence on community-level risk of and vulnerability to GI infections amongst individuals seeking care through NHS telephone-based healthcare advice services. Alongside future planned analyses, these results could ultimately be used to provide further evidence to inform policies to address inequalities in risk, vulnerability and consequences of GI infections.

6.5 Interpretation

In this chapter I have demonstrated evidence to suggest differential risk of calling NHS telephone-based healthcare advice lines for GI symptoms in England, with higher risk of calling for GI symptoms amongst people from more disadvantaged areas compared to people from less disadvantaged areas. This finding agrees with the results reported in Chapter 4 which suggested that more disadvantaged children experience a greater burden of GI infections, although disagrees with the results of Chapter 5 which suggested that disadvantaged individuals in the community had a lower hazard of IID symptoms compared to advantaged individuals. The reason for this discrepancy is not yet clear, although it may be that this dataset provides a more reliable assessment of the relationship between SES and GI infections as the IID2 study experienced very low response rates which varied by SES. To accurately estimate the incidence of GI infections, it is important to get as close to the true population incidence as possible therefore telephone helplines represent the nearest community incidence measure for which there is high usage from a representative population. This is likely to estimate incidence more accurately than other datasets, such as HES, which would record more severe infection only and would possibly introduce a larger healthcare seeking bias.

This finding could also suggest differential symptom recognition or healthcare-seeking behaviour. This hypothesis will be tested further in Chapter 7 through the analysis of social patterning of risk factors and exposures for a severe GI infection. Finally, this finding could suggest that those who are more disadvantaged do have a

higher risk of GI infections. This may be related to social patterning of risk factors and exposures which could influence vulnerability to GI infections. This hypothesis will also be further tested in Chapter 7.

Chapter 7 – Study 4 (Objective 4)
Social patterning of clinical outcomes, healthcare contact, risk factors and development of severe complications in a diagnosed GI infection

Abstract

Background

Shiga toxin-producing *Escherichia coli* (STEC) infection, such as that caused by *E. coli* O157, is a relatively rare but potentially serious cause of gastrointestinal illness in England. Symptoms can range from mild gastroenteritis to severe bloody diarrhoea and can cause the potentially fatal Haemolytic Uraemic Syndrome (HUS), the leading cause of acute renal failure in children in the UK. Many infections are socially patterned; but the role of socioeconomic status (SES) in differential clinical outcomes, healthcare contact and exposure to potential risk factors; and the role of socioeconomic conditions (SECs) in the risk of subsequent HUS development have not been explored in England.

Methods

An observational study using data collected on all primary, symptomatic, STEC cases identified in the PHE National Enhanced Surveillance System for STEC (NESSS) from January 2010 to December 2015 was undertaken. Multivariable logistic regression was used to assess the relationship between SES, clinical factors, healthcare contact and exposure to known risk factors for STEC. A separate retrospective cohort of paediatric HUS cases, generated by an active surveillance system, was analysed to estimate the odds of progression to HUS on the basis of socio-demographic risk factors.

Results

The odds of a case of STEC infection visiting A&E and hospitalisation were significantly higher among disadvantaged, compared to less disadvantaged cases (OR 1.35, 95% CI 1.10-1.75 and OR 1.71, 95% CI 1.36-2.15 respectively). Odds of exposure to known risk factors for STEC were significantly lower among more disadvantaged STEC cases; foreign travel (OR 0.57, 95% CI 0.44-0.73), fish/shellfish (OR 0.68, 95% CI 0.52-0.87), salad, fruit, vegetables or herbs (OR 0.63, 95% CI 0.46-0.86), fresh water (OR 0.65, 95% CI 0.44-0.95), walking in a paddock (OR 0.55, 95% CI 0.37-0.82), day trip (OR 0.64, 95% CI 0.46-0.89), contact with soil (OR 0.66, 95% CI 0.47-0.92), and UK travel (OR 0.54, 95% CI 0.39-0.74).

Odds of progression to HUS among paediatric STEC cases was non-significantly lower among more disadvantaged children, compared to less disadvantaged children (OR 0.57, 95% CI 0.25-1.31) in the fully adjusted model.

Conclusion

Social patterning of clinical outcomes, healthcare contact and some risk factors were observed, with higher odds of reporting A&E and hospitalisation in more disadvantaged individuals and lower odds of exposure to known risk factors for STEC infection. Non-significant lower HUS development was identified among more disadvantaged children. Further studies are required corroborate the findings of this study to better identify and tackle socioeconomically driven inequalities in GI infections.

7.1 Introduction

This study was designed to address gaps in the literature identified in the literature review (Chapter 2), and to further investigate potential differences in exposure or healthcare interaction which may partially explain the findings of Studies 1-3 of this thesis. The objectives of the research within this thesis are detailed in Chapter 1. This chapter seeks to meet Objective 4: To explore the social patterning of clinical outcomes, healthcare contact and risk factors for a laboratory-confirmed, potentially severe, GI infection (STEC) and socio-demographic inequalities in risk of development of a serious sequela (HUS) in order to suggest hypotheses for testing in future studies which will lead to a better understanding of the mechanisms leading to socioeconomic inequalities and may identify important links in the causal chain which could be addressed more effectively.

Shiga toxin-producing *Escherichia coli* (STEC; also known as vero cytotoxin-producing *E. coli* (VTEC)) are a group of bacteria that cause infectious gastroenteritis, with STEC O157 being the most frequently reported strain to cause illness in England (Byrne et al., 2015). Symptoms can range from mild gastroenteritis (David et al., 2004) through to severe bloody diarrhoea (Ackers et al., 1998, Sutcliffe et al., 2004). On rare occasions, STEC infection can cause the serious condition of haemolytic uraemic syndrome (HUS) (Payne et al., 2003, Rangel et al., 2005, Tarr et al., 2005), affecting the blood, kidneys and, in the most severe cases, the central nervous system. Children and the elderly are most susceptible to severe illness and HUS is recognised as the most common cause of acute renal failure among children in the UK (Lynn et al., 2005). A previous study conducted by the British Paediatric Surveillance Unit (BPSU), in conjunction with PHE (formerly the HPA), identified 413 cases of HUS between 1997 and 2001, 330 of which were STEC related (Lynn et al., 2005). It is estimated that 5-8% of individuals with STEC infection will progress to HUS (Tarr et al., 2005) however, it is acknowledged that this could be as high as 15% in young children (Byrne et al., 2015, Gould et al., 2009). Strains of STEC encoding s*tx2* toxin genes are more often associated with HUS than other strains (Byrne et al., 2015, Dallman et al., 2015, Ethelberg et al., 2004, Lynn et al., 2005, Milford et al., 1990, Persson et al., 2007).

Infection with STEC is a relatively rare cause of gastrointestinal illness in England, with around 900 cases diagnosed annually (Public Health England, 2016). Several

large and severe outbreaks have occurred, notably three outbreaks which led to inquiries including; an outbreak in Central Scotland in 1996 associated with a butchers which resulted in 496 cases, 127 individuals admitted to hospital and 18 deaths (The Pennington Group, 1997); an outbreak in South Wales in 2009 associated with contaminated meat sourced from an abattoir which resulted in 157 cases, 31 hospital admissions and one infant death (Pennington, 2009) and an outbreak at a petting farm in England in 2009 resulted in 93 cases, mostly children, 17 of whom developed HUS (Griffin, 2010).

Transmission to humans occurs through consumption of contaminated food (Ackers et al., 1998, Gillespie et al., 2005, Rangel et al., 2005) or exposure to a contaminated environment involving direct or indirect contact with animals or their faeces, including petting farm visits (David et al., 2004, Griffin, 2010, Payne et al., 2003) and swimming in contaminated water (Rangel et al., 2005, Verma et al., 2007). Furthermore, the low infectious dose of STEC (Teunis et al., 2004, Tuttle et al., 1999) means that once in a population, person-to-person spread is common (Adams et al., 2016b, Al-Jader et al., 1999, Swerdlow and Griffin, 1997).

Risk factors for STEC infection are well documented and include a variety of foodborne, waterborne and environmental factors as well as foreign travel (Gillespie et al., 2005, Locking et al., 2001, Parry et al., 1998). There is evidence to suggest that those who are disadvantaged have a lower risk of STEC infection (Chang et al., 2009, Jalava et al., 2011, Whitney et al., 2015), and potentially a lower risk of progression to HUS outside of England (Rowe et al., 1991, Whitney et al., 2015), however no studies have looked at the relationship between SES, STEC and HUS in England; studies reporting differences in risk of GI infection by SES have hypothesised that these differences may be due to differential exposure such as through travel, eating outside of the home or dietary preferences (Chang et al., 2009, Jalava et al., 2011, Whitney et al., 2015) or related to healthcare interaction (Whitney et al., 2015). There is limited evidence as to whether these hypothesised associations are real, and no studies exploring this in England. Despite many interventions to reduce the incidence of STEC infection over the last 30 years, which have resulted in changes in risk factors and exposures, levels of infection have remained relatively stable (Adams et al., 2016b).

Few studies have explored the social patterning of risk factors for STEC (Bentancor et al., 2012) or the socio-demographic risk factors associated with progression to HUS, and no such studies have been undertaken in England. In this chapter I will present the results from the analysis of the social patterning of STEC risk factors and the analysis of a paediatric cohort of STEC cases, including details of sensitivity and robustness analyses and the implications of the findings for future work and on our understanding for public health.

7.2 Methods – Social patterning of clinical outcomes, healthcare contact and risk factors for STEC

The methods for this analysis are found in greater detail in Chapter 3. A brief overview is provided below.

Design, setting and data source

An observational study design was used to assess the relationship between SES and a variety of risk factors for STEC infection. Data were extracted from the PHE National Enhanced Surveillance System for STEC (NESSS), described in detail in Chapter 3.

Participants

All microbiologically confirmed, probable or clinically suspect, symptomatic STEC cases with onset dates between 1st January 2010 and 31st December 2015 (inclusive) recorded in NESSS were included. Both sporadic and outbreak-associated cases were included. Figure 3.6 (Chapter 3) described the selection of participants for inclusion in this study.

Asymptomatic, microbiologically suspect or shiga-toxin (*stx*) negative individuals, or those lost-to-follow-up or without an ESQ were excluded prior to data extraction as follow-up information is not routinely obtained for these groups and therefore it was not possible to obtain details of exposures for these individuals. Individuals for whom postcode could not be linked to an IMD score were excluded. The presence of *stx* genes was used as the main microbiological variable.

Outcomes and covariates

The outcomes of interest were a range of reported exposures including foodborne, waterborne and environmental risk factors. The association between each of these and the primary exposure of interest (SES) were tested.

Socioeconomic status was determined using a small area deprivation measure, the Index of Multiple Deprivation 2010 (IMD) (Department for Communities and Local Government, 2011), assigned to each individual based on their full postcode. Other covariates of interest included in the analysis were age; sex; ethnicity; rurality; whether the case was associated with an outbreak; clinical symptoms; healthcare contact; antibiotic and anti-diarrhoeal use; and microbiology. Where symptoms, travel status, healthcare contact or exposure variables were blank or unknown, these were recoded as a negative response.

Analysis strategy

Analyses were conducted in Stata 13.1 (Statacorp, Texas). The distribution of the population by IMD quintile overall and by age and sex was compared to the study population and crude incidence rates, as well as age- and sex-adjusted incidence rates, accounting for the population 'at risk' were produced to explore whether the patterns identified in the descriptive analysis could be explained as following the expected distribution within the population. A descriptive analysis of the distribution of risk factors and clinical presentation by IMD quintiles was undertaken and chi square test for trend used to assess whether there was a statistically significant relationship between IMD quintile and each of the variables in turn. Variables identified as significantly related to IMD in the descriptive analysis ($p=<0.05$) were included in subsequent analyses. Foreign travel associated cases were excluded from the analysis of risk factors, as follow-up information on other risk factors is not sought for these cases in NESSS.

Univariate and multivariable logistic regression models were used to assess the relationship between IMD quintile and clinical presentation, adjusting for age, sex, ethnicity, rurality and the presence of *stx* gene as a marker of severity.

The relationship between IMD and foreign travel was also assessed in this way, however the multivariable model did not control for s*tx* genes as this is not on the causal pathway between SES and exposure to risk factors.

Foreign travel associated cases were then excluded and univariate and multivariable logistic regression was then used to assess the relationship between IMD quintile and each risk factor variable adjusting for age, sex, ethnicity and rurality; as for foreign travel, *stx* gene was not included in these models. Due to missing data for the ethnicity variable (19.1%), multiple imputation using chained equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was used to impute values where ethnicity (White/non-White) was missing. Fifty imputed datasets were generated. The distribution of ethnicity by age and sex was assessed to check the missing at random (MAR) assumption.

To assess the robustness of our findings, the risk factor analysis was repeated for sporadic cases and for cases aged less than 16 years to determine whether there were differences in risk factors by SES for children. The clinical presentation analyses were repeated on restricted datasets for i) non-travel associated cases, ii) sporadic cases and iii) children aged less than 16 years.

7.3 Methods – socio-demographic risk factors in development of HUS

The methods for this analysis are found in greater detail in Chapter 3. A brief overview is provided below.

Design, setting and data source

For this study, a retrospective cohort analysis was conducted using two linked data sources; PHE NESSS and the British Paediatric Surveillance Unit (BPSU) HUS Study in conjunction with PHE. Data were extracted on all symptomatic confirmed, probable or suspect (clinical) STEC cases aged 0-15 years (inclusive) identified in NESSS and clinical data on paediatric (aged <16 years) HUS cases, collected by BPSU HUS Study during the period of the BPSU HUS Study (1st October 2011 to 31st October 2014). Both surveillance systems are described in greater detail in Chapter 3. Cases in the BPSU dataset were linked to those in the NESSS dataset to create a retrospective cohort.

Participants

The linkage of two robust datasets, both of which can record HUS status, ensures high ascertainment of HUS cases. Figure 3.7 (Chapter 3) described the selection of participants for inclusion in this study.

Outbreak cases were included. Asymptomatic, suspect or *stx*-negative individuals were treated as non-cases and excluded as follow-up information is not routinely obtained for these groups. Individuals without an ESQ were also excluded as it was not possible to obtain details of illness for these individuals. Cases identified in the absence of microbiological confirmation of STEC or with a serological result were retained for summarising but excluded from the regression model as it is not always possible to obtain a stool sample in HUS cases and excluding these individuals may have led to an under ascertainment of HUS. All HUS cases included in this study were diagnosed with 'typical' HUS as they experienced a diarrhoeal prodrome.

Outcome and covariates

The main outcome of interest was HUS. This was determined either by the presence of the case in the BPSU dataset, conforming to a set of clinical criteria for HUS (Byrne, 2017) or by the completion of the HUS field in the ESQ and recorded within NESSS. As a proxy for childhood SEC a small area deprivation measure, the Index of Multiple Deprivation 2010 (IMD) (Department for Communities and Local Government, 2011), was used and assigned to each individual based on their full postcode. The IMD score was divided into population-level quintiles, with the first quintile representing the least disadvantaged and the fifth quintile representing the most disadvantaged. Other covariates of interest included in the analysis were age group (<1, 1-4, 5-9, 10-15 years); sex (male/female); ethnicity (White/non-White); travel (yes/no); rurality (rural/urban); microbiology (*stx1*, *stx2*, *stx1+2*); antibiotic use (yes/no); clinical symptoms (diarrhoea, bloody diarrhoea, nausea, vomiting, abdominal pain, fever) and region (regions of England). The presence of *stx* genes was used as the main microbiological variable.

Analysis strategy

Analyses were conducted in Stata 13.1 (Statacorp, Texas). The main analysis investigated the relationship between SEC, as measured by IMD, and development of

HUS among paediatric STEC cases operationalised as a binary variable (0=no progression to HUS; 1=progression to HUS). Firstly this was investigated through a descriptive analysis of the patterning of sociodemographic risk factors by IMD quintiles.

Univariate relationships were then explored between SECs and the covariates of interest; age; sex; ethnicity; travel; rurality; microbiology; antibiotic use; clinical symptoms; and region before fitting a multivariable logistic regression model, adjusting for these covariates. Healthcare contact variables were excluded from the main analysis as these are not regarded as potential confounders. Cases identified via serological testing only (n=66) or for whom no microbiological information was available (n=4) were excluded in order to assess the role of s*tx* genes. All variables were retained in this model in order to control for any potential confounding. Interaction terms between variables (IMD, ethnicity, age and sex) were tested for inclusion to investigate whether the strength of any relationship was moderated by the inclusion of another variable. Due to missing data for the ethnicity variable, multiple imputation using chained equations (MICE) (UCLA Institute for Digital Research and Education, 2017a) was used to impute values where ethnicity (White/non-White) was missing (20.1%). The distribution of ethnicity by age, sex and region was assessed to check the missing at random (MAR) assumption.

To further explore the relationship between age and sex in this cohort, a fractional polynomial prediction plot was fitted to detect the best functional form for age (as a continuous variable) and sex by HUS. A likelihood ratio test was performed to test whether the fractional polynomial model provided a better fit in comparison to the linear model.

To determine whether there were differences in progression to HUS by SECs for children who travelled abroad during their incubation period compared to those who did not, the main analysis was repeated excluding travel-related cases.

Finally, to explore the impact of ethnicity, due to the 20.1% of records which were of unknown ethnicity, a multivariable logistic regression model was fitted with all covariates excluding ethnicity. To further explore potential issues of multicollinearity between IMD and ethnicity a post-hoc matched analysis on ethnicity using conditional logistic regression and penalised logistic regression on the multiply

imputed dataset were conducted. The post-hoc matched analysis was conducted on a smaller number of variables in order to provide the simplest but most complete model possible.

7.4 Results – Social patterning of clinical outcomes, healthcare contact and risk factors for STEC

Characteristics of participants

A total of 4115 primary, symptomatic cases of STEC in England were recorded in NESSS between 1st January 2010 and 31st December 2015. Of these, it was not possible to ascertain the IMD for 143 cases due to invalid or incomplete postcode information. Therefore, 3,972 cases of STEC were included in this analysis (Table 7.1 and 7.2). Information on exposure to risk factors was available for 2,961 non-travel cases. Ethnicity was missing for 19.1% of cases (n=759); ethnicity was imputed for these cases for the main analysis.

Overall crude incidence was significantly lower among the most disadvantaged compared to the least disadvantaged (Table 7.1); this pattern was also reflected in the age- and sex-adjusted incidence rates for those over 5 years of age but, the 0-4 age group showed the opposite pattern, although the observed IRRs were not statistically significant for any age-group.

Table 7.1: Incidence rate ratio for exposed (most disadvantaged) compared to unexposed (least disadvantaged) - STEC

	Rate/100,000 in most disadvantaged	Rate/100,000 in least disadvantaged	Incidence rate difference	Incidence rate ratio	95% CI
Overall	0.97	1.49	-0.51	0.65	0.50-0.84
Female	1.00	1.65	-0.66	0.61	0.42-0.86
Male	0.95	1.32	-0.37	0.71	0.48-1.06
0-4	3.84	3.25	0.59	1.17	0.63-2.14
5-9	1.89	2.57	-0.68	0.78	0.33-1.79
10-15	1.04	2.25	-1.21	0.48	0.18-1.17
16-19	0.89	2.27	-1.38	0.42	0.12-1.27
20-59	0.76	0.96	-0.20	0.79	0.51-1.19
60+	0.54	1.65	-1.10	0.34	0.16-0.66

Table 7.2 shows the study population stratified by IMD quintile. This shows that when case numbers rather than incidence rates are used, there is a similar social gradient in that the population burden and the individual risk are both higher in more

affluent areas and therefore justifies the use of case numbers in the analysis. There were more cases from the least disadvantaged quintile and fewer cases from the most disadvantaged quintile in comparison to the distribution of approximately 20% in each quintile within the general population.

A higher proportion of those in the most disadvantaged quintile were infected with s*tx*2 only, associated with more severe disease and progression to HUS, and a lower proportion of those in the most disadvantaged quintile were infected with *stx*1+2. Vomiting, diarrhoea, s*tx* gene and healthcare contact variables (GP, A&E and hospitalisation) were independently associated with SES. Water and environmental exposures, as well as travel within and outside of the UK were also independently associated with SES.

Table 7.2: Characteristics of STEC cases by Index of Multiple Deprivation quintile

	Q1 (least disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (most disadvantaged) n (%)	p value[a]
Total	945 (23.8)	913 (23.0)	821 (20.7)	699 (17.6)	594 (15.0)	<0.01
Demographics (n=3,972)						
Age group						<0.01
<1	14 (1.5)	18 (2.0)	9 (1.1)	13 (1.9)	19 (3.2)	
1-4	133 (14.1)	142 (15.6)	154 (18.8)	125 (17.9)	118 (19.9)	
5-9	97 (10.3)	95 (10.4)	83 (10.1)	60 (8.6)	63 (10.6)	
10-15	99 (10.5)	74 (8.1)	67 (8.2)	55 (7.9)	45 (7.6)	
16-19	69 (7.3)	57 (6.2)	38 (4.6)	33 (4.7)	27 (4.6)	
20-59	357 (37.8)	353 (38.7)	321 (39.1)	294 (42.1)	239 (40.2)	
60+	176 (18.6)	174 (19.1)	149 (18.2)	119 (17.0)	83 (14.0)	
Sex						0.11
Male	412 (43.6)	381 (41.7)	345 (42.0)	288 (41.2)	284 (47.8)	
Female	533 (56.4)	532 (58.3)	476 (58.0)	411 (58.8)	310 (52.2)	
Ethnicity						<0.001
White	758 (80.2)	706 (77.3)	634 (77.2)	475 (68.0)	341 (57.4)	
Non-White	28 (3.0)	32 (3.5)	34 (4.1)	73 (10.4)	132 (22.2)	
Unknown	159 (16.8)	175 (19.2)	153 (18.6)	151 (21.6)	121 (20.4)	
Rurality						<0.001
Urban	671 (71.0)	542 (59.4)	512 (62.4)	605 (86.6)	582 (98.0)	
Rural	274 (29.0)	371 (40.6)	309 (37.6)	94 (13.5)	12 (2.0)	
Outbreak	154 (16.3)	145 (15.9)	131 (16.0)	83 (11.9)	72 (12.1)	0.02
Clinical (n=3,972)						
Diarrhoea	917 (97.0)	858 (94.0)	785 (95.6)	669 (95.7)	573 (96.5)	0.02
Bloody diarrhoea	602 (63.7)	572 (62.6)	558 (68.0)	443 (63.4)	386 (65.0)	0.17
Nausea	470 (49.7)	436 (47.8)	403 (49.1)	328 (46.9)	295 (49.7)	0.76
Vomiting	318 (33.7)	344 (37.7)	298 (36.3)	254 (36.3)	259 (43.6)	<0.01
Abdominal pain	815 (86.2)	771 (84.5)	683 (83.2)	580 (83.0)	486 (81.8)	0.15
Fever	316 (33.4)	291 (31.9)	283 (34.5)	246 (35.2)	208 (35.0)	0.60
HUS	28 (3.0)	36 (3.9)	40 (4.9)	29 (4.2)	21 (3.5)	0.33

	Q1 (least disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (most disadvantaged) n (%)	p value[a]
Healthcare (n=3,972)						
Antibiotics	133 (14.1)	128 (14.0)	102 (12.4)	102 (14.6)	77 (13.0)	0.72
Antidiarrhoeals	235 (24.9)	213 (23.3)	172 (21.0)	147 (21.0)	125 (21.0)	0.19
NHS Direct	113 (12.0)	96 (10.5)	87 (10.6)	68 (9.7)	47 (7.9)	0.15
GP	665 (70.4)	637 (69.8)	576 (70.2)	452 (64.7)	367 (61.8)	0.001
A&E	204 (21.6)	195 (21.4)	168 (20.5)	199 (28.5)	174 (29.3)	<0.001
Hospital	284 (30.1)	299 (32.8)	286 (34.8)	251 (35.9)	232 (39.1)	<0.01
Microbiology (n=3,972)						
Stx						
*Stx*1+2	385 (40.7)	354 (38.8)	306 (37.3)	249 (35.6)	186 (31.3)	<0.01
*Stx*1	35 (3.7)	26 (2.9)	20 (2.4)	27 (3.9)	27 (4.6)	
*Stx*2	525 (55.6)	533 (58.4)	495 (60.3)	423 (60.5)	381 (64.1)	
Food Exposures (n=2,961)						
Ate outside the home	492 (75.5)	515 (74.3)	458 (71.5)	376 (71.7)	307 (68.1)	0.06
Any meat	569 (87.3)	602 (86.9)	552 (86.1)	465 (88.7)	391 (86.7)	0.75
Any fish/shellfish	354 (54.3)	325 (46.9)	315 (49.1)	254 (48.5)	200 (44.4)	0.01
Any dairy	575 (88.2)	612 (88.3)	572 (89.2)	452 (86.3)	386 (85.6)	0.32
Any salad/fruit/vegetables/herbs	540 (82.8)	562 (81.1)	512 (79.9)	413 (78.8)	338 (74.9)	0.02
Juice	267 (41.0)	261 (37.7)	245 (38.2)	191 (36.5)	178 (39.5)	0.56
Water Exposures (n=2,961)						
Mains water	586 (89.9)	621 (89.6)	553 (86.3)	455 (86.8)	391 (86.7)	0.13
Private water supply	15 (2.3)	35 (5.1)	43 (6.7)	17 (3.2)	1 (0.2)	<0.001
Bottled water	272 (41.7)	224 (32.3)	255 (39.8)	197 (37.6)	175 (38.8)	<0.01
Unboiled water	11 (1.7)	12 (1.7)	12 (1.5)	8 (1.5)	2 (0.4)	0.36
Recreational freshwater exposure	123 (18.9)	129 (18.6)	108 (16.9)	67 (12.8)	49 (10.9)	<0.001
Recreational seawater exposure	37 (5.7)	39 (5.6)	36 (5.6)	28 (5.3)	13 (2.9)	0.22
Environmental Exposures (n=2,961)						
Any animal contact	447 (68.6)	471 (68.0)	432 (67.4)	340 (64.9)	252 (55.9)	<0.001
Walked in paddock	149 (22.9)	205 (29.6)	175 (27.3)	99 (18.9)	40 (8.9)	<0.001
Visited a farm	91 (14.0)	122 (17.6)	104 (16.2)	62 (11.8)	48 (10.6)	<0.01
Day trip	156 (23.9)	165 (23.8)	150 (23.4)	102 (19.5)	71 (15.7)	<0.01
Contact with soil	168 (25.8)	205 (29.6)	154 (24.0)	104 (19.9)	61 (13.5)	<0.001

	Q1 (least disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (most disadvantaged) n (%)	p value[a]
Travel (n=3,972)						
UK travel	203 (21.5)	198 (21.7)	163 (19.9)	139 (19.9)	73 (12.3)	<0.001
Non-UK travel	293 (31.0)	220 (24.1)	180 (21.9)	175 (25.0)	143 (24.1)	<0.001

[a] Statistical significance of relationship between IMD quintile and each variable, tested using χ^2 test

HUS – haemolytic uraemic syndrome; s*tx* – shiga toxin type; NHS Direct – National Health Service telephone advice line, now NHS 111; GP – General Practitioner; A&E – accident and emergency

Main analysis

In univariate analysis, those who were more disadvantaged were less likely to be associated with an outbreak, more likely to report vomiting, visit A&E and be hospitalised, but less likely to visit their GP or have a travel-associated infection (Table 7.3).

In multivariable analysis accounting for age, sex, ethnicity, rurality and potential severity (defined by s*tx* gene), those who were more disadvantaged were more likely to report vomiting (OR 1.61, p<0.001). They were also more likely to visit A&E or be hospitalised for their illness (OR 1.35, p=0.02; OR 1.71, p<0.001 respectively), but less likely to visit their GP (OR 0.67, p<0.01). Those who were more disadvantaged were also less likely to have reported foreign travel (OR 0.57, p<0.001, Table 7.4).

For indigenously-acquired STEC cases, those who were more disadvantaged were significantly less likely to report exposure to; fish/shellfish; salad, fruit, vegetables or herbs; recreational freshwater; walking in a paddock; taking a day trip; contact with soil or; travel within the UK. There were no significant differences identified between the most compared to the least disadvantaged for food exposures, nor for commonly cited reasons for differential STEC infection by SES such as eating out, animal contact or visiting a farm (Table 7.5).

Table 7.3: Univariate and multivariable regression analysis – clinical outcomes and healthcare contact (n=3,972)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Outbreak	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.97	(0.76-1.24)	1.03	(0.80-1.32)	0.84
	3	0.98	(0.76-1.26)	1.03	(0.80-1.34)	0.80
	4	0.69	(0.52-0.92)	0.71	(0.53-0.96)	0.02
	5 (most disadvantaged)	0.71	(0.52-0.96)	0.79	(0.58-1.09)	0.16
Diarrhoea	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.48	(0.30-0.76)	0.46	(0.29-0.74)	<0.01
	3	0.67	(0.40-1.10)	0.67	(0.40-1.11)	0.12
	4	0.68	(0.40-1.15)	0.78	(0.46-1.34)	0.37
	5 (most disadvantaged)	0.83	(0.47-1.48)	1.12	(0.61-2.06)	0.72
Vomiting	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.19	(0.99-1.44)	1.20	(0.99-1.46)	0.06
	3	1.12	(0.92-1.37)	1.13	(0.93-1.38)	0.23
	4	1.13	(0.92-1.38)	1.17	(0.95-1.45)	0.15
	5 (most disadvantaged)	1.52	(1.23-1.88)	1.61	(1.28-2.02)	<0.001
GP	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.97	(0.80-1.19)	0.95	(0.77-1.16)	0.58
	3	0.99	(0.81-1.21)	0.97	(0.79-1.19)	0.74
	4	0.77	(0.63-0.95)	0.76	(0.62-0.95)	0.01
	5 (most disadvantaged)	0.68	(0.55-0.85)	0.67	(0.53-0.84)	<0.01
A&E	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.99	(0.79-1.23)	1.05	(0.84-1.32)	0.67
	3	0.93	(0.74-1.18)	0.99	(0.78-1.24)	0.90
	4	1.45	(1.15-1.81)	1.39	(1.10-1.75)	0.01
	5 (most disadvantaged)	1.50	(1.19-1.90)	1.35	(1.05-1.74)	0.02
Hospital	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.13	(0.93-1.38)	1.15	(0.94-1.40)	0.17
	3	1.24	(1.02-1.52)	1.27	(1.04-1.56)	0.02
	4	1.30	(1.06-1.61)	1.41	(1.14-1.74)	<0.01
	5 (most disadvantaged)	1.49	(1.20-1.85)	1.71	(1.36-2.15)	<0.001

[a] Adjusted for age group, sex, ethnicity, rurality and *stx* gene; *stx* – Shiga toxin gene; GP – General Practice, A&E – Accident and Emergency

Table 7.4: Univariate and multivariable regression analysis – foreign travel as a risk factor (n=3,972)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Non-UK travel	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.71	(0.58-0.87)	0.74	(0.60-0.91)	0.01
	3	0.62	(0.50-0.77)	0.65	(0.52-0.81)	<0.001
	4	0.74	(0.60-0.93)	0.64	(0.51-0.81)	<0.001
	5 (most disadvantaged)	0.71	(0.56-0.89)	0.57	(0.44-0.73)	<0.001

[a]Adjusted for age group, sex, ethnicity and rurality

Table 7.5: Univariate and multivariable univariate regression analysis – risk factors (n=2,961)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Any fish/shellfish	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.74	(0.60-0.92)	0.73	(0.59-0.90)	<0.01
	3	0.81	(0.65-1.01)	0.80	(0.64-0.99)	0.04
	4	0.79	(0.63-1.00)	0.78	(0.62-0.99)	0.04
	5 (most disadvantaged)	0.67	(0.53-0.85)	0.68	(0.52-0.87)	<0.01
Any salad/fruit/vegetables/herbs	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.89	(0.67-1.18)	0.84	(0.64-1.12)	0.24
	3	0.82	(0.62-1.09)	0.76	(0.57-1.02)	0.06
	4	0.77	(0.58-1.03)	0.76	(0.56-1.02)	0.07
	5 (most disadvantaged)	0.62	(0.46-0.82)	0.63	(0.46-0.86)	<0.01
Private water supply	1 (least disadvantaged)	1.0 (reference)				
	2	2.26	(1.22-4.18)			
	3	3.05	(1.68-5.55)	Omitted due to small numbers		
	4	1.42	(0.70-2.88)			
	5 (most disadvantaged)	0.09	(0.01-0.72)			
Bottled water	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.67	(0.53-0.83)	0.66	(0.53-0.83)	<0.001
	3	0.92	(0.74-1.15)	0.90	(0.72-1.13)	0.36
	4	0.84	(0.67-1.07)	0.78	(0.61-1.00)	0.05
	5 (most disadvantaged)	0.89	(0.69-1.13)	0.80	(0.61-1.04)	0.10
Recreational freshwater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.98	(0.75-1.29)	1.02	(0.76-1.37)	0.89
	3	0.87	(0.65-1.16)	0.90	(0.66-1.22)	0.50
	4	0.63	(0.46-0.87)	0.70	(0.49-0.98)	0.04
	5 (most disadvantaged)	0.52	(0.37-0.75)	0.65	(0.44-0.95)	0.03
Recreational seawater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.99	(0.62-1.57)	1.03	(0.64-1.65)	0.92
	3	0.99	(0.62-1.59)	1.02	(0.63-1.66)	0.93
	4	0.94	(0.57-1.55)	1.02	(0.61-1.72)	0.93
	5 (most disadvantaged)	0.49	(0.26-0.94)	0.60	(0.31-1.17)	0.13

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Any animal contact	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.97	(0.77-1.22)	0.95	(0.75-1.21)	0.68
	3	0.95	(0.75-1.20)	0.95	(0.74-1.21)	0.67
	4	0.85	(0.66-1.08)	1.09	(0.84-1.42)	0.52
	5 (most disadvantaged)	0.58	(0.45-0.74)	1.01	(0.76-1.33)	0.97
Walked in paddock	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.42	(1.11-1.81)	1.36	(1.05-1.75)	0.02
	3	1.27	(0.99-1.63)	1.23	(0.94-1.60)	0.13
	4	0.79	(0.59-1.05)	1.04	(0.77-1.40)	0.81
	5 (most disadvantaged)	0.33	(0.23-0.48)	0.55	(0.37-0.82)	<0.01
Visited a farm	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.32	(0.98-1.77)	1.30	(0.95-1.77)	0.10
	3	1.19	(0.88-1.62)	1.17	(0.85-1.61)	0.34
	4	0.83	(0.59-1.17)	0.90	(0.63-1.29)	0.57
	5 (most disadvantaged)	0.73	(0.51-1.07)	0.94	(0.62-1.40)	0.74
Day trip	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.99	(0.77-1.28)	1.03	(0.80-1.33)	0.83
	3	0.97	(0.75-1.26)	1.00	(0.77-1.30)	0.99
	4	0.77	(0.58-1.02)	0.79	(0.59-1.06)	0.11
	5 (most disadvantaged)	0.59	(0.44-0.81)	0.64	(0.46-0.89)	0.01
Contact with soil	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.21	(0.95-1.54)	1.15	(0.90-1.46)	0.28
	3	0.91	(0.71-1.17)	0.88	(0.68-1.14)	0.32
	4	0.71	(0.54-0.94)	0.87	(0.65-1.16)	0.33
	5 (most disadvantaged)	0.45	(0.33-0.62)	0.66	(0.47-0.92)	0.02
UK travel	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.89	(0.70-1.13)	0.91	(0.71-1.16)	0.44
	3	0.78	(0.60-1.00)	0.80	(0.62-1.03)	0.09
	4	0.80	(0.62-1.05)	0.86	(0.66-1.13)	0.28
	5 (most disadvantaged)	0.46	(0.33-0.62)	0.54	(0.39-0.74)	<0.001

[a]Adjusted for age group, sex, ethnicity and rurality

Sensitivity analyses

Multiple sensitivity analyses were conducted to assess the robustness of the findings; in particular to explore the social patterning of clinical outcomes and healthcare contact excluding foreign travel associated cases (Table 7.6); to explore potential impact of excluding outbreak cases on the social patterning of clinical outcomes, healthcare contact (Table 7.7) and exposure to known risk factors for sporadic and non-travel associated cases (Table 7.8); and to explore potential differences when investigating the association in children only (Tables 7.9 and 7.10).

These analyses in the main did not alter the overall findings of this research (Tables 7.3-7.5); namely that there are differences in healthcare contact by SES, when controlling for age, sex, ethnicity, rurality and severity; that commonly cited explanations for differences in STEC infection by SES, namely eating out and visiting farms, were not reported significantly differently by SES. Other exposures such as foreign travel, water exposures, walking in a paddock, contact with soil and day trips and travel within the UK were less likely to be reported by more disadvantaged STEC cases.

Sensitivity analysis performed on clinical outcomes and healthcare contact for non-foreign travel related cases (indigenous cases) suggested that those who were more disadvantaged were significantly less likely to be associated with an outbreak or visit their GP but significantly more likely to report vomiting, visit A&E or be hospitalised (Table 7.6).

Sensitivity analysis performed on sporadic (non-outbreak) cases echoed the main findings for both clinical outcomes and healthcare contact as well as reported risk factors. Disadvantaged individuals with a sporadic infection were almost twice as likely to report hospitalisation as the least disadvantaged (Table 7.7).

Restricting the analysis to children aged less than 16 years did yield different results (Table 7.9); contact with a GP and visiting A&E were no longer statistically significant. Difference in reporting of foreign travel (Table 7.10) and salad, fruit, vegetables or herbs (Table 7.11) were also no longer statistically significant.

Table 7.6: Univariate and multivariable regression analysis – clinical outcomes and healthcare contact non-travel cases (n=2,961)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Outbreak	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.84	(0.65-1.09)	0.92	(0.70-1.21)	0.55
	3	0.83	(0.63-1.08)	0.90	(0.68-1.18)	0.43
	4	0.59	(0.44-0.80)	0.58	(0.43-0.79)	<0.01
	5 (most disadvantaged)	0.61	(0.45-0.84)	0.65	(0.47-0.91)	0.01
Diarrhoea	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.60	(0.36-1.00)	0.59	(0.36-0.99)	0.05
	3	0.91	(0.52-1.58)	0.61	(0.52-1.59)	0.73
	4	0.73	(0.42-1.28)	0.85	(0.48-1.50)	0.57
	5 (most disadvantaged)	0.91	(0.49-1.67)	1.26	(0.66-2.42)	0.48
Vomiting	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.20	(0.96-1.50)	1.21	(0.96-1.52)	0.10
	3	1.14	(0.91-1.43)	1.17	(0.93-1.48)	0.19
	4	1.19	(0.94-1.51)	1.22	(0.95-1.56)	0.12
	5 (most disadvantaged)	1.56	(1.22-1.99)	1.63	(1.25-2.13)	<0.001
GP	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.99	(0.79-1.25)	0.96	(0.76-1.21)	0.73
	3	1.08	(0.85-1.36)	1.05	(0.83-1.33)	0.69
	4	0.84	(0.66-1.07)	0.87	(0.68-1.11)	0.26
	5 (most disadvantaged)	0.72	(0.56-0.92)	0.73	(0.56-0.95)	0.02
A&E	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.97	(0.75-1.25)	1.03	(0.79-1.34)	0.82
	3	0.87	(0.67-1.13)	0.91	(0.70-1.19)	0.50
	4	1.48	(1.14-1.93)	1.37	(1.05-1.79)	0.02
	5 (most disadvantaged)	1.67	(1.28-2.18)	1.43	(1.07-1.90)	0.01
Hospital	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.16	(0.93-1.46)	1.19	(0.95-1.50)	0.13
	3	1.16	(0.92-1.46)	1.20	(0.95-1.51)	0.13
	4	1.40	(1.03-1.66)	1.36	(1.06-1.74)	0.01
	5 (most disadvantaged)	1.63	(1.27-2.08)	1.79	(1.38-2.34)	<0.001

[a]Adjusted for age group, sex, ethnicity, rurality and *stx* gene; *stx* – Shiga toxin gene; GP – General Practice, A&E – Accident and Emergency

Table 7.7: Univariate and multivariable regression analysis – clinical outcomes and healthcare contact sporadic cases only (n=3,387)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Diarrhoea	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.64	(0.38-1.08)	0.63	(0.37-1.07)	0.09
	3	0.67	(0.39-1.14)	0.67	(0.39-1.15)	0.15
	4	0.68	(0.39-1.20)	0.76	(0.43-1.34)	0.34
	5 (most disadvantaged)	0.93	(0.49-1.75)	1.18	(0.60-2.30)	0.63
Vomiting	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.17	(0.95-1.44)	1.18	(0.96-1.46)	0.12
	3	1.14	(0.92-1.41)	1.14	(0.96-1.41)	0.26
	4	1.12	(0.90-1.40)	1.16	(0.93-1.46)	0.19
	5 (most disadvantaged)	1.52	(1.21-1.91)	1.61	(1.26-2.06)	<0.001
GP	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.01	(0.81-1.26)	0.99	(0.79-1.24)	0.92
	3	0.92	(0.73-1.15)	0.90	(0.72-1.13)	0.37
	4	0.74	(0.59-0.93)	0.72	(0.57-0.91)	0.01
	5 (most disadvantaged)	0.61	(0.48-0.78)	0.57	(0.44-0.73)	<0.001
A&E	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.07	(0.84-1.37)	1.15	(0.89-1.47)	0.28
	3	1.03	(0.80-1.33)	1.10	(0.85-1.42)	0.48
	4	1.48	(1.16-1.89)	1.42	(1.10-1.83)	0.01
	5 (most disadvantaged)	1.54	(1.19-1.99)	1.37	(1.04-1.81)	0.02
Hospital	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.18	(0.95-1.46)	1.19	(0.96-1.49)	0.12
	3	1.28	(1.03-1.60)	1.30	(1.04-1.63)	0.02
	4	1.35	(1.08-1.69)	1.48	(1.18-1.87)	<0.01
	5 (most disadvantaged)	1.59	(1.26-2.01)	1.91	(1.48-2.46)	<0.001

[a] Adjusted for age group, sex, ethnicity, rurality and *stx* gene; *stx* – Shiga toxin gene; GP – General Practice, A&E – Accident and Emergency

Table 7.8: Univariate and multivariable regression analysis – risk factors for non-travel sporadic cases only (n=2,396)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Any fish/shellfish	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.75	(0.59-0.96)	0.74	(0.58-0.95)	0.03
	3	0.85	(0.66-1.08)	0.83	(0.65-1.07)	0.15
	4	0.76	(0.59-0.99)	0.74	(0.57-0.97)	0.03
	5 (most disadvantaged)	0.64	(0.49-0.84)	0.62	(0.46-0.82)	<0.01
Any salad/fruit/vegetables/herbs	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.89	(0.64-1.23)	0.84	(0.60-1.18)	0.32
	3	0.81	(0.58-1.13)	0.75	(0.54-1.05)	0.09
	4	0.67	(0.48-0.93)	0.64	(0.45-0.90)	0.01
	5 (most disadvantaged)	0.55	(0.40-0.77)	0.54	(0.38-0.78)	0.001
Private water supply	1 (least disadvantaged)	1.0 (reference)				
	2	2.06	(1.10-3.83)			
	3	2.59	(1.41-4.77)	Omitted due to small numbers		
	4	1.05	(0.50-2.21)			
	5 (most disadvantaged)	0.09	(0.01-0.65)			
Bottled water	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.73	(0.57-0.93)	0.72	(0.56-0.93)	0.01
	3	0.95	(0.74-1.22)	0.91	(0.71-1.18)	0.49
	4	0.86	(0.66-1.11)	0.77	(0.59-1.01)	0.06
	5 (most disadvantaged)	0.87	(0.66-1.14)	0.75	(0.55-1.00)	0.05
Recreational freshwater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.93	(0.68-1.26)	0.96	(0.69-1.34)	0.83
	3	0.76	(0.55-1.05)	0.79	(0.56-1.11)	0.17
	4	0.58	(0.41-0.83)	0.62	(0.42-0.91)	0.01
	5 (most disadvantaged)	0.47	(0.32-0.70)	0.56	(0.36-0.87)	0.01
Recreational seawater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.99	(0.60-1.63)	1.06	(0.64-1.75)	0.83
	3	0.94	(0.57-1.57)	1.01	(0.60-1.70)	0.97
	4	0.91	(0.53-1.55)	0.97	(0.56-1.68)	0.91
	5 (most disadvantaged)	0.36	(0.17-0.75)	0.42	(0.19-0.93)	0.03
Any animal contact	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.13	(0.89-1.47)	1.13	(0.86-1.50)	0.38
	3	0.94	(0.72-1.23)	0.97	(0.73-1.28)	0.81
	4	0.86	(0.66-1.13)	1.14	(0.85-1.52)	0.39
	5 (most disadvantaged)	0.54	(0.41-0.72)	1.00	(0.73-1.37)	0.99

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Walked in paddock	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.22	(0.94-1.60)	1.15	(0.87-1.52)	0.32
	3	1.10	(0.84-1.45)	1.04	(0.78-1.39)	0.77
	4	0.60	(0.44-0.83)	0.79	(0.57-1.09)	0.15
	5 (most disadvantaged)	0.22	(0.15-0.34)	0.39	(0.25-0.61)	<0.001
Visited a farm	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.15	(0.82-1.59)	1.15	(0.82-1.63)	0.42
	3	1.14	(0.82-1.60)	1.16	(0.82-1.66)	0.40
	4	0.71	(0.48-1.04)	0.77	(0.51-1.14)	0.19
	5 (most disadvantaged)	0.62	(0.41-0.94)	0.79	(0.50-1.23)	0.30
Day trip	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.00	(0.76-1.33)	1.07	(0.80-1.43)	0.65
	3	0.91	(0.68-1.21)	0.97	(0.72-1.30)	0.82
	4	0.75	(0.55-1.02)	0.76	(0.55-1.04)	0.09
	5 (most disadvantaged)	0.53	(0.37-0.75)	0.56	(0.39-0.81)	<0.01
Contact with soil	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.24	(0.95-1.62)	1.18	(0.90-1.55)	0.24
	3	0.96	(0.72-1.27)	0.93	(0.70-1.24)	0.62
	4	0.71	(0.53-0.97)	0.87	(0.64-1.19)	0.40
	5 (most disadvantaged)	0.42	(0.29-0.60)	0.63	(0.43-0.92)	0.02
UK travel	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.77	(0.59-1.01)	0.79	(0.60-1.04)	0.09
	3	0.74	(0.56-0.97)	0.76	(0.58-1.01)	0.06
	4	0.67	(0.50-0.89)	0.71	(0.52-0.95)	0.02
	5 (most disadvantaged)	0.36	(0.25-0.51)	0.42	(0.29-0.60)	<0.001

[a]Adjusted for age group, sex, ethnicity and rurality

Table 7.9: Univariate and multivariable regression analysis – clinical outcomes and healthcare contact for cases aged <16 (n=1,483)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Outbreak	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.01	(0.69-1.49)	1.18	(0.79-1.76)	0.42
	3	0.95	(0.64-1.41	1.08	(0.72-1.63)	0.70
	4	0.49	(0.30-0.80)	0.53	(0.32-0.88)	0.02
	5 (most disadvantaged)	0.51	(0.31-0.83)	0.65	(0.38-1.11)	0.11
Diarrhoea	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.46	(0.23-0.91)	0.45	(0.22-0.90)	0.02
	3	0.91	(0.41-1.99)	0.94	(0.42-2.10)	0.88
	4	1.07	(0.45-2.54)	1.40	(0.57-3.44)	0.47
	5 (most disadvantaged)	0.76	(0.34-1.71)	1.32	(0.52-3.32)	0.56
Vomiting	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.26	(0.93-1.71)	1.23	(0.90-1.68)	0.19
	3	1.22	(0.90-1.66)	1.22	(0.89-1.68)	0.21
	4	1.08	(0.78-1.51)	1.12	(0.80-1.58)	0.51
	5 (most disadvantaged)	1.47	(1.06-2.04)	1.61	(1.11-2.33)	0.01
GP	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.26	(0.91-1.74)	1.24	(0.89-1.71)	0.20
	3	1.45	(1.04-2.01)	1.46	(1.04-1.04)	0.03
	4	1.20	(0.85-1.70)	1.28	(0.89-1.82)	0.18
	5 (most disadvantaged)	0.96	(0.68-1.35)	1.00	(0.68-1.47)	0.99
A&E	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.89	(0.62-1.28)	0.93	(0.64-1.34)	0.69
	3	1.11	(0.78-1.59)	1.15	(0.80-1.65)	0.46
	4	1.52	(1.06-2.19)	1.43	(0.98-2.09)	0.06
	5 (most disadvantaged)	1.40	(0.96-2.02)	1.27	(0.84-1.91)	0.26
Hospital	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.19	(0.86-1.63)	1.19	(0.86-1.66)	0.30
	3	1.49	(1.08-2.05)	1.52	(1.09-2.11)	0.01
	4	1.47	(1.05-2.06)	1.66	(1.16-2.37)	0.01
	5 (most disadvantaged)	1.31	(0.93-1.85)	1.77	(1.20-2.61)	<0.01

[a]Adjusted for age group, sex, ethnicity, rurality and *stx* gene; *stx* – Shiga toxin gene; GP – General Practice, A&E – Accident and Emergency

Table 7.10: Univariate and multivariable regression analysis – foreign travel as a risk factor for cases aged <16 (n=1,483)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Non-UK travel	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.62	(0.41-0.94)	0.64	(0.42-0.97)	0.03
	3	0.64	(0.43-0.97)	0.64	(0.42-0.98)	0.04
	4	0.85	(0.56-1.29)	0.68	(0.44-1.05)	0.08
	5 (most disadvantaged)	1.01	(0.76-1.51)	0.65	(0.41-1.03)	0.07

[a] Adjusted for age group, sex, ethnicity and rurality

Table 7.11: Univariate and multivariable regression analysis – risk factors for non-travel cases aged <16 (n=1,224)

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Any fish/shellfish	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.66	(0.47-0.93)	0.63	(0.45-0.88)	0.01
	3	0.80	(0.57-1.12)	0.76	(0.54-1.07)	0.11
	4	0.72	(0.50-1.03)	0.73	(0.50-1.06)	0.09
	5 (most disadvantaged)	0.62	(0.43-0.90)	0.64	(0.43-0.97)	0.04
Any salad, fruit, vegetables or herbs	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.01	(0.67-1.51)	0.96	(0.63-1.45)	0.83
	3	0.96	(0.64-1.45)	0.88	(0.58-1.34)	0.55
	4	0.87	(0.57-1.35)	0.86	(0.55-1.35)	0.52
	5 (most disadvantaged)	0.74	(0.48-1.14)	0.76	(0.47-1.23)	0.26
Private water supply	1 (least disadvantaged)	1.0 (reference)				
	2	2.42	(0.99-5.93)			
	3	3.39	(1.42-8.07)	Omitted due to small numbers		
	4	1.13	(0.37-3.41)			
	5 (most disadvantaged)	0.20	(0.02-1.61)			
Bottled water	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.68	(0.48-0.96)	0.68	(0.48-0.97)	0.04
	3	0.77	(0.54-1.09)	0.76	(0.53-1.08)	0.12
	4	0.75	(0.52-1.10)	0.71	(0.48-1.05)	0.09
	5 (most disadvantaged)	0.87	(0.60-1.27)	0.79	(0.52-1.21)	0.28
Recreational freshwater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.23	(0.86-1.76)	1.28	(0.89-1.85)	0.18
	3	1.08	(0.75-1.56)	1.16	(0.80-1.70)	0.43
	4	0.65	(0.42-0.99)	0.75	(0.49-1.17)	0.21
	5 (most disadvantaged)	0.58	(0.38-0.91)	0.90	(0.56-1.47)	0.68
Recreational seawater exposure	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.30	(0.73-2.32)	1.38	(0.77-2.49)	0.28
	3	1.07	(0.58-1.96)	1.14	(0.62-2.12)	0.67
	4	1.08	(0.56-2.08)	1.15	(0.59-2.24)	0.69
	5 (most disadvantaged)	0.62	(0.29-1.36)	0.81	(0.36-1.84)	0.62
Any animal contact	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.98	(0.67-1.43)	1.05	(0.70-1.57)	0.80
	3	0.98	(0.67-1.43)	1.15	(0.76-1.74)	0.52
	4	1.07	(0.71-1.61)	1.82	(1.14-2.90)	0.01
	5 (most disadvantaged)	0.40	(0.27-0.59)	1.14	(0.71-1.85)	0.58

Exposure variable	IMD Quintile	Univariate		Multivariable[a]		p value
		Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Walked in paddock	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.50	(1.05-2.14)	1.52	(1.04-2.20)	0.03
	3	1.28	(0.89-1.85)	1.34	(0.91-1.96)	0.14
	4	0.76	(0.50-1.15)	1.08	(0.70-1.69)	0.72
	5 (most disadvantaged)	0.30	(0.17-0.51)	0.67	(0.38-1.19)	0.17
Visited a farm	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.22	(0.84-1.78)	1.18	(0.80-1.73)	0.41
	3	0.92	(0.62-1.37)	0.89	(0.59-1.34)	0.58
	4	0.54	(0.34-0.86)	0.54	(0.34-0.88)	0.01
	5 (most disadvantaged)	0.63	(0.39-0.99)	0.86	(0.51-1.43)	0.55
Day trip	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.15	(0.81-1.65)	1.19	(0.83-1.71)	0.35
	3	1.10	(0.76-1.58)	1.13	(0.78-1.64)	0.52
	4	0.77	(0.51-1.16)	0.79	(0.52-1.20)	0.27
	5 (most disadvantaged)	0.63	(0.41-0.97)	0.73	(0.46-1.17)	0.19
Contact with soil	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	1.44	(0.99-2.07)	1.40	(0.96-2.03)	0.08
	3	0.96	(0.65-1.42)	0.96	(0.64-1.42)	0.83
	4	0.89	(0.59-1.36)	1.13	(0.73-1.75)	0.58
	5 (most disadvantaged)	0.64	(0.40-1.00)	1.05	(0.64-1.74)	0.85
UK travel	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)		
	2	0.75	(0.51-1.09)	0.79	(0.54-1.15)	0.22
	3	0.76	(0.52-1.11)	0.81	(0.55-1.18)	0.27
	4	0.86	(0.58-1.28)	0.94	(0.62-1.43)	0.79
	5 (most disadvantaged)	0.51	(0.33-0.80)	0.69	(0.42-1.12)	0.13

[a]Adjusted for age group, sex, ethnicity and rurality

7.5 Results – socio-demographic risk factors in development of HUS

Characteristics of participants

A total of 1059 paediatric STEC cases were included in the study. Table 7.12 shows the population stratified by IMD quintile. Of the 1059 paediatric STEC cases included in the study, 207 (19.5%) developed HUS as recorded by the BPSU study and/or via the ESQ. In the least disadvantaged quintile the progression of STEC cases to HUS was 19.2% (47/245) compared with 15.3% (29/189) in the most disadvantaged quintile although this difference was non-significant (p=0.27). Progression to HUS varied by age and gender (Table 7.12, Figure 7.1). A higher proportion of progression to HUS was observed in girls aged less than 1 year compared to boys of the same age (14.3% compared to 4.8%) but this was non-significant, and amongst girls aged 10-15 years compared to boys of the same age (19.3% compared to 7.1%); a statistically significant difference (p=0.01). The highest proportion of progression was observed in girls aged 1-4 years (26.0%). Vomiting and s*tx* gene were independently associated with SES.

There was no difference in missing ethnicity by sex, however there were some differences by age group (57.3% of missing ethnicity in 1-4 age group; n=114/199) and region (31.2% of missing ethnicity in London; n=62/199); these were not regarded as problematic as, given the observed data for other variables, the missing data are considered independent

Table 7.12: Characteristics of cohort participants by Index of Multiple Deprivation quintile (n=1,059)

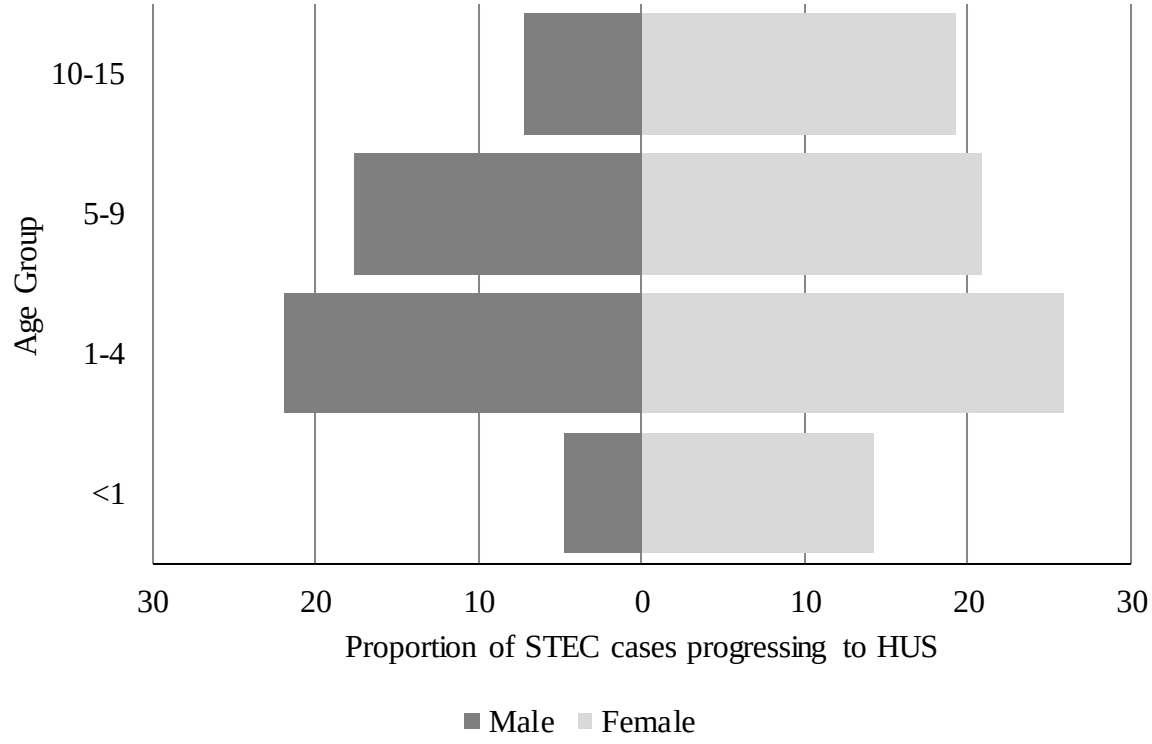
	Q1 (least disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (most disadvantaged) n (%)	p value[a]
Total	245 (23.1)	221 (20.9)	219 (20.7)	185 (17.5)	189 (17.9)	0.05
HUS (BPSU and/or ESQ)	47 (19.2)	35 (15.8)	53 (24.2)	43 (23.2)	29 (15.3)	0.96
Demographics						
Age group						
<1	19 (7.8)	17 (7.7)	9 (4.1)	8 (4.3)	17 (9.0)	0.18
1-4	104 (42.5)	96 (43.4)	103 (47.0)	90 (48.7)	93 (49.2)	
5-9	65 (26.5)	67 (30.3)	64 (29.2)	54 (29.2)	46 (24.3)	
10-15	57 (23.3)	41 (18.6)	43 (19.6)	33 (17.8)	33 (17.5)	
Sex						
Female	111 (45.3)	119 (53.9)	117 (53.4)	79 (42.7)	90 (47.6)	0.67
Male	134 (54.7)	102 (46.2)	102 (46.6)	106 (57.3)	99 (52.4)	
Ethnicity						
White	185 (75.5)	158 (71.5)	150 (68.5)	121 (65.4)	72 (38.1)	<0.001
Non-white	14 (5.7)	19 (8.6)	15 (6.9)	26 (14.1)	82 (43.4)	
Unknown	46 (18.8)	44 (19.9)	54 (24.7)	38 (20.5)	35 (18.5)	
Travel						
Yes	37 (15.1)	28 (12.7)	27 (12.3)	23 (12.4)	35 (18.5)	0.46
Rurality						
Rural	84 (34.3)	101 (45.7)	79 (36.1)	20 (10.8)	2 (1.1)	<0.001
Urban	161 (65.7)	120 (54.3)	140 (63.9)	165 (89.2)	187 (98.9)	
Region						
East Midlands	24 (9.8)	21 (9.5)	17 (7.8)	8 (4.3)	10 (5.3)	0.79
East of England	22 (9.0)	24 (10.9)	13 (5.9)	7 (3.8)	5 (2.7)	
London	6 (2.5)	16 (7.2)	21 (9.6)	29 (15.7)	42 (22.2)	
North East	15 (6.1)	15 (6.8)	8 (3.7)	29 (15.7)	16 (8.5)	
North West	51 (20.8)	39 (17.7)	34 (15.5)	34 (18.4)	39 (20.6)	
South East	52 (21.2)	22 (10.0)	27 (12.3)	7 (3.8)	9 (4.8)	
South West	26 (10.6)	28 (12.7)	50 (22.8)	21 (11.4)	8 (4.2)	
West Midlands	21 (8.6)	26 (11.8)	18 (8.2)	28 (15.1)	21 (11.1)	
Yorkshire and Humber	28 (11.4)	30 (13.6)	31 (14.2)	22 (11.9)	39 (20.6)	

	Q1 (least disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (most disadvantaged) n (%)	p value [a]
Stx						
*Stx*1	3 (1.2)	1 (0.5)	3 (1.4)	3 (1.6)	8 (4.2)	0.04
*Stx*2	160 (65.3)	158 (71.5)	160 (73.1)	133 (71.9)	134 (70.9)	
*Stx*1+2	68 (27.8)	51 (23.1)	41 (18.7)	34 (18.4)	32 (16.9)	
Serology	14 (5.7)	10 (4.5)	15 (6.9)	14 (7.6)	13 (6.9)	
Unknown	-	1 (0.5)	-	1 (0.5)	2 (1.1)	
Symptoms						
Diarrhoea	232 (94.7)	208 (94.1)	209 (95.4)	175 (94.6)	176 (93.1)	0.62
Bloody diarrhoea	134 (54.7)	120 (54.3)	138 (63.0)	105 (56.8)	87 (46.0)	0.23
Nausea	83 (33.9)	73 (33.0)	80 (36.5)	64 (34.6)	67 (35.5)	0.64
Vomiting	104 (42.5)	105 (47.5)	102 (46.6)	85 (46.0)	103 (54.5)	0.04
Abdominal pain	177 (72.2)	151 (68.3)	161 (73.5)	122 (66.0)	123 (65.1)	0.11
Fever	84 (34.3)	73 (33.0)	70 (32.0)	55 (29.7)	74 (39.2)	0.59
Healthcare contact						
Antibiotics	32 (13.1)	29 (13.1)	30 (13.7)	20 (10.8)	19 (10.1)	0.27
NHS Direct	21 (8.6)	24 (10.9)	24 (11.0)	14 (7.6)	10 (5.3)	0.15
GP	161 (65.7)	150 (67.9)	144 (65.8)	112 (60.5)	114 (60.3)	0.10
A&E	57 (23.3)	52 (23.5)	59 (26.9)	54 (29.2)	56 (29.6)	0.06
Other healthcare contact	29 (11.8)	27 (12.2)	28 (12.8)	20 (10.8)	27 (14.3)	0.63
Hospital	90 (36.7)	85 (38.5)	98 (44.8)	83 (44.9)	70 (37.0)	0.44

[a] Statistical significance of relationship between IMD quintile and each variable, tested using χ^2 test

HUS – haemolytic uraemic syndrome; BPSU – British Paediatric Surveillance Unit; ESQ – enhanced surveillance questionnaire; *stx* – shiga toxin gene; NHS Direct – National Health Service telephone advice line, now NHS 111; GP – General Practitioner; A&E – accident and emergency

Figure 7.1: Proportions of STEC cases progressing to HUS by age and gender (n=1,059)



STEC – Shiga toxin-producing *Escherichia* coli; HUS – haemolytic uraemic syndrome

Overall crude incidence was lower among the most disadvantaged compared to the least disadvantaged (Table 7.13); this pattern was also reflected in the age- and sex-adjusted incidence rates although the incidence rate ratios were non-significant.

Table 7.13: Incidence rate ratio for exposed (most disadvantaged) compared to unexposed (least disadvantaged) - HUS

	Rate/100,000 in most disadvantaged	Rate/100,000 in least disadvantaged	Incidence rate difference	Incidence rate ratio	95% CI
Overall	0.41	0.86	-0.45	0.48	0.21-1.08
Female	0.48	1.03	-0.55	0.46	0.16-1.36
Male	0.34	0.69	-0.35	0.50	0.14-1.75
0-4	0.77	1.53	-0.76	0.50	0.18-1.38
5-9	0.32	0.68	-0.36	0.48	0.09-2.40
10-14	0.05	0.42	-0.37	0.12	0.00-4.25

NB: population data only available by IMD, age and gender for age groups up to 14 years of age

Main analysis

For the purposes of statistical analysis we excluded cases identified via serological testing only (n=66) or for whom no microbiological information was available (n=4); 989 STEC cases remained, of which 144 (15%) developed HUS. Ethnicity was missing for 20.1% of cases (n=199); ethnicity was imputed for these cases.

In univariate regression, age group, travel, s*tx2* only, antibiotics, diarrhoea, bloody diarrhoea, nausea, vomiting and fever were associated with higher odds of development of HUS.

In the fully adjusted model (Table 7.14), the most disadvantaged children had (non-significantly) lower odds of progression to HUS compared to the least disadvantaged children (OR 0.57, 95% CI 0.25-1.31). This analysis also identified significantly lower odds amongst <1, 5-9 and 10-15 year olds compared to 1-4 year olds (OR 0.21, 95% CI 0.05-0.82; OR 0.43, 95% CI 0.25-0.74; and OR 0.20, 95% CI 0.09-0.43 respectively) and significantly higher odds amongst those infected with s*tx*2-only (OR 5.92, 95% CI 2.49-14.10), prescribed antibiotics (OR 8.46, 95% CI 4.71-15.18) and among those who had experienced bloody diarrhoea (OR 3.56, 95% CI 2.04-6.24) or vomiting (OR 4.47, 95% CI 2.62-7.63). There were no significant interactions identified.

Table 7.14: Univariate and multivariable regression analysis - HUS (n=989)

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
IMD Quintile	1 (least disadvantaged)	231 (23.4)	1.0 (reference)		1.0 (reference)		
	2	210 (21.2)	0.83	(0.48-1.42)	0.64	(0.32-1.27)	0.20
	3	204 (20.6)	1.28	(0.77-2.12)	1.01	(0.54-1.91)	0.97
	4	170 (17.2)	1.10	(0.64-1.90)	1.10	(0.54-2.26)	0.79
	5 (most disadvantaged)	174 (17.6)	0.57	(0.30-1.06)	0.57	(0.25-1.31)	0.18
Age group	<1	67 (6.8)	0.19	(0.06-0.62)	**0.21**	**(0.05-0.82)**	**0.03**
	1-4	456 (46.1)	1.0 (reference)		1.0 (reference)		
	5-9	274 (27.7)	0.62	(0.40-0.94)	**0.43**	**(0.25-0.74)**	**0.002**
	10-15	192 (19.4)	0.34	(0.19-0.61)	**0.20**	**(0.09-0.43)**	**<0.001**
Sex	Male	513 (51.9)	1.0 (reference)		1.0 (reference)		
	Female	476 (48.1)	1.37	(0.96-1.96)	1.38	(0.88-2.14)	0.16
Ethnicity	White	797 (80.6)	1.0 (reference)		1.0 (reference)		
	Non-White	192 (19.4)	0.39	(0.18-0.81)	**0.28**	**(0.11-0.74)**	**0.01**
Travel	No	850 (86.0)	1.0 (reference)		1.0 (reference)		
	Yes	139 (14.0)	0.46	(0.24-0.88)	0.64	(0.28-1.45)	0.28
Rurality	Urban	719 (72.7)	1.0 (reference)		1.0 (reference)		
	Rural	270 (27.3)	1.21	(0.82-1.77)	0.88	(0.52-1.48)	0.63
Region	East Midlands	72 (7.3)	0.62	(0.24-1.59)	0.59	(0.18-1.92)	0.39
	East of England	66 (6.7)	1.03	(0.44-2.42)	1.12	(0.37-3.37)	0.84
	London	108 (10.9)	1.0 (reference)		1.0 (reference)		
	North East	76 (7.7)	1.19	(0.53-2.64)	0.71	(0.26-1.97)	0.51
	North West	185 (18.7)	1.20	(0.63-2.31)	1.02	(0.44-2.37)	0.97
	South East	107 (10.8)	1.09	(0.52-2.28)	1.31	(0.48-3.63)	0.60
	South West	127 (12.8)	1.48	(0.75-2.93)	1.25	(0.50-3.13)	0.63
	West Midlands	104 (10.5)	0.54	(0.23-1.29)	0.53	(0.18-1.53)	0.24
	Yorkshire and Humber	144 (14.6)	0.62	(0.29-1.33)	0.52	(0.20-1.34)	0.17
Stx	*Stx*1+2	226 (22.9)	1.0 (reference)		1.0 (reference)		
	*Stx*1	18 (1.8)	1.84	(0.21-15.84)	5.53	(0.53-57.42)	0.15
	*Stx*2	745 (75.3)	6.99	(3.22-15.17)	**5.92**	**(2.49-14.10)**	**<0.001**
Antibiotics	No	887 (89.7)	1.0 (reference)		1.0 (reference)		
	Yes	102 (10.3)	8.54	(5.48-13.30)	**8.46**	**(4.71-15.18)**	**<0.001**

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Diarrhoea	No	49 (5.0)	1.0 (reference)		1.0 (reference)		
	Yes	940 (95.0)	8.61	(1.18-62.89)	4.04	(0.50-32.59)	0.19
Bloody diarrhoea	No	440 (44.5)	1.0 (reference)		1.0 (reference)		
	Yes	549 (55.5)	4.85	(3.07-8.00)	**3.56**	**(2.04-6.24)**	**<0.001**
Nausea	No	653 (66.0)	1.0 (reference)		1.0 (reference)		
	Yes	336 (34.0)	1.52	(1.06-2.18)	1.12	(0.67-1.86)	0.66
Vomiting	No	549 (55.5)	1.0 (reference)		1.0 (reference)		
	Yes	440 (44.5)	6.05	(3.95-9.26)	**4.47**	**(2.62-7.63)**	**<0.001**
Abdominal pain	No	309 (31.2)	1.0 (reference)		1.0 (reference)		
	Yes	680 (68.8)	1.49	(0.99-2.25)	0.82	(0.46-1.46)	0.50
Fever	No	657 (66.4)	1.0 (reference)		1.0 (reference)		
	Yes	332 (33.6)	1.50	(1.05-2.16)	1.05	(0.67-1.66)	0.82

aAdjusted for all other covariates in the model; s*tx* – shiga toxin gene

Sensitivity analyses

Multiple sensitivity analyses were conducted to assess the robustness of the findings; in particular to explore the age and sex differences and assess the potential collinearity between IMD and ethnicity. These analyses did not alter the overall conclusions of this research; namely there was no significant difference in odds of progression to HUS by SEC.

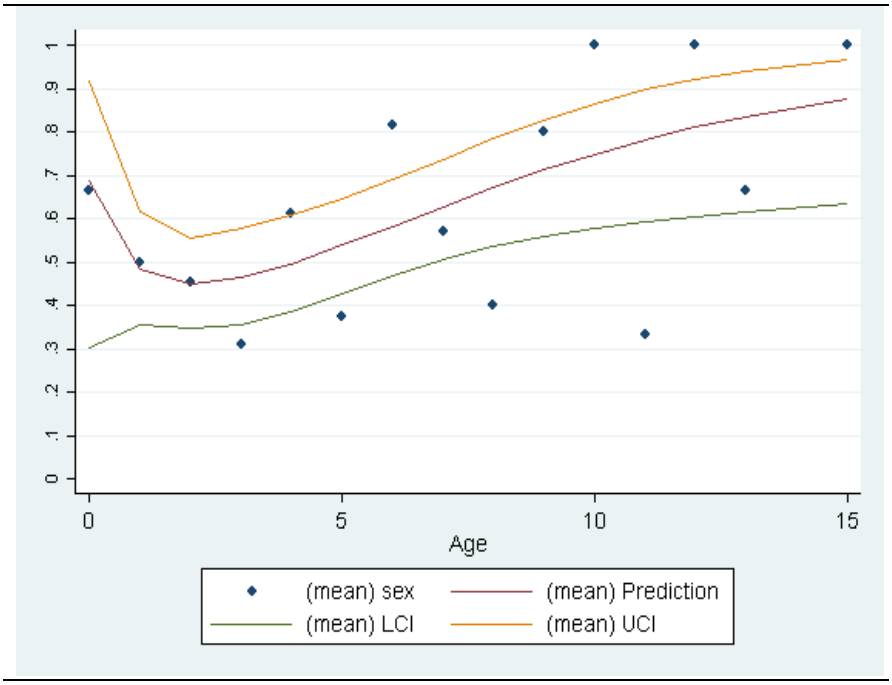
As no significant interaction was observed between age and sex in the main analysis (Table 7.14) despite the observed differences in HUS progression in descriptive analysis (Table 7.12), fractional polynomial prediction plots were fitted to explore this relationship further (Figure 7.2). The fractional polynomial prediction plots by age and sex show no evidence for non-linearity and a likelihood ratio test performed on the HUS cases suggested that the fractional polynomial model was not an improvement on the linear model ($p=0.19$) which further supports the absence of evidence for a non-linear distribution.

Sensitivity analyses excluding children thought to have acquired their infection through travel to foreign countries (Table 7.15) did not differ from the primary analysis, suggesting that travel outside of the UK is not a significant factor in HUS development in the paediatric STEC population.

Sensitivity analyses excluding the ethnicity variable due to the high level of missing data found significantly lower odds of progression to HUS amongst the most disadvantaged children compared to the least disadvantaged children (Table 7.16; OR 0.37, 95% CI 0.17-0.79). To test whether this was due to collinearity between IMD and ethnicity, a post-hoc matched analysis on ethnicity was performed (Table 7.17). This analysis produced similar results to the main model suggesting that this was not problematic for the main analysis. For further reassurance, a penalised logistic regression was performed on the multiply imputed dataset (Table 7.18). This analysis also produced similar results to the main model, again suggesting that collinearity was not an issue in this analysis.

Figure 7.2: Fractional polynomial prediction plots for age and sex by HUS Status

HUS



Non-HUS

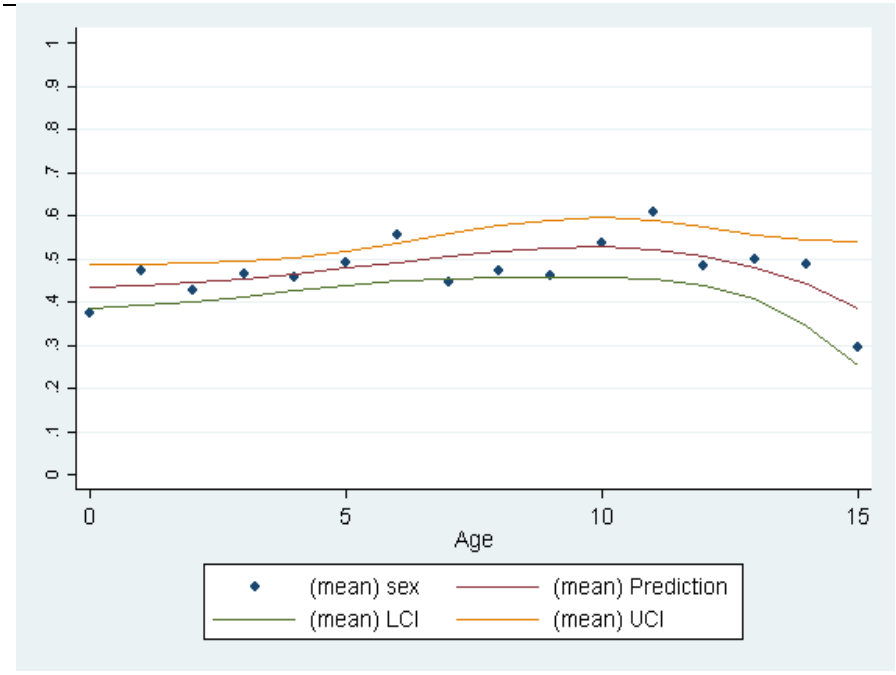


Table 7.15: Univariate and multivariable regression analysis - Sensitivity analysis excluding travel cases (n subjects=850)

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
IMD Quintile	1 (least disadvantaged)	196 (23.1)	1.0 (reference)		1.0 (reference)		
	2	184 (21.7)	0.88	(0.50-1.54)	0.75	(0.36-1.53)	0.43
	3	180 (21.2)	1.24	(0.73-2.10)	1.06	(0.54-2.10)	0.86
	4	150 (17.7)	1.07	(0.61-1.90)	1.14	(0.53-2.46)	0.74
	5 (most disadvantaged)	140 (16.5)	0.52	(0.26-1.04)	0.57	(0.23-1.43)	0.23
Age group	<1	55 (6.5)	0.14	(0.03-0.59)	**0.12**	**(0.02-0.62)**	**0.01**
	1-4	400 (47.1)	1.0 (reference)		1.0 (reference)		
	5-9	242 (28.5)	0.61	(0.39-0.94)	**0.45**	**(0.25-0.79)**	**0.005**
	10-15	153 (18.0)	0.32	(0.17-0.60)	**0.17**	**(0.07-0.39)**	**<0.001**
Sex	Male	445 (52.4)	1.0 (reference)		1.0 (reference)		
	Female	405 (47.7)	1.31	(0.91-1.90)	1.30	(0.81-2.08)	0.27
Ethnicity	White	700 (82.4)	1.0 (reference)		1.0 (reference)		
	Non-White	150 (17.6)	0.41	(0.20-0.83)	**0.27**	**(0.10-0.73)**	**0.01**
Rurality	Urban	606 (71.3)	1.0 (reference)		1.0 (reference)		
	Rural	244 (28.7)	1.28	(0.86-1.90)	1.07	(0.62-1.86)	0.80
Region	East Midlands	61 (7.2)	0.54	(0.19-1.49)	0.49	(0.14-1.80)	0.29
	East of England	55 (6.5)	0.84	(0.33-2.16)	1.02	(0.30-3.44)	0.98
	London	83 (9.8)	1.0 (reference)		1.0 (reference)		
	North East	71 (8.4)	0.90	(0.38-2.14)	0.59	(0.19-1.81)	0.35
	North West	172 (20.2)	1.13	(0.56-2.25)	1.06	(0.42-2.67)	0.90
	South East	73 (8.6)	1.27	(0.58-2.86)	1.78	(0.57-5.56)	0.32
	South West	110 (12.9)	1.38	(0.66-2.86)	1.14	(0.42-3.11)	0.80
	West Midlands	95 (11.2)	0.52	(0.21-1.26)	0.57	(0.18-1.78)	0.33
	Yorkshire and Humber	130 (15.3)	0.59	(0.27-1.32)	0.53	(0.19-1.47)	0.22
Stx	*Stx*1+2	183 (21.5)	1.0 (reference)		1.0 (reference)		
	*Stx*1	8 (0.9)	4.21	(0.45-39.89)	**24.71**	**(1.86-328.34)**	**0.02**
	*Stx*2	659 (77.5)	6.97	(3.02-16.10)	**6.08**	**(2.32-15.92)**	**<0.001**
Antibiotics	No	766 (90.1)	1.0 (reference)		1.0 (reference)		
	Yes	84 (9.9)	10.0	(6.18-16.32)	**10.89**	**(5.65-20.97)**	**<0.001**
Diarrhoea	No	48 (6.6)	1.0 (reference)		1.0 (reference)		
	Yes	802 (94.4)	9.26	(1.27-67.70)	4.00	(0.48-33.16)	0.20

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Bloody diarrhoea	No	360 (42.4)	1.0 (reference)		1.0 (reference)		
	Yes	490 (57.7)	5.10	(3.10-8.38)	**3.70**	**(2.00-6.85)**	**<0.001**
Nausea	No	562 (66.1)	1.0 (reference)		1.0 (reference)		
	Yes	288 (33.9)	1.52	(1.04-2.22)	1.03	(0.60-1.76)	0.92
Vomiting	No	465 (54.7)	1.0 (reference)		1.0 (reference)		
	Yes	385 (45.3)	6.50	(4.13-10.23)	**5.25**	**(2.94-9.38)**	**<0.001**
Abdominal pain	No	259 (30.5)	1.0 (reference)		1.0 (reference)		
	Yes	591 (69.5)	1.39	(0.91-2.13)	0.73	(0.39-1.36)	0.32
Fever	No	569 (66.9)	1.0 (reference)		1.0 (reference)		
	Yes	281 (33.1)	1.71	(1.17-2.50)	1.28	(0.79-2.09)	0.32

aAdjusted for all other covariates in the model; *stx* – shiga toxin gene

Table 7.16: Univariate and multivariable regression analysis - Sensitivity analysis excluding ethnicity variable (n subjects=989)

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
IMD Quintile	1 (least disadvantaged)	231 (23.4)	1.0 (reference)		1.0 (reference)		
	2	210 (21.2)	0.83	(0.48-1.42)	0.65	(0.33-1.27)	0.21
	3	204 (20.6)	1.28	(0.77-2.12)	0.98	(0.51-1.79)	0.89
	4	170 (17.2)	1.10	(0.64-1.89)	0.96	(0.48-1.94)	0.91
	5 (most disadvantaged)	174 (17.6)	0.57	(0.30-1.06)	**0.41**	**(0.19-0.90)**	**0.03**
Age group	<1	67 (6.8)	0.19	(0.06-0.62)	**0.24**	**(0.06-0.92)**	**0.04**
	1-4	456 (46.1)	1.0 (reference)		1.0 (reference)		
	5-9	274 (27.7)	0.62	(0.40-0.94)	**0.45**	**(0.26-0.76)**	**0.003**
	10-15	192 (19.4)	0.34	(0.19-0.61)	**0.22**	**(0.11-0.47)**	**<0.001**
Sex	Male	513 (51.9)	1.0 (reference)		1.0 (reference)		
	Female	476 (48.1)	1.37	(0.96-1.96)	1.40	(0.91-2.16)	0.13
Travel	No	850 (86.0)	1.0 (reference)		1.0 (reference)		
	Yes	139 (14.0)	0.46	(0.24-0.88)	0.58	(0.26-1.28)	0.18
Rurality	Urban	719 (72.7)	1.0 (reference)		1.0 (reference)		
	Rural	270 (27.3)	1.21	(0.82-1.77)	0.97	(0.58-1.62)	0.90
Region	East Midlands	72 (7.3)	0.62	(0.24-1.59)	0.65	(0.21-2.02)	0.45
	East of England	66 (6.7)	1.03	(0.44-2.42)	1.37	(0.47-4.00)	0.56
	London	108 (10.9)	1.0 (reference)		1.0 (reference)		
	North East	76 (7.7)	1.19	(0.53-2.64)	1.02	(0.38-2.71)	0.97
	North West	185 (18.7)	1.20	(0.63-2.31)	1.29	(0.57-2.90)	0.54
	South East	107 (10.8)	1.09	(0.52-2.28)	1.74	(0.65-4.63)	0.27
	South West	127 (12.8)	1.48	(0.75-2.93)	1.64	(0.68-3.99)	0.27
	West Midlands	104 (10.5)	0.54	(0.23-1.29)	0.67	(0.24-1.86)	0.44
	Yorkshire and Humber	144 (14.6)	0.62	(0.29-1.33)	0.60	(0.24-1.52)	0.28
Stx	*Stx*1+2	226 (22.9)	1.0 (reference)		1.0 (reference)		
	*Stx*1	18 (1.8)	1.84	(0.21-15.84)	5.34	(0.54-52.82)	0.15
	*Stx*2	745 (75.3)	6.99	(3.22-15.17)	**5.76**	**(2.43-13.67)**	**<0.001**
Antibiotics	No	887 (89.7)	1.0 (reference)		1.0 (reference)		
	Yes	102 (10.3)	8.54	(5.48-13.30)	**7.46**	**(4.27-13.03)**	**<0.001**
Diarrhoea	No	49 (5.0)	1.0 (reference)		1.0 (reference)		
	Yes	940 (95.0)	8.61	(1.18-62.89)	4.09	(0.51-32.47)	0.18

Variable	Category	n (%)	Univariate		Multivariable[a]		p value
			Odds Ratio	(95% CI)	Odds Ratio	(95% CI)	
Bloody diarrhoea	No	440 (44.5)	1.0 (reference)		1.0 (reference)		
	Yes	549 (55.5)	4.85	(3.07-7.67)	**3.74**	**(2.15-6.49)**	**<0.001**
Nausea	No	653 (66.0)	1.0 (reference)		1.0 (reference)		
	Yes	336 (34.0)	1.52	(1.06-2.18)	1.11	(0.67-1.83)	0.69
Vomiting	No	549 (55.5)	1.0 (reference)		1.0 (reference)		
	Yes	440 (44.5)	6.05	(3.95-9.26)	**4.38**	**(2.59-7.40)**	**<0.001**
Abdominal pain	No	309 (31.2)	1.0 (reference)		1.0 (reference)		
	Yes	680 (68.8)	1.49	(0.99-2.25)	0.83	(0.47-1.46)	0.52
Fever	No	657 (66.4)	1.0 (reference)		1.0 (reference)		
	Yes	332 (33.6)	1.50	(1.05-2.16)	1.04	(0.66-1.63)	0.86

aAdjusted for all other covariates in the model; s*tx* – shiga toxin gene

Table 7.17: Comparison between logistic regression model and post-hoc matched analysis on ethnicity

Variable	Category	Main model[a]		Post-hoc matched analysis[a]	
		Odds Ratio	**(95% CI)**	**Odds Ratio**	**(95% CI)**
IMD Quintile	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)	
	2	0.64	(0.32-1.27)	0.93	(0.44-2.15)
	3	1.01	(0.54-1.91)	1.40	(0.63-3.30)
	4	1.10	(0.54-2.26)	1.39	(0.58-3.82)
	5 (most disadvantaged)	0.57	(0.25-1.31)	0.63	(0.23-1.70)
Age group	<1	**0.21**	**(0.05-0.82)**	0.40	(0.13-0.90)
	1-4	1.0 (reference)		1.0 (reference)	
	5-9	**0.43**	**(0.25-0.74)**	0.66	(0.31-1.35)
	10-15	**0.20**	**(0.09-0.43)**	0.28	(0.01-1.01)
Sex	Male	1.0 (reference)		1.0 (reference)	
	Female	1.38	(0.88-2.14)	1.76	(0.96-3.34)
Travel	No	1.0 (reference)		1.0 (reference)	
	Yes	0.64	(0.28-1.45)	0.71	(0.24-1.99)
Stx	*Stx*1+2	1.0 (reference)		1.0 (reference)	
	*Stx*2	**5.92**	**(2.49-14.10)**	**14.37**	**(6.66-67.74)**

aAdjusted for all other covariates in the model; *stx* – shiga toxin gene

Table 7.18: Comparison between logistic regression model and penalised logistic regression model

Variable	Category	Main model[a]		Penalised logistic regression[a]	
		Odds Ratio	**(95% CI)**	**Odds Ratio**	**(95% CI)**
IMD Quintile	1 (least disadvantaged)	1.0 (reference)		1.0 (reference)	
	2	0.64	(0.32-1.27)	0.64	(0.32-1.27)
	3	1.01	(0.54-1.91)	1.01	(0.54-1.91)
	4	1.10	(0.54-2.26)	1.09	(0.53-2.23)
	5 (most disadvantaged)	0.57	(0.25-1.31)	0.55	(0.24-1.26)
Age group	<1	**0.21**	**(0.05-0.82)**	**0.21**	**(0.05-0.82)**
	1-4	1.0 (reference)		1.0 (reference)	
	5-9	**0.43**	**(0.25-0.74)**	**0.20**	**(0.09-0.44)**
	10-15	**0.20**	**(0.09-0.43)**	**0.43**	**(0.25-0.74)**
Sex	Male	1.0 (reference)		1.0 (reference)	
	Female	1.38	(0.88-2.14)	1.37	(0.88-2.13)
Ethnicity	White	1.0 (reference)		1.0 (reference)	
	Non-White	**0.28**	**(0.11-0.74)**	**0.32**	**(0.11-0.96)**
Travel	No	1.0 (reference)		1.0 (reference)	
	Yes	0.64	(0.28-1.45)	0.64	(0.28-1.44)
Rurality	Urban	1.0 (reference)		1.0 (reference)	
	Rural	0.88	(0.52-1.48)	0.88	(0.52-1.49)
Region	East Midlands	0.59	(0.18-1.92)	0.59	(0.18-1.91)
	East of England	1.12	(0.37-3.37)	1.16	(0.39-3.47)
	London	1.0 (reference)		1.0 (reference)	
	North East	0.71	(0.26-1.97)	0.73	(0.26-2.05)
	North West	1.02	(0.44-2.37)	1.03	(0.44-2.41)
	South East	1.31	(0.48-3.63)	1.34	(0.48-3.73)
	South West	1.25	(0.50-3.13)	1.28	(0.51-3.21)
	West Midlands	0.53	(0.18-1.53)	0.54	(0.91-1.56)
	Yorkshire and Humber	0.52	(0.20-1.34)	0.52	(0.20-1.35)
Stx	*Stx*1+2	1.0 (reference)		1.0 (reference)	
	*Stx*1	5.53	(0.53-57.42)	5.58	(0.54-57.77)
	*Stx*2	**5.92**	**(2.49-14.10)**	**5.89**	**(2.38-14.02)**
Antibiotics	No	1.0 (reference)		1.0 (reference)	
	Yes	**8.46**	**(4.71-15.18)**	**8.36**	**(4.65-15.03)**

Variable	Category	Main model[a]		Penalised logistic regression[a]	
		Odds Ratio	**(95% CI)**	**Odds Ratio**	**(95% CI)**
Diarrhoea	No	1.0 (reference)		1.0 (reference)	
	Yes	4.04	(0.50-32.59)	3.97	(0.49-31.93)
Bloody diarrhoea	No	1.0 (reference)		1.0 (reference)	
	Yes	**3.56**	**(2.04-6.24)**	**3.54**	**(2.02-6.21)**
Nausea	No	1.0 (reference)		1.0 (reference)	
	Yes	1.12	(0.67-1.86)	1.12	(0.67-1.86)
Vomiting	No	1.0 (reference)		1.0 (reference)	
	Yes	**4.47**	**(2.62-7.63)**	**4.48**	**(2.63-7.64)**
Abdominal pain	No	1.0 (reference)		1.0 (reference)	
	Yes	0.82	(0.46-1.46)	0.81	(0.46-1.45)
Fever	No	1.0 (reference)		1.0 (reference)	
	Yes	1.05	(0.67-1.66)	0.81	(0.67-1.66)

aAdjusted for all other covariates in the model; *stx* – shiga toxin gene

7.6 Discussion – Social patterning of clinical outcomes, healthcare contact and risk factors for STEC

In this cross-sectional observational analysis of STEC cases reported to NESSS over a six-year period, the social patterning of risk factors for STEC infection using IMD as an area-level measure of SES was explored.

Crude incidence of STEC was significantly lower in the most disadvantaged compared to the least disadvantaged. This also echoes the findings from Chapter 5, which suggested that the incidence of IID in the community was lower in the most disadvantaged group. Descriptive analysis suggested a higher proportion of those in the most disadvantaged quintile being infected with *stx*2, associated with more severe disease and progression to HUS, however the relationship between SES and reporting of HUS and symptoms was not significant. There were differences in healthcare contact by SES, with lower odds of reporting contact with GP and higher odds of reporting visiting A&E or being hospitalised amongst the most disadvantaged group. There were also differences in exposure to known risk factors by SES; with lower odds of reporting exposure to salad, fruit, vegetables or herbs; recreational freshwater; walking in a paddock; day trips; contact with soil; UK travel and; foreign travel amongst the most disadvantaged group. Lower odds of reported exposure to fish/shellfish were also identified; although this is asked about in the ESQ it is not widely regarded as a risk factor for STEC and likely to reflects population-level exposure.

As this is an observational study it is not possible to determine causation. Therefore the associations presented here need to be interpreted with caution. Furthermore, without accurate population-level comparators for social patterning of the risk factors included within this study, it is possible that the differences detected may reflect the patterning of these risk factors in the general population. As exposure to a private water supply is relatively rare, despite it being a known STEC risk factor, it was not possible to perform a multivariable analysis for this due to very small numbers in some covariates. Finally, the lack of significant findings for risk factors in children aged less than 16 years is likely to be due to low study power.

Despite this, this is a novel analysis of data collected through a country-wide, representative surveillance system to explore the social patterning of risk factors

amongst STEC cases in England. This study captures extensive risk factor and exposure data in a well-characterised STEC population. To the best of my knowledge, following a review of the literature and discussion with national experts, this is the first study to explore the social patterning of risk factors for STEC in England. Infection with STEC is rare but potentially very serious infection and, particularly in children and the elderly, is likely to result in interaction with healthcare services. It is therefore likely that NESSS captures data on a high proportion of STEC cases in England and hence this study is likely to be representative of STEC cases nationally.

Most disadvantaged individuals had a lower crude incidence rate of STEC infection compared to least disadvantaged and this was consistent for men and women as well as all age groups with the exception of those aged 0-4 years who had a non-significant difference in incidence rate ratio. This finding has also been seen in other studies. Chang et al. (2009) and Jalava et al. (2011) found lower incidence in communities with lower education levels and hypothesise that this could relate to poorer food safety practices and differential food consumption habits among more highly educated individuals. Jalava et al. (2011) also identified increased incidence associated with the proportion of low income households with children and suggest this could relate to transmission of infection, suggesting that the relationship between SES and GI infections across the life course is important. A study by Sakuma et al. (2006) found that areas with lower average income were positively associated with STEC infection although this was an ecological study, and studies by Simonsen et al. (2008) and Pearl et al. (2009) found no significant difference although Simonsen et al. (2008) suggest this was an artefact of small numbers in their study.

In other studies which found lower risk amongst disadvantaged individuals for other GI pathogens researchers have hypothesised that higher risk of GI infections in individuals of higher socioeconomic status could relate to; differential exposure such as consumption of higher risk food amongst those of a higher SES (Patil et al., 2005, Simonsen et al., 2008, Tarr et al., 2005, Taylor, 2012, Whitney et al., 2015, Younus et al., 2007); more foreign travel; greater exposure to farm animals or the rural environment; or it could be related to differential healthcare interaction (Simonsen et al., 2008, Younus et al., 2007).

In our study, the most disadvantaged individuals were less likely to visit their GP but more likely to visit A&E and be hospitalised compared to the least disadvantaged. Despite limited research on healthcare interaction for GI infections, there is some evidence to suggest that, for general ill-health, disadvantaged individuals are generally more likely to interact with all three of these healthcare services; however more advantaged individuals are also more likely to seek care at an earlier stage of disease progression (Cookson et al., 2016). There is also evidence to suggest that, amongst individuals with GI infections, those who are more disadvantaged are more likely to be hospitalised (Olowokure et al., 1999, Rose, 2017). The discrepancy at the GP level could suggest that accessing GP care varies by type of illness and SES. It could suggest differences in healthcare interaction amongst disadvantaged STEC cases in terms of interaction with or to access to healthcare services or potentially recognition of symptoms and seeking care at a more advanced stage of illness, as observed for general ill health (Cookson et al., 2016). Further studies would be required to robustly assess these hypotheses.

One study was identified which explored the social patterning of a limited number of risk factors for STEC (consumption of hamburgers, type of water available for consumption, recreational water exposure and personal hygiene) in Argentina (Bentancor et al., 2012). The study looked at young school children but did not assess STEC infection within this study population and risk factors in Argentina may differ from those in the UK. This study identified significantly higher consumption of hamburgers in higher socioeconomic groups but no significant difference in consumption of drinking water or exposure to recreational water was identified. Of interest, there was a significant difference in hand hygiene by SES, with children from low socioeconomic groups washing their hands more frequently compared to those from other socioeconomic groups and a higher proportion of children from low socioeconomic who reported washing their hands before eating compared to other socioeconomic groups.

It is possible that the pattern of lower odds of reporting foreign travel amongst more disadvantaged groups reflects a population-level pattern however travel is a known risk factor for STEC infection (Byrne et al., 2015, Parry et al., 1998); therefore lower exposure to this risk factor may confer a lower risk of STEC infection. There is also evidence to suggest that perceptions of attractiveness and safety of local walking

environments may differ by SES, with those in more disadvantaged areas perceiving that their local walking environment is less attractive and less safe compared to those in less disadvantaged areas (The Ramblers' Association, 2010); this could partly explain the lower odds of reporting exposure to walking in a paddock or contact with soil in the more disadvantaged group. A rural-urban difference was also noted between quintiles in our study which may explain the differential exposure. Studies finding a higher risk of STEC infection and HUS development amongst the least disadvantaged individuals cite likely reasons for this finding as related to differential exposures through consumption of higher risk food amongst those of a higher SES (Patil et al., 2005, Simonsen et al., 2008, Tarr et al., 2005, Taylor, 2012, Whitney et al., 2015, Younus et al., 2007); more foreign travel; greater exposure to farm animals or the rural environment; or differential healthcare interaction (Simonsen et al., 2008, Younus et al., 2007). In this study there was evidence for differences in foreign travel, rural environmental exposures and in healthcare contact which would support these assertions. There was no evidence to suggest differences in consumption of known foodborne risk factors or greater exposure to animals. None of the standard risk factors cited to explain differences in risk by SES were more common in the more disadvantaged group. This potentially suggests that difference risk factors, such as person-to-person spread may be more important in disadvantaged groups; a suggestion that has also been postulated by others (Newman et al., 2015). This hypothesis warrants further investigation in future studies.

There is some evidence to suggest that risk factors for STEC are socially patterned, with lower odds of reported exposure to certain known risk factors such as foreign travel and walking in a paddock, but there was no evidence of a relationship between SES and eating out or visiting farms which are often suggested as a possible explanation for differential STEC or GI infection by SES. This could suggest that certain population groups may be more likely to be exposed to risk factors for STEC and which may, in part, explain the distribution of STEC cases by socioeconomic status.

Further research is required to assess the validity of the associations identified in this study to improve our understanding of populations at increased risk of STEC infection and the mechanisms driving inequalities in risk in order to more effectively address these.

7.7 Discussion – socio-demographic risk factors in development of HUS

This is a novel data linkage and analysis of two previously collected datasets to explore the role of socioeconomic and socio-demographic inequalities in the development of HUS following STEC infection. This study captures the progression of HUS in a well-characterised paediatric STEC population. To the best of my knowledge, following a review of the literature and discussion with national experts, this is the first study to explore the relationship between childhood SEC and progression to HUS in England and is one of the largest cohorts of HUS cases to be explored. Infection with STEC is potentially very serious and is likely to result in interaction with healthcare services, particularly in children as it is the most common cause of renal failure in children in the UK (Lynn et al., 2005). The BPSU HUS study was an active surveillance system and it is therefore expected that this cohort captured a high proportion of paediatric HUS in England during the study period. Furthermore, the linkage of the STEC and HUS datasets ensures high ascertainment of HUS cases. The results of this study are likely to be generalisable to other high-income countries.

In this retrospective cohort study examining progression to HUS among paediatric STEC cases, the relationship between childhood SEC and other socio-demographic factors and progression to HUS was investigated using IMD as an area-level measure of SEC. Given the compelling evidence that the risk and consequences of GI infection is higher for more disadvantaged children identified in Chapter 4 (Study 1) and other studies (Olowokure et al., 1999, Phillips et al., 2011, Pockett et al., 2011, Rose et al., 2017) – the finding in our study suggests that SES is unlikely to be a contributor for HUS. Despite the association between SEC and HUS development not reaching significance, the point estimates and much of the confidence intervals fall into a region considered sufficient to represent a clinically important difference and therefore reject the null hypothesis. This finding requires confirmation in larger cohorts.

Sensitivity analysis excluding ethnicity suggested significantly lower odds of progression to HUS among more disadvantaged children. This indicated that ethnicity may be intrinsically linked with IMD, which has been suggested previously (Jivraj and Khan, 2013). Sensitivity analyses using a post-hoc matched analysis on ethnicity and penalised logistic regression suggested that potential collinearity

between these two variables was not enough to alter the results of the main analysis. There were differences identified by ethnicity, with non-White ethnic groups having significantly lower odds of progression to HUS; further research is required to assess this relationship and improved collection of ethnicity in ESQs would be beneficial to this assessment.

This study has several strengths. Multivariable logistic regression explored the relationship between progression to HUS following STEC infection and childhood SEC. Despite a high proportion of missing ethnicity data, we were able to use multiple imputation to impute values where data were missing for this variable and demonstrate that this did not significantly alter the results. We were able to include a number of potentially confounding variables, such as travel to adjust for potential confounding caused by travel-associated strains of STEC tending to be milder compared to indigenous strains. The characteristics of paediatric STEC cases by IMD were broadly comparable to the population distribution. Amongst HUS cases, a much lower proportion of these cases were in the most disadvantaged quintiles. The BPSU HUS study from which the HUS cohort data were derived is an active surveillance system and is therefore likely to have high ascertainment; as such it is also unlikely that differential ascertainment of HUS cases or an unrepresentative population explains our results.

It is possible that there is residual confounding that could not be controlled for and which might explain the lack of a significant relationship between progression to HUS and SEC, such as intrinsic characteristics which may increase differential vulnerability or susceptibility by SEC such as genetic predisposition, co-morbidities, and clinical or treatment characteristics. Further, an area-level measure of SEC was used which uses postcode to assign SECs to study participants. It is possible that the use of an area-level measure is not sensitive enough to detect socioeconomic inequalities, particularly if individual factors rather than area-level factors have more influence over risk of acquiring more severe strains of STEC which would increase risk of progression to HUS. Person-to-person spread is an important risk factor for GI infections and therefore community or area-level risk would be a factor in considering individual risk. Excluding individuals with a serological result from the analysis is a potential bias which could lead to an underestimate of HUS incidence although this is unlikely to affect SES distribution. Finally, it was not always

possible to determine whether antibiotics had been prescribed during treatment for STEC infection or following a diagnosis of HUS therefore this finding should be interpreted with caution.

Previous studies have found that HUS is associated with lower deprivation (Rowe et al., 1991, Whitney et al., 2015). The analysis by Rowe et al. (1991) was similar to this study but included all HUS cases <16 years regardless of their association with STEC and used area-level income as a measure of SEC. The analysis by Whitney et al. (2015) presented age-adjusted HUS incidence but on a small number of HUS cases (n=49). The authors of both studies suggest that the finding of higher risk HUS in children of higher socioeconomic status could relate to differential exposure such as; consumption of higher risk food amongst those of a higher SES; more foreign travel; greater exposure to farm animals or the rural environment; or differential healthcare interaction. Cleary and Lopez (1989), in a study exploring the major questions relating to STEC and HUS, suggest that infection earlier in life in more disadvantaged children might allow for the development of immunity prior to the age of peak risk of severe illness; differential immunity across the life course has been suggested as a potential explanation for differential susceptibility by others (Karmali, 1989).

Studies finding no association between SEC and risk of HUS suggest that this could be related to; differential ascertainment of HUS cases (Bell et al., 1997); potentially unrepresentative populations; or the inability of area-level SEC to detect individual factors contributing to risk (Tarr and Hickman, 1987) in these studies – factors which are not likely to be an issue in this study.

Previous studies in England have suggested that children aged 1-4 years, women and white ethnic groups have the highest incidence of STEC infection (Adams et al., 2016b, Byrne et al., 2015). Higher progression to HUS in children (particularly aged 1-4), women (particularly aged over 10 years) and in those of White ethnicity has also been reported (Bell, 1997, Byrne et al., 2015, Kinney et al., 1988, Launders et al., 2016, Milford et al., 1990, Rogers et al., 1986, Tarr and Hickman, 1987). This study, despite the inevitable changes in strains and exposures during the intervening time period, echoes the findings of by Milford et al (Milford et al., 1990), conducted in Britain, which demonstrated higher progression to HUS amongst children aged 1-

4 years. No difference in risk of HUS by sex was identified in this study, a finding echoed in several other studies (Bell et al., 1997, Cimolai et al., 1994, Rowe et al., 1998, Tserenpuntsag et al., 2005); this is an area of disagreement in the literature with several studies finding higher risk amongst women (Chang et al., 2004, Gould et al., 2009, Rowe et al., 1991) although two of these studies finding higher risk in women did not look specifically among children (Chang et al., 2004, Gould et al., 2009). In this study, differences by age and sex were observed, with a greater proportion of progression to HUS amongst girls less than 1 year of age and 10-15 years of age compared to boys of the same age groups (Figure 7.1), although no interaction between age and sex could be identified. The reasons for the differential risk by age, and the discrepant findings by sex, are currently unclear and there is a need to understand differences in risks and exposures between these groups in order to better understand the mechanisms leading to socioeconomic inequalities and identify links in the causal chain which can be addressed more effectively.

The finding of 15% of STEC cases progressing to HUS is higher than previous studies, which have estimated the proportion of cases of STEC O157 progressing to HUS to be 5.2% in England (Launders et al., 2016) and 9% in Scotland (Locking et al., 2011) however these estimates were for all ages combined. A recent systematic review suggested the proportion of STEC cases progressing to HUS, as identified via cohort studies, was 7.8% (Fischer Walker et al., 2012). This study presents the proportion of paediatric STEC cases of any serotype progressing to HUS which provides a more specific estimate.

Whilst rurality has been documented to be an important factor in risk of STEC in the literature (Byrne et al., 2015, Chang et al., 2009), the findings of this study suggest that rurality is not a significant driver of progression to HUS.

There is some evidence from this study and others to suggest the existence of a relationship between STEC, HUS and SES, with lower SEC associated with non-significant lower risk of developing HUS. There were also demographic differences by age, gender and ethnicity. This could suggest increased risk of exposure or vulnerability to more severe strains of STEC which confers a greater risk of developing HUS.

Further research is required to confirm the relationship between SEC and HUS in larger cohorts. Further research is also required to elucidate the populations at risk of STEC infection and HUS in terms of deprivation, ethnicity, age and sex, in order to better understand whether there are real differences in risk or whether the skew towards less disadvantaged groups is an artefact. These results contribute to the evidence on inequalities in the risk of and vulnerability to GI infections. Alongside future planned analyses, this could ultimately be used to provide further evidence to inform policies to address inequalities in risk, vulnerability and consequences of GI infections.

7.8 Interpretation

In this chapter I have demonstrated evidence to suggest that there is differential healthcare contact and risk factors by SES, with disadvantaged individuals more likely to present at hospital and less likely to present at their GP and lower exposure to environmental risk factors, including travel abroad. I have also demonstrated differential risk of HUS among paediatric STEC cases in England, with lower risk of HUS amongst the most disadvantaged children. Despite compelling evidence to suggest a relationship between GI infections and SES such as higher rates of hospitalisation and increased severity with low SES (Baker et al., 2012, Olowokure et al., 1999, Phillips et al., 2011, Pockett et al., 2011, Rose et al., 2017) and STEC being the most severe indigenous IID in the UK, neither STEC infection nor progression to HUS are contributors to the increased severe IID risk in lower SES. Further, STEC, despite being more common in children, does not have the same SES profile as found for children with GI infections generally.

The finding of differential healthcare contact by SES, with more disadvantaged individuals being more likely to present to hospital and less likely to present at their GP independently of severity, could reflect differential healthcare seeking behaviour. This could go some way to explaining the bias towards higher risk in less disadvantaged individuals if this group is more likely to present to their GP and therefore have a sample taken. It has been hypothesised throughout that children are likely to be presented at healthcare services regardless of deprivation, and therefore that patterns identified in this age group is likely to reflect true patterns regardless of differential healthcare seeking behaviour; disadvantaged children were significantly more likely to report hospitalisation compared to the least disadvantaged children

suggesting that there could be a genuine increased risk of hospitalisation among the most disadvantaged.

The finding of differential reporting of exposure to known risk factors for STEC such as lower reporting of foreign travel and some environmental risk factors could suggest that those who are less disadvantaged are therefore at lower risk of STEC infection. It is of note that oft cited reasons for the social patterning of STEC infections such as lower levels of eating out; animal contact: and visiting farms amongst more disadvantaged individuals were not significant in this analysis. Furthermore, given the social patterning of risk factors identified, the finding of lower risk of HUS in the most disadvantaged could reflect lower likelihood of exposure to known risk factors for STEC and subsequent lower risk of STEC.

Chapter 8
Discussion

8.1 Introduction

The overall aim of this thesis is to investigate the existence, extent and nature of socioeconomic inequalities in risk of GI infections in the UK. The review of the literature (Chapter 2) revealed significant gaps in the knowledge on the risk of GI infections in high income countries, particularly with regard to explanations for the observed findings which included artefact explanations as a result of methodological and measurement of GI infections and SES, as well as hypothesised differential exposure to GI infections and differential healthcare interactions.

The objectives of my thesis were:

1. To conduct a systematic review of existing evidence of socioeconomic inequalities in risk of GI infections in high income countries.
2. To investigate the extent and nature of socioeconomic inequalities in risk of GI infections in the community in the UK, with estimates derived from the most up-to-date population-based household survey.
3. To analyse the extent of, and mechanisms underlying, socioeconomic inequalities in risk of GI infections in the community, with estimates derived from routine data on members of the public seeking telephone-based healthcare advice in England.
4. To explore the social patterning of clinical outcomes, healthcare contact and risk factors for a laboratory-confirmed, potentially severe, GI infection (STEC) and socio-demographic inequalities in risk of development of a serious sequela (HUS) in England.
5. To draw out policy implications and recommendations for further research into the role of socioeconomic inequalities in GI infections.

In this chapter I first summarise the key findings of my work with respect to the objectives of this thesis. I describe the contribution of each empirical study to the existing literature in four main areas: evidence for an association between SES and GI infections; evidence for a differential association between SES and GI infections across the life course; evidence for differential healthcare contact by SES and; evidence for differential exposures for GI infections by SES. I will then describe the strengths and limitations of the studies in this thesis. Finally, I present the conclusions of this thesis, reflect on the policy implications of the findings, offer

some recommendations for further research including work that is on-going and reflect on the PhD experience.

8.2 Key findings

This is the first comprehensive analysis of multiple data sources in order to improve our understanding of the role of socioeconomic inequalities in the risk of GI infections in the UK. The main findings with respect to each study are summarised below, followed by a more detailed discussion of each key finding. Objective 5, to draw out policy implications and recommendations for further research into the role of socioeconomic inequalities in GI infections, will be addressed throughout the subsequent sections of this chapter.

Objective 1: To conduct a systematic review of existing evidence of socioeconomic inequalities in risk of GI infections in high income countries

Study 1 (Chapter 4) addressed the first objective of this thesis through a systematic review and meta-analysis of studies conducted in high income countries. In total, 102 studies were included in the review of which 77 were included in the meta-analysis.

Disadvantaged children, in comparison to their more advantaged counterparts are at higher risk of GI infection.

Clear social patterning of risk of GI infection in children was evident in the harvest plots, with studies reporting higher risk of GI infection in disadvantaged children or no association between GI infection risk and SES. Disadvantaged children had 1.5 times the risk of GI infection compared to advantaged children. It is possible that this reflects differential exposures by SES in children. It is also possible that this reflects differential immunity by SES which has been reported for GI infections by other studies. This will be discussed in more detail later in this chapter.

There was no difference in risk for disadvantaged compared to advantaged adults.

For adults, the pattern illustrated in the harvest plots was not as distinct although most studies demonstrated lower risk of GI infection in disadvantaged adults or no association. There was no significant difference in overall risk of GI infections in disadvantaged compared to advantaged individuals for all ages combined (RR 1.06, 95% CI 0.95–1.19) or disadvantaged compared to advantaged adults (RR 0.83, 95%

CI 0.61–1.14). Study 1 highlighted the importance of exploring GI infections across the life course in subsequent analyses when considering the role of socioeconomic inequalities and the importance of exploring the role of inequalities in cases of GI infections in children specifically.

Objective 2: To investigate the extent and nature of socioeconomic inequalities in risk of GI infections in the community in the UK, with estimates derived from the most up-to-date population-based household survey

Study 2 (Chapter 5) addressed the second objective of this thesis through the analysis of a large prospective community cohort in the UK using self-reported GI symptoms and an individual-level measure of SES using occupation. Response rate for this study was low, and individuals of lower SES were underrepresented which is likely to have influenced the results.

In the UK, the incidence of GI infections is significantly lower in disadvantaged than in advantaged individuals in the community.

Study 2 was conducted in order to assess the role of socioeconomic inequalities in the community not necessarily accessing healthcare. Incidence of GI infections was significantly lower in the most disadvantaged individuals (166 per 1,000 person-years) compared to the least disadvantaged individuals (235 per 1,000 person-years). This finding was consistent when using the area-level measure of SES as well as the individual-level measure. As previously mentioned, the response rate for this study was low and disadvantaged individuals were underrepresented so the risk of GI infections in disadvantaged groups is likely to be underestimated.

There was a significantly lower hazard ratio of GI infections for disadvantaged compared to advantaged individuals.

In Study 2, disadvantaged individuals had significantly lower risk of GI infection compared to advantaged individuals (HR 0.75, 95% CI 0.61-0.91). When stratified by age, this finding was consistent amongst adults but there was no significant difference in risk of GI infection amongst disadvantaged children compared to advantaged children. For the main analysis in Study 2, an individual-level measure of SES was used however when the study was repeated using an area-level measure of SES there was no significant difference in risk for disadvantaged compared to

advantaged individuals. The findings of this study may reflect the low and differential response rate described above although it could also reflect differential exposure or symptom recognition by SES.

Objective 3: To analyse the extent of, and mechanisms underlying, socioeconomic inequalities in risk of GI infections in the community, with estimates derived from routine data on members of the public seeking telephone-based healthcare advice in England

Study 3 (Chapter 6) addressed the third objective of this thesis through an analysis of over 24 million calls to the NHS telephone-based healthcare advice helplines; NHS Direct and NHS 111 between 2010 and 2015.

In England, rate of calls was significantly higher in the most disadvantaged areas compared to the least disadvantaged areas.

In NHS Direct, the rate of calls for both GI and non-GI symptoms was higher in the most disadvantaged areas compared to the least disadvantaged areas (IRR 1.07, 95% CI 1.05-1.08 and IRR 1.20, 95% CI 1.19-1.20 respectively). In NHS 111, the effect of deprivation was more pronounced for GI calls rates, with significantly higher rates in the most disadvantaged areas compared to the least disadvantaged areas (IRR 1.50, 95% CI 1.49-1.52) and was lower for non-GI calls although still significantly higher in the most disadvantaged areas (IRR 1.06, 95% CI 1.05-1.06).

The rate of calls in the most disadvantaged compared to the least disadvantaged areas varied across the life course for GI and non-GI calls.

In NHS Direct, the rate of GI and non-GI calls referring to disadvantaged children aged 0-4 was significantly lower than the rate of calls referring to less disadvantaged children (IRR 0.79, 95% CI 0.78-0.81 and IRR 0.77, 95% CI 0.76-0.77 respectively). For callers aged 15 and over, there was a significantly higher rate of GI calls and non-GI calls to NHS Direct in the most disadvantaged areas.

In NHS 111, the rate of GI calls referring to disadvantaged children and adults was significantly higher than the rate of calls referring to less disadvantaged children. This was not observed for non-GI calls in children aged under 15 for which rates were significantly lower in the most disadvantaged compared to the least

disadvantaged children. For adults less than 60 years of age, rates of non-GI calls were significantly higher in the most disadvantaged areas.

People from more disadvantaged areas had a higher risk of calling for GI symptoms compared to people from less disadvantaged areas and this varied across the life course.

In agreement with Study 1 (Chapter 4), Study 3 (Chapter 6) found that disadvantaged children were at greater risk of GI infections than their less disadvantaged counterparts for NHS 111 calls. Study 3 also presents evidence to suggest that disadvantaged adults are at greater risk of GI infection than their less disadvantaged counterparts. This finding could reflect differential exposure, symptom recognition or healthcare interaction by SES.

Objective 4: To explore the social patterning of clinical outcomes, healthcare contact and risk factors for a laboratory-confirmed, potentially severe, GI infection (STEC) and socio-demographic inequalities in risk of development of a serious sequela (HUS) in England

Study 4 (Chapter 7) addressed the fourth objective of this thesis through the analysis of a national enhanced surveillance system for STEC infection and a national surveillance study of HUS.

In England, disadvantaged individuals had a lower crude incidence of STEC infection compared to advantaged individuals.

In Chapter 7, the social patterning of clinical outcomes (symptoms), healthcare contact and risk factors among STEC cases were explored. Overall crude incidence was significantly lower amongst the most disadvantaged compared to the least disadvantaged (IRR 0.65, 95% CI 0.58-0.73). this pattern was also reflected in the age- and sex-adjusted incidence rates for those over 5 years of age but, the 0-4 age group showed the opposite pattern, although the observed IRRs were not statistically significant for any age-group.

Disadvantaged individuals with STEC infection had higher odds of presentation to A&E and hospitalisation but lower odds of visiting their GP.

Disadvantaged individuals were more likely to report visiting A&E (OR 1.35, 95% CI 1.05-1.74) and being hospitalised (OR 1.71, 95% CI 1.36-2.15) but less likely to visit their GP (OR 0.67, 95% CI 0.53-0.84). These estimates were adjusted for potential severity and suggest that there may be differential interaction with healthcare services by SES. This could go some way to explaining the bias towards higher risk in less disadvantaged individuals if this group is more likely to present to their GP and therefore have a sample taken.

Exposure to known risk factors for STEC was lower in more disadvantaged groups.

Disadvantaged individuals were significantly less likely to report exposure to known risk factors for STEC infection including foreign travel, fresh water, walking in a paddock, contact with soil, taking a day trip and UK travel. This lower exposure to risk factors could lead to lower risk of infection compared to their more advantaged counterparts. There were no significant differences by SES for other commonly cited potential explanations for the differential risk by SES for STEC and other GI infections, such as eating out and consumption of high risk foods (e.g. fresh produce), with the exception of salad, fruit, vegetables or herbs.

There was no significant difference in risk of progression from STEC infection to HUS amongst disadvantaged children compared to advantaged children.

There was no significant difference in risk of progression to HUS in disadvantaged children with STEC compared to advantaged children with STEC. Lower risk of STEC and lack of evidence of a differential risk of HUS among more disadvantaged individuals is an interesting finding, given that evidence from the other studies in this thesis and the wider literature suggest higher risk of GI infection in disadvantaged children and generally higher risk of more severe consequences of GI infection in disadvantaged individuals. Lower risk of exposure to known risk factors for STEC amongst disadvantaged individuals, described above, could explain the lower risk of HUS progression amongst these groups.

Summary

In summary, disadvantaged children are at greater risk of GI infections compared to their less disadvantaged counterparts across high income countries and in England (Studies 1 and 3). In adults in high income countries, there was no difference in risk

by SES (Study 1); however, individuals aged less than 60 years in disadvantaged areas had higher odds of GI infection symptoms than those in less disadvantaged areas based on over 24 million calls to NHS helplines in England (Study 3). Although Study 2 found that disadvantaged individuals were at lower risk of GI infections compared to their less disadvantaged counterparts, the response rate for this study was very low and disadvantaged individuals were underrepresented which is likely to have underestimated the risk of GI infection in this group. There was no evidence for a difference in risk of developing HUS in a paediatric population in England (Study 4).

Differences in types of healthcare utilisation by SES were observed, with evidence that disadvantaged STEC cases were more likely to seek care through A&E and hospitalisation than their GP compared to less disadvantaged cases (Study 4). Study 4 also provided evidence that known risk factors for STEC infection are socially patterned, with disadvantaged individuals less likely to report exposure to known risk factors.

8.3 Contribution to knowledge

As indicated previously, the review of the literature (Chapter 2) revealed significant gaps in the literature and limitations in the design of previously conducted studies. The four empirical studies in this thesis sought to address these methodological limitations and gaps through updating estimates of differential risk by SES in the UK, exploring community level socioeconomic inequalities in risk of GI infections without the requirement for healthcare interaction, exploring differential risk across the life course for children and adults separately and exploring differential risk factors for exposure to GI infections that were hypothesised in the literature but in the absence of evidence.

The studies in this thesis have shown that socioeconomic inequalities in GI infections in the UK do exist; in particular that disadvantaged children are at greater risk of GI infections than their less disadvantaged counterparts. Whilst the community cohort analysis from the IID2 study found that disadvantaged individuals were at lower risk of IID, there was a low response rate, the effect of which was compounded by the fact that the sample of responders was biased in favour of more advantaged individuals and as such the risk of GI infection in disadvantaged groups is likely to

be an underestimate. Amongst users of the NHS telephone-based healthcare advice service, people from more disadvantaged areas were more likely to call for GI symptoms compared to people calling from less disadvantaged areas. These findings provide robust estimates of the association between SES and GI infections at several levels of healthcare interaction and contribute substantially to the wider literature on this topic by providing robust evidence of socioeconomically-driven inequalities in GI infections and evidence to inform further studies exploring this relationship in other countries.

Establishing evidence for an association between SES and GI infections

The studies in this thesis show that SES-driven inequalities in risk of GI infections exist, and as such are potentially amenable to policy interventions. Building on the inconsistent and conflicting findings identified by the review of the literature, Study 1 (Chapter 4) constitutes an original contribution to the literature as it is the first systematic review and meta-analysis on this topic. The review showed that disadvantaged children, but not adults, are at greater risk of GI infections in high income countries. A previous but more limited systematic review explored the relationship between SES and foodborne illness using laboratory confirmed pathogens and found no consistent trends (Newman et al., 2015).

Study 3 showed that people from more disadvantaged areas were more likely to call NHS 111 for GI symptoms compared to people calling from less disadvantaged areas. There have been limited studies exploring differential contact with telephone-based health advice services by SES and none, to the best of my knowledge, which have explored this relationship with a specific focus on GI-related calls, thus the findings of Study 3 constitute an original contribution to the literature. Previous limited studies exploring the relationship between calls in general, not specifically GI calls, to healthcare services and SES reported inconclusive results. Two studies found higher rates of calls for medical advice in more disadvantaged areas (Cooper et al., 2005, Shah and Cook, 2008), although this differed by age and was not seen in calls regarding children (Cooper et al., 2005). In contrast, another study found a non-linear relationship between use of telephone-based healthcare advice and SES, with lower rates in the most and least disadvantaged areas (Burt et al., 2003).

A limitation of many studies exploring the relationship between SES and GI infection is the reliance on laboratory confirmation of a GI pathogen for inclusion in a study. There are therefore few population-based cohort studies that have been conducted in high income countries to investigate differences in risk of GI infection by SES. As with Study 1, studies finding higher risk in disadvantaged children could reflect differential severity. Indeed, analysis of the IID2 dataset to explore socioeconomic inequalities in consequences through symptom severity and sickness absence, found that more disadvantaged individuals were more likely to experience greater symptom severity and, largely as a result of this increased severity, a higher level of sickness absence (Rose et al., 2017).

Study 2, a population-level cohort (IID2 Study) using an individual-level measure of SES, found that the most disadvantaged individuals were less likely to report GI symptoms compared to the least disadvantaged. As previously described, the response rate to this study was very low (9%) and varied by SES, with disadvantaged individuals being underrepresented and as such the risk of GI infection in disadvantaged individuals is likely to be underestimated. Study 4 suggested lower risk of STEC amongst the most disadvantaged and no significant difference in risk of developing HUS. Whilst this study was reliant on laboratory reporting, it also suggested evidence for differential healthcare seeking behaviour and exposure to risk factors by SES which could help to explain the differential risk of GI infection by SES identified in this thesis.

Establishing evidence for an association between SES and GI infections across the life course

The studies in this thesis provide consistent evidence of higher risk of GI infection in disadvantaged children. It is hypothesised that children may be more likely to present to healthcare services, regardless of SES, compared to adults, therefore it is possible that any ascertainment bias due to healthcare seeking behaviours is minimised in children and as such the findings from Study 1 could reflect a genuine difference in risk by SES. It is also possible that this finding reflects differential consequences of GI infection rather than differential risk. Arguably, this could relate back to differential contact with healthcare services, for example, if more disadvantaged individuals access care differently, or differential symptom recognition, for example,

if more disadvantaged individuals recognise or act upon symptoms at a more advanced stage of illness.

It has also been hypothesised previously that differential immunity may influence the observed differential risk. For example, in a study looking at STEC infection and HUS, Cleary and Lopez (1989) hypothesise that GI infection earlier in life amongst more disadvantaged children may allow these children to develop immunity prior to the age at which they would become most susceptible to illness and its consequences and thus reduce the burden of disease in disadvantaged adults. Drawing on studies conducted low income countries on *Campylobacter* (Fernández et al., 2008, Kakai et al., 1995, Lloyd-Evans et al., 1983, Quetz et al., 2010), as well as other bacteria and parasites (Nematian et al., 2004, O'Ryan et al., 2005), suggests that this could be a plausible explanation as such infections are seen almost exclusively in disadvantaged children (Fernández et al., 2008, Kakai et al., 1995, Lloyd-Evans et al., 1983, Quetz et al., 2010) but rarely in adults (Coker et al., 2002). In Study 1 the relationship between SES and GI infection was less clear for adults than for children, while in Study 3, there was evidence to suggest that the association with SES and GI infection was less clear for adults aged 60 and over. It is possible that this could reflect waning immunity and further research to explore whether differential immunity by SES exists and how this impacts on vulnerability to infection across all age groups is required.

Establishing evidence for an association between SES and healthcare interaction for GI infections

Study 4 presented evidence of differential reporting of healthcare contact by SES; more disadvantaged individuals were less likely to report contact with their GP and more likely to report contact with A&E or hospitalisation. This finding supports the hypothesis that differential healthcare contact may be mediating the relationship between SES and risk of GI infection. It could also reflect differential recognition of symptoms if the reason for the need for healthcare services providing a more urgent response, such as A&E compared to a GP, relates to symptom severity, although potential severity was adjusted for in this study. It could also be that access to, and quality of, GP services is poorer in more disadvantaged areas and residents may therefore bypass primary care and go straight to A&E.

As discussed earlier, studies in this thesis have suggested that this may differ by type of healthcare contact, with evidence to suggest higher levels of referral to urgent healthcare services and use of NHS telephone-based healthcare advice (Study 3) and higher levels of presentation at A&E and hospitalisation compared to GP (Study 4). A study by Carlisle et al. (2002) found a higher proportion of out-of-hours, surgery consultations and same-day consultations in a disadvantaged GP practice compared to a less disadvantaged practice; routine visits by doctors and practice nurses did not have an association with deprivation. Higher referral rates in disadvantaged areas have also been identified (Hippisley-Cox et al., 1997). Deprivation is associated with shorter length of consultation (Stirling et al., 2001). Despite not investigating this specifically for GI infections, it raises the possibility of a potential differential interaction with healthcare services amongst disadvantaged compared to less disadvantaged individuals and a potential 'inverse care law' effect. The 'inverse care law' first described by Tudor Hart (1971) suggests that availability of medical care varies inversely with need for care. It is possible that whilst the most disadvantaged might be more likely to seek care, they may actually have greater barriers to accessing care and thus be underrepresented in national records of healthcare contact, including laboratory reporting.

Limited analysis of differential healthcare contact has been undertaken for GI infections however where this has been explored, studies have found higher levels of healthcare seeking for diarrhoeal illness amongst more disadvantaged individuals. One study conducted in two areas of Tyneside, UK found that children in the less affluent area were more likely to present to primary care with an episode of diarrhoea (Edwards, 1996). The author of this study suggests that this could relate to greater difficulty in self-managing symptoms. Similarly a study conducted in the US found that among people with acute diarrhoea, lower household income and lower level of education was associated with higher odds of seeking medical care (Scallan et al., 2006). Tam et al. (2003) found higher levels of GP attendance amongst more disadvantaged individuals with GI infections using the IID2 study data, with authors speculating that this could relate to an individual's level of education influencing their health beliefs and awareness or ability to self-care. de Wit et al. (2001b) found that those with the lowest and highest education levels had the lowest levels of consulting their GP.

Establishing evidence for an association between SES and exposures for GI infections

In this section I will discuss the contribution of the studies in this thesis to our understanding of the potential mechanisms that may be driving these inequalities. Whilst I have been able to demonstrate evidence to suggest differential exposure using the case study of STEC infection, it is important to note that Study 4 explores exposures for one specific pathogen; exposures are likely to differ by pathogen and as such there are likely to be important exposures that are not captured in this dataset.

The studies in this thesis contribute to the literature on GI epidemiology through improving our understanding of socioeconomic inequalities in exposure to GI infections. These studies have provided evidence that inequalities in exposure to GI infections exist, and as such are potentially amenable to policy interventions; by reducing exposure to GI infections there is an opportunity to modify risk of developing illness.

Study 4 is a novel study exploring differential reporting of key risk factors for STEC infection. The results of Study 4 provide more robust evidence to answer the question of why inequalities in risk exist. Through Study 4 specific risk factors were identified which could provide evidence, rather than the hypotheses presented in other studies, of differential exposure to GI infections. Testing these hypotheses, Study 4 found that more disadvantaged groups were less likely to report foreign travel. Whilst it is possible that this reflects a social gradient in foreign travel at a population level, this is a known risk factor for STEC infection (Byrne et al., 2015, Parry et al., 1998) and therefore may reflect a genuine differential risk which contributes to the lower risk of STEC infection in disadvantaged individuals. Contrary to speculation in the literature cited above, Study 4 did not identify any differences by SES in reporting of food-related exposures, with the exception of fish/shellfish and salad/fruit/vegetables/herbs consumption (which were lower in more disadvantaged groups). More disadvantaged individuals were also less likely to report walking in a paddock or contact with soil. It is possible that these findings reflect social patterning of exposures in the general population although, if these results reflect genuine differential exposures by SES, these findings could suggest that inequalities in exposure to environmental risk factors, including travel, play a more important role in driving inequalities in risk of STEC infection. As there was

no evidence to suggest differential exposure to other risk factors postulated as contributing to the differential risk, such as food related risk exposures; this finding provides greater clarity and evidence for the explanation of lower risk in more disadvantaged population groups. It could be that individuals with STEC do have different exposure to risk factors compared to the healthy population. Further research is required to test this hypothesis.

Application of the Diderichsen model to GI infections

An adaptation of the Diderichsen model (Diderichsen et al., 2001) was used as a theoretical and analytical framework to guide the analyses within this thesis. In this model differential exposure, differential vulnerability and differential consequences are key mechanisms through which SES may influence GI infections. Following on from the findings of the empirical studies within this thesis, an adapted Diderichsen model is presented (Figure 8.1) to demonstrate these mechanisms for GI infections.

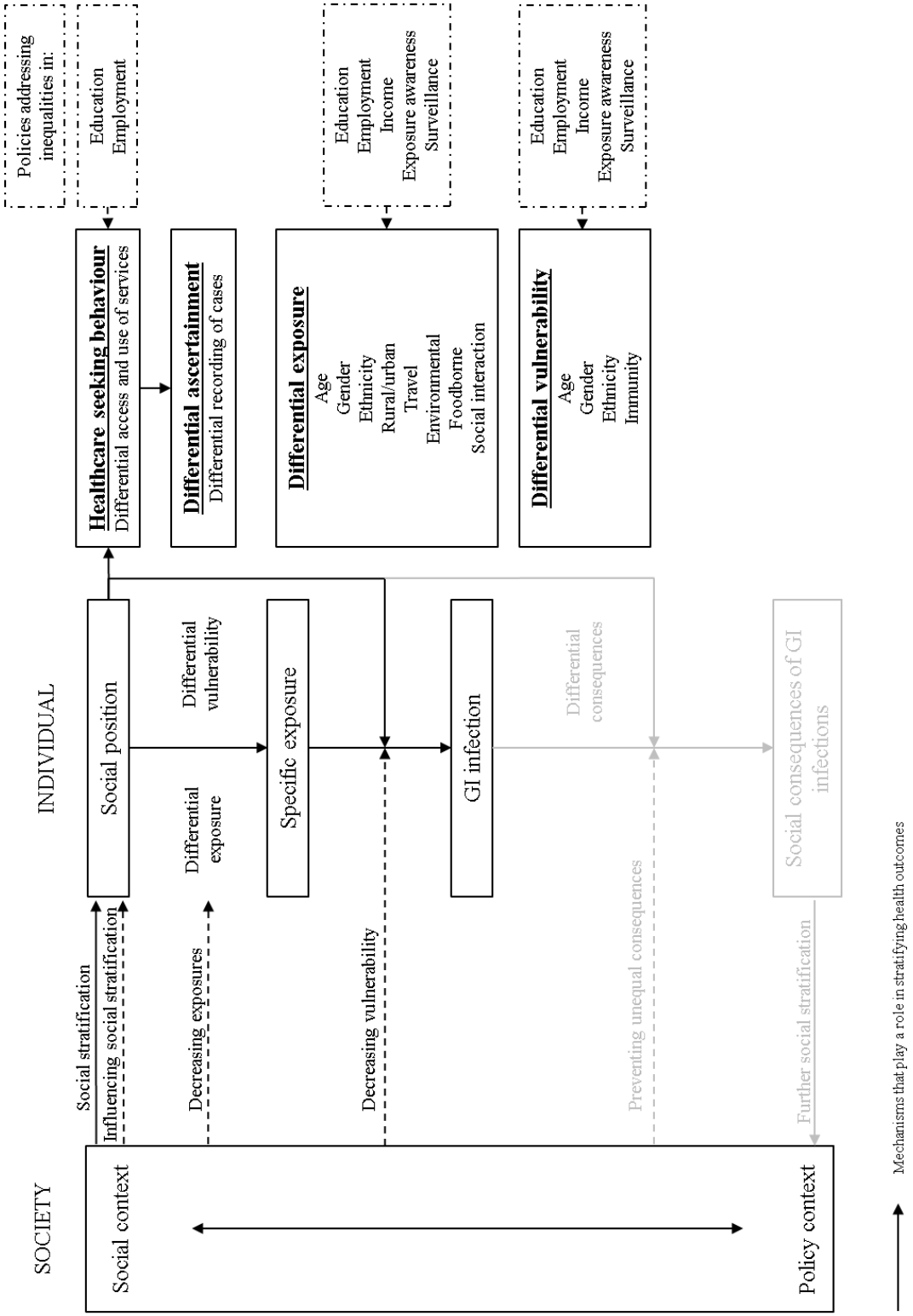
Differential exposure was explored in Study 4 through analysis STEC cases. Study 4 suggested differential exposure to known risk factors for STEC by SES existed and could play a part in the differential risk observed within STEC infection and across the other studies in this thesis. Studies 1 and 3 suggested that disadvantaged children were at greater risk of GI infections – Study 1 also suggested that there was no differential risk for adults whilst Study 3 suggested disadvantaged groups of all ages are at greater risk of GI infections. Conversely Studies 2 and 4 suggested that disadvantaged individuals and groups were at lower risk of GI infections although biases in the study population in Study 2 may underestimate risk in disadvantaged individuals. Although it was not possible to explore differential vulnerability empirically in the studies in this thesis, the observed discrepancy between the studies in terms of differential risk could relate to differential vulnerability to infection through differential exposure or differential immunity. The relationship may also be explained by artefact explanations such as differential ascertainment at different levels of healthcare contact caused by differential healthcare interaction or differential access to services.

Finally, work to explore the differential consequences of GI infections has also been undertaken by my PhD colleague, Tanith Rose. The findings from Rose's studies indicate that those who are more disadvantaged experience more severe illness,

greater sickness absence as a result (Rose et al., 2017), and higher risk of hospitalisation (Rose, 2017). Policies could be implemented to support appropriate access to healthcare services which would go some way to resolving any effect of an inverse care law in access. This could also help to reduce inequalities in consequences of GI infections by care being given at an earlier stage of illness and may help to reduce ascertainment bias and therefore present a clearer view of the association between GI infection and SES. Educational interventions to highlight potential risk factors for GI infections and ways to mitigate the risk of exposure may help to reduce differential risk of infection.

Figure 8.1: Adaptation of the Diderichsen model outlining differential risk, vulnerability and consequences of GI infections and policy entry points

Source: Adapted from Diderichsen et al. (2001) by Adams, this thesis



8.4 Critique of overall study design

The strengths and limitations of each study were discussed in their respective chapters (Chapters 4-7). The strengths and limitations of the overall thesis will be explored briefly below. This thesis brings together a series of novel analyses of existing datasets including novel data-linkages and application of statistical methods to explore the role of SES in influencing risk of and exposure to GI infections.

Key strengths of the datasets

One of the main strengths of this thesis is the measurement of inequalities in GI infections across a range of datasets. Triangulating the results across the four studies and multiple datasets enhances the validity of the findings and validates the consistency of the results, allowing several hypotheses to be tested. Furthermore, examining the issue from several perspectives - individuals not seeking healthcare; individuals seeking telephone-based healthcare advice; and individuals with a diagnosed GI infection, provides estimates of the association between socioeconomic inequalities and GI infections at different points in the healthcare pathway. The ability to include syndromic definitions of GI infection, which has not been widely considered previously potentially due to lack of robust datasets, is a particular strength of this work as it has allowed for consideration of populations who would not be identified via routine surveillance and would therefore not ordinarily be included in statistics. The inclusion of these individuals is vital if the hypothesis that there is differential healthcare interaction by SES is correct, as these associations would be challenging to detect if the study were restricted solely to laboratory confirmed infections.

The studies in this thesis have also made wider contributions to our understanding of health inequalities through the exploration and consideration of healthcare interaction and risk factors for infection. This multi-level approach which triangulates the findings across several datasets and analyses could be applied to the exploration of inequalities in other infectious diseases. Across all studies, it was possible to take a number of key confounders into consideration and statistically account for missing data to increase the robustness of the findings and ensure high quality results. The identification of such confounders, as well as important considerations such as

stratifying by age are valuable for future work exploring the association between SES and GI infections in high income countries.

External validity is also high. These results provide more robust estimates and evidence for the associations described in previously published literature. The datasets included within this thesis are large, comprehensive and generally representative of both the wider population of the UK and other high income countries.

Key limitations of the datasets

The datasets were subject to a number of limitations including potential for misclassification of GI infections, lack of population-level data for risk-factor comparisons and lack of individual-level SES measures, response bias and unavailability of data on comorbidities, GP and hospitalisation data and data to explore differential vulnerability.

A major limitation in Study 2 was that the participation in the IID2 cohort study was only 9% of the original number recruited and screened, and differed by SES as those who were most disadvantaged were underrepresented in comparison to the UK population. It is therefore possible that the estimates derived from this study are not reliable for the assessment of the role of SES in GI infections; this is supported by the weight of evidence from Study 3 which was an extremely large and representative dataset. It is also possible that some of the other datasets were not representative of the population. For the two telephone-based healthcare advice datasets (NHS Direct and NHS 111), the characteristics of individuals who do not make use of these services is not known. Postcode districts where no calls were made were included to minimise this issue.

The results from the first three studies in this thesis were not pathogen specific. Including studies for which laboratory confirmation was not necessary may have led to misclassification but this possibility is small. In addition, although the wider literature is suggestive of differential risk between pathogens, one of the main hypotheses explored in this thesis was that of differential healthcare contact (including no contact) by SES therefore it was considered important to include data

which did not require a diagnosis of a specific GI infection to assess whether there was any evidence for differential healthcare interaction.

For several of the datasets included in this thesis there was a lack of population comparison. In the STEC dataset, it was not possible to compare the reporting of risk factors within the STEC population with the reporting of these in the general population and therefore it is possible that the differential reporting of risk factors identified may reflect differential reporting at the population-level.

It was not possible to obtain details of comorbidities, an important covariate but one that is not routinely or robustly collected, which could have acted as a confounder or as an effect modifier of the relationship between SES and GI infections. There were also issues with missing data across several of the datasets. In the IID2 Study, the individual measure of SES, NS-SEC, was not classifiable for a relatively high proportion of participants (n=1,112, 16.3%). In the STEC and HUS datasets, ethnicity was missing for 19.1% and 20.1% respectively. As discussed in each of the relevant chapters and above, the use of multiple imputation to statistically impute values for each missing record allows participants with missing data to be included in the analysis which is a more robust approach than exclusion through complete-case analysis.

Across all studies with the exception of IID2, there were no individual-level measures of SES. It is possible that area-level measures of SES do not adequately capture individual-level inequalities in risk of GI infections and as such the use of the area-level measures may mask some of these inequalities, although it is important to note that analyses at group- or area-level for diseases spread via person-to-person transmission may provide a more realistic estimate of risk compared with individual-based interpretations which do not account for this – it is therefore possible that the ecological fallacy does not hold as strong for GI infections. Further research in this area should consider robust methods of capturing individual-level measures of SES despite the challenges this poses for routine data collection as often this is not the primary focus and a balance between data collection and non-response needs to be sought.

Despite being one of the largest studies conducted on inequalities in HUS, no significant differential risk was identified. The results were suggestive of a lower risk

of HUS amongst more disadvantaged children with STEC infection although it could be that this dataset was not large enough to detect a significant difference. Similarly, in the IID2 dataset, there was no significant difference in risk for children and it is possible that this subset of the main IID2 dataset was not large enough to detect a significant difference.

Unfortunately, it was not possible to assess data to explore differential risk of GI infections at the GP or hospital level. This was mitigated somewhat by exploring healthcare contact by SES within the STEC cohort, using this diagnosed GI infection as a case-study. Furthermore, other work associated with this thesis explored inequalities in GI infection at the hospital level and found higher risk of hospitalisation in the most disadvantaged areas (Rose, 2017). Finally, it was also not possible to directly explore differential vulnerability to GI infections through these datasets.

Key strengths of the analyses

The analyses in this thesis were theoretically informed through the analytical framework of the Diderichsen model (Diderichsen et al., 2001) of the mechanisms of differential exposure and vulnerability in the pathway to health inequalities. Furthermore, the limitations of previously published studies have been taken into account when designing the analyses included in this thesis. The multidisciplinary expertise of the collaborators on the analyses within this thesis is also a key strength. Collaborators included experts in GI infections, epidemiology, public health, health inequalities and economics. Public involvement and engagement is a central part of the National Institute for Health Research Health Protection Unit in Gastrointestinal infections (NIHR HPRU in GI infections) and as such, lay members of the external advisory panel have also been involved in shaping the work in this thesis through feedback and participation at meetings. Although it was not possible to directly engage patients in the analyses due to use of secondary datasets, I participated in a workshop for schools to raise awareness of GI infections (National Institute for Health Research Health Protection Unit in Gastrointestinal infections, 2017).

A variety of observational study designs were utilised. In Study 1, a systematic review and meta-analysis was used. This approach allows for a robust exploration of existing literature including statistically pooling results to provide estimates of risk of

GI infections by SES using a much larger sample size than would otherwise be possible as well as harvest plots which allowed for a visual assessment of the strength and weight of the evidence. It provided a timely update of the situation in high income countries and the first meta-analysis on this topic. The remaining analyses were conducted on large national datasets. In Study 1, a prospective cohort study (IID2) was used. Prospective cohort studies are normally regarded as expensive and time consuming; however this was an efficient use of an existing dataset which provided robust measures of the incidence of GI infections by SES. As SES was collected at the beginning of the follow-up period, recall bias was minimised. Finally, Studies 3 and 4 were cross-sectional studies which were able to explore the relationship between SES and GI infection in defined populations with high quality exposure data, providing robust estimates of the associations between SES and GI infection.

Key limitations of the analyses

It was also not possible to classify a relatively high proportion of individuals by SES in the IID2 Study. Multiple imputation was used in order to account for this however this does not address the low response rate for this study which differed by SES.

In several of the studies, with the exception of the IID2 Study, area-level SES was necessary to use due to the absence of individual-level measures. Despite being a commonly used measure of SES, it is important to acknowledge that using an area-level measure could lead to ecological fallacy, particularly if inequalities at the individual-level are driving differential risk. This was largely countered by assessing the relationship across number of datasets and levels of healthcare contact and, as described above, area-level measures may more accurately estimate the relationship for diseases such as GI infections which have person-to-person transmission routes. The use of different SES measures across the studies is also a potential issue. This is particularly the case in the systematic review (Study 1) although across the other studies IMD was used to assess the relationship which improves the comparability between studies. This required individual postcodes to link to IMD scores from which population-level quintiles of deprivation were calculated. In the NHS Direct and NHS 111 datasets, only postcode district was available which is a larger

geographical area than postcode. As such, this presents a larger risk of ecological fallacy.

A lack of multiple measures of SES in Studies 3 and 4 meant that it was not possible to corroborate the findings within each study. Study 1 highlighted that multiple SES measures were used and in Study 2, the use of both an area-level and individual-level measure produced differing results; both found significant differences in incidence rates in the most disadvantaged compared to the least disadvantaged although there was no significant difference identified in hazard of GI infection in the most disadvantaged compared to least disadvantaged when using IMD. Further research should consider including multiple measures of SES. The studies in this thesis consistently use IMD, which is a commonly used measure of SES; however triangulating findings using multiple measures of SES may help to understand the role played by the measure of SES used.

The IMD scores were converted to population-level IMD quintiles. This allowed for an assessment of the characteristics of the study population by SES in the context of the general population although grouping IMD scores in this way may have reduced the study power.

The studies within this thesis were observational. Observational studies have a number of limitations. In the systematic review and meta-analysis (Study 1), there was a high level of heterogeneity between the studies and a large amount of unexplained heterogeneity, despite stratified analyses being conducted to explore this. This suggests that residual confounding caused by unmeasured variables are a potential issue and this is a possible limitation across all the studies within this thesis.

With prospective cohort studies (Study 2), differential loss to follow up is regarded as a potential area for bias however differential follow-up time was accounted for in the analysis of this dataset. Participation bias was a major issue in Study 2, participation in this cohort study was low (9%) and higher SES groups were over-represented, while lower SES were under-represented, which may have biased the results and provided an under-estimation of risk in more disadvantaged groups. Finally, the use of cross-sectional analyses (Studies 3 and 4) meant that it was not possible to determine causation; the risk factor associations identified (Study 4)

could not be proven to be the cause of infection and it was not possible to assess the risk factor associations identified alongside a population-level comparator.

Methodological developments

The studies in this thesis have contributed a number of methodological developments for the investigation of socioeconomic inequalities in risk of and exposure to GI infections. I have applied systematic review and meta-analysis methods; making use of a novel data synthesis method, harvest plots. I have performed novel data linkage of large and routinely collected datasets; linked with population-level inequalities data to answer novel questions. A variety of analytical approaches have been used and where possible comparable analyses have been performed across the datasets. The studies in this thesis represent an in depth exploration of inequalities in GI infection. The approach used in this thesis could be applied to explore the role of inequalities in other infectious diseases.

Consistency and conflicts between studies

The studies in this thesis have provided consistent evidence that disadvantaged children are at greater risk of GI infection in comparison to their more advantaged counterparts; however the findings for adults were much less consistent.

A large and methodologically robust systematic review and meta-analysis (Study 1) found quantitative evidence to support disadvantaged children being at greater risk of GI infections compared to more advantaged children, however there was no significant difference in risk by SES for adults. A systematic review is a robust and evidence-based method of collating empirical evidence based on pre-specified criteria to address a specific research question (Higgins and Green, 2011) and as such, the results of Study 1 provide strong evidence of the association between SES and GI infection in children.

By contrast, a large prospective cohort study was used to assess this relationship (Study 2) in the community and found significantly lower risk in disadvantaged adults but no significant difference in risk for children. The point estimate for lower risk for disadvantaged adults was consistent with the point estimate for the reduced risk in disadvantaged adults in the systematic review (Study 1), although this was non-significant. Prospective cohort studies are widely regarded as the 'gold-standard'

of epidemiological study designs and, as such, this is a robust dataset and the findings within this study may reveal real associations. There is a particular problem with the IID2 cohort study, however, in relation to the SES studies. There was substantial selection bias by SES within this study, with over-representation of professional and managerial classes and under-representation of lower manual and routine classes in recruitment and retention (Appendix 3.6). There is further potential for selection bias in relation to health outcomes, which could bias the results of the SES associations. This study adds to the methodological evidence base by identifying and highlighting the importance of considering selection bias by SES when analysing even gold standard cohort studies.

Similarly, despite strong evidence of a significantly higher risk of GI infection in disadvantaged children in the systematic review, a retrospective paediatric cohort study (Study 4) was conducted and also found no significant difference in risk of developing HUS among paediatric STEC cases by SES. Retrospective cohort studies, despite having greater potential for bias than prospective cohort studies, are also considered to be robust study designs and the findings in this study are also likely to be real associations.

Furthermore, a cross-sectional study of the social patterning of risk factors among STEC cases was undertaken and found disadvantaged individuals overall had a significantly lower risk of STEC compared to advantaged individuals, in agreement with Study 2. This pattern was also reflected in the age- and sex-adjusted incidence rates for those over 5 years of age however, the 0-4 age group showed the opposite pattern of higher risk in disadvantaged children although the observed IRRs were not statistically significant for any age-group.

In agreement with the systematic review (Study 1), a large cross-sectional study exploring NHS telephone helpline data (Study 3) found significantly higher risk of calls for GI symptoms in children, as well as adults, from disadvantaged areas compared to more advantaged areas in NHS 111, although there was a significantly lower risk for children in NHS Direct and no difference for adults. NHS 111 is a freephone number and acts as an out-of-hours GP service, unlike NHS Direct, and resulted in much higher call volume which could explain this discrepancy. Furthermore, the lack of symptom data for infants under 1 in NHS Direct from

November 2011 may have contributed to these findings. This was an extremely large study comprising data from over 24 million calls which increases the likelihood that the findings are true and not due to chance. These datasets are nationally representative in terms of coverage although it is possible that they may not be representative in terms of the use of telephone helplines by SES. Cross-sectional studies are considered to be less robust than the study designs described above but are useful to explore associations between an exposure and an outcome. While the finding of higher risk for disadvantaged children is consistent with the findings in Study 1 and the non-significant finding for higher risk of STEC infection in young children (0-4 years) in Study 4, this study was the only study in this thesis to find a higher risk of GI infection in adults.

8.5 Implications of findings for policy and practice

As has been clearly established throughout this thesis, GI infections place a considerable burden on the economy, society and individuals. There are an estimated 17 million cases in the UK each year (Tam et al., 2011a), and this burden is not distributed equally across society with evidence that disadvantaged individuals, and disadvantaged children in particular, experience a greater share of the burden. Various agencies and individuals have specific responsibilities to identify and seek to address socioeconomic inequalities in health including clinicians, primary and secondary care, PHE, local authorities, policy makers, education systems and social care. In the following sections I will make some reflections on the policy entry points using the framework of the Diderichsen model (Figure 8.1).

This research forms part of the NIHR HPRU in GI infections, which aims to:

> *"explore and explain the distribution of diarrhoeal diseases in the population, establishing for whom the disease burden is greatest and why...This integrated, inter-disciplinary research programme will generate new strategies for control, meeting Public Health England's main objectives of addressing inequalities, protecting the country from infectious diseases, and being an evidence-led organisation that provides answers to public health problems."*

(National Institute for Health Research Health Protection Unit in Gastrointestinal infections, 2014)

The studies in this thesis have provided evidence to suggest that inequalities exist for GI infections; further evidence of inequalities in health more broadly in the UK; and has provided evidence of the magnitude of the problem. The evidence is particularly strong for GI infections in children; inequalities are most evident in the most vulnerable of populations; children living in disadvantaged circumstances.

As David Taylor-Robinson stated as the keynote speaker at the 2017 PHE Conference:

> *"[it is] absolutely baffling that we allow an exposure as toxic as child poverty - a modifiable exposure - to wash over such a large proportion of the children in this country. The central challenge to practitioners is to advocate for policies and practices that are going to improve the developmental trajectories of children especially those living in disadvantaged conditions as a way to improve their health and reduce inequalities over their life course."*

(Taylor-Robinson, 2017)

While risk factors for GI infections are generally well-understood and have resulted in multiple education and awareness campaigns on good food hygiene, for example, how to avoid cross-contamination (Food Standards Agency, 2014), little attention has been paid to the socioeconomic context in which personal behaviour takes place and wider barriers to infection control. This may be one reason why such general public education campaigns have not been particularly effective. Improvements in and ongoing routine data collection efforts to inform policy could be implemented to help inform action to tackle socioeconomic inequalities in GI infections. Such improvements in data collection through enhanced surveillance systems will be crucial in developing evidence for policy.

Decreasing exposure

As part of the PHE remit to protect and improve the nation's health and wellbeing, reducing exposure to infectious diseases such as GI infections is important. Without on-going research to provide evidence of inequalities in GI infections, particularly

with regards to these areas of differential exposure, immunity and healthcare interaction, it is not possible to develop effective policy responses. It could be beneficial to collect these data systematically and at an individual-level on representative population-level datasets as a priority for surveillance and research.

More broadly, surveillance of GI infections needs to consider the potential for disadvantaged individuals to be underrepresented, particularly for milder GI infections where necessity for contact with healthcare services is reduced. Consideration of other data sources which do not require laboratory confirmation of pathogens would be beneficial in order to develop the evidence base for potential inequalities in healthcare access and utilisation. Improving our understanding of the level of underreporting by SES and different GI pathogens as well as by the different healthcare access points will also enable the burden of GI infections to be more accurately estimated.

The Gastrointestinal Infections Department at PHE, responsible for national surveillance of GI pathogens, is considering how best to explore inequalities in GI infections and capture inequalities in routine data collection such as enhanced surveillance systems. In particular, I am in discussion with the surveillance leads to inform and advise on the improvements needed in data collection. Through these discussions, consideration is being given to changes in the collection of data on SES through routine enhanced surveillance questionnaires and questionnaires used in outbreaks in order to explore some of the findings of this thesis in detail for STEC and for other GI infections such as listeriosis. Furthermore, the methods used in this thesis can be used to study the role of socioeconomic inequalities in other infectious diseases. As was highlighted in the analysis of the IID2 data in Study 2 (Chapter 5), disadvantaged individuals are often much less likely to engage in studies compared to less disadvantaged individuals. Active recruitment of disadvantaged individuals, who would otherwise be lost of follow-up, would help to improve the availability of data through which to explore differential exposure. Surveillance of GI infections could also take into consideration the potential for disadvantaged individuals to access care differently, or to not access care at all.

Preventing unequal consequences

Whilst the focus of this thesis was on differential risk of and exposure to GI infections, differential consequences of GI infections were also identified. Policies aimed at reducing risk and exposures are likely to also reduce the risk of consequences of GI infection. Furthermore if, for example, social stratification results in seeking care at a later stage of illness leading to more severe disease (Rose et al., 2017), policies addressing awareness of healthcare access and improving employment situations could be crucial in redressing the inequalities in the consequences of GI. Policies which seek to reduce the impact of sickness absence on children's education and on family circumstances whilst ensuring disadvantaged individuals are not further disadvantaged could also be crucial in redressing the imbalance in the consequences of GI infections.

8.6 Conclusions

The studies in this thesis suggest that disadvantaged individuals and disadvantaged children in particular, are at greater risk of GI infections. Although Study 2 (Chapter 5) found a lower risk in disadvantaged individuals, due to the low and biased response rate it is possible that this result is misleading; an assertion supported by the results of Study 3 (Chapter 6) which was an extremely large and representative study.

There is evidence to suggest differential healthcare interaction or access by SES. This finding could reflect the need for more flexible healthcare services for disadvantaged groups if it is not possible for them to schedule appointments with their GP in accordance with other commitments such as lack of sick-leave or childcare. It is also possible that there are differential symptom recognition or self-care abilities by SES leading to illness becoming more severe and requiring more urgent treatment.

Differential exposure to GI infection by SES is likely to play a role in driving the differential risk of infection observed by SES. Although disadvantaged children have a higher risk of GI infections in general, there is no evidence to suggest they are at greater risk of developing HUS following infection with STEC, which could reflect the lower risk of exposure to known risk factors for STEC infection among disadvantaged individuals identified in Study 4.

The studies in this thesis make an original contribution to the literature and to the study of socioeconomic inequalities in GI infections. The findings suggest the existence of inequalities which should be borne in mind by policy makers, clinicians and epidemiologists working in GI surveillance to reduce the burden of GI infections. Practically, steps should be taken to explore interventions and policies aimed at reducing the risk of, exposure to and consequences of GI infections, as well as considering the individual risk factors which may be contributing to increased risk. Given the high burden of GI infections both to society and the NHS, it could be beneficial to explore the efficacy of such policies in order to reduce this burden.

8.7 Recommendations for further research

GI infections are common and certain sociodemographic characteristics may affect the risk of or exposure to GI infections. It is of concern that disadvantaged children are at greater risk of infection in high income countries. This was not apparent with progression to HUS following STEC infection but further research is required in larger datasets to confirm these findings.

Further work is currently in progress from this thesis. A complementary thesis exploring inequalities in the consequences of GI infections in the UK has been undertaken which found that disadvantaged individuals are at greater risk of the consequences of GI infections including sickness absence and symptom severity (Rose et al., 2017) and hospitalisation (Rose, 2017).

The analysis of socioeconomic inequalities is challenging due to their multifactorial nature and as such, quantitative studies can be limited in their ability to provide some of the answers for the results found. A number of hypotheses using the quantitative analyses within this thesis have been suggested for testing using mixed methods and qualitative approaches as part of the further work of the HPRU:

- Disadvantaged individuals may be treated differently by healthcare professionals; for example they may be less likely to request or have a stool sample take, or disadvantaged children may be more readily admitted to hospital because clinicians judge that their living conditions are not suitable for treatment at home.

- Disadvantaged individuals may interact differently with healthcare services; for example they may have more barriers to accessing healthcare through their GP.

- Disadvantaged individuals may have differential symptom recognition or self-care capabilities; for example they may delay seeking care until symptoms are more severe or be unable to rehydrate effectively.

- Disadvantaged individuals may be exposed differently; for example they may undertake different activities.

- Disadvantaged individuals may be more vulnerable to more severe consequences of GI infections; for example they may have comorbidities.

Robust studies are also required to assess differential exposure to risk factors in the context of exposure to these factors in the wider public; the use of a control group in future studies would be beneficial. Assessment of differential exposure for other GI infections is also warranted as risk factors are likely to vary by pathogen. Expanding the methods in this thesis to explore inequalities in pathogen-specific risk could provide answers to some of the questions raised by the literature review and also the conflicting evidence presented for adults in the systematic review and meta-analysis.

It would be beneficial to explore the hypothesis that disadvantaged children are exposed earlier, experience GI infection at an earlier age and at a greater rate and thus develop immunity prior to becoming most susceptible to the potential for severe disease. Understanding how childhood exposure impacts on risk in adulthood would also potentially explain the differential risk across the life course. As it was not possible to explore differential vulnerability in this thesis, research into this potential mechanism is crucial to complete the picture of the pathways to inequalities for GI infections. Future studies could make use of the methods proposed by Nordahl et al. (2014) to test for both differential exposure and vulnerability to GI infections as these have been demonstrated to operate at the same time and are not mutually exclusive (Nordahl et al., 2014). Through this approach it is possible to quantify the role of differential vulnerability and differential exposure separately.

The differential healthcare contact observed among STEC cases could suggest differences in healthcare interaction amongst disadvantaged STEC cases in terms of interacting with or access to services or potentially recognition of symptoms at a more advanced stage of illness. Further studies are now needed to robustly assess these hypotheses, both for STEC and for other GI pathogens including non-laboratory confirmed illnesses. Assessing inequalities in GI infection at the GP level would be beneficial to completing the healthcare contact patterns observed through the studies in this thesis. Including distance to GP in future analyses may improve our understanding of inequalities in healthcare access. Additionally, assessment of the relationship between SES and GI infections following the rotavirus vaccine introduction in datasets such as HES or in GP datasets would be informative.

Finally, exploration of the role of SES in GI infection risk, exposure and consequences could also be expanded to assess inequalities amongst older age groups, which would be complemented by a greater understanding of the role of comorbidities in risk and consequences of GI infections. Further work in this area will help to identify those at greatest risk of exposure to and consequences of GI infections which will help to inform policies and interventions to reduce the risk, vulnerability and social, economic and healthcare consequences of GI infections in the UK.

8.8 Reflections on the PhD experience

Undertaking this research has been an invaluable learning experience. I have gained skills and experience in conducting academically rigorous research and in the use of statistical methods with broad applicability including systematic review, meta-analysis, missing data methods and survival analysis. I have also had the opportunity to develop analytical skills in both Stata and R statistical software packages.

If I were to undertake this PhD again, I would have liked to conduct primary data collection of a community cohort ensuring that participation of disadvantaged individuals was central to the data collection methodology. Unfortunately it was not possible to obtain the planned linked dataset, due to unforeseen complications in access to data, which would have contained data from PHE laboratory reports, HES and GP data. Whilst this meant it was necessary to rethink the structure of this

research, it also made it possible to analyse other distinct datasets which have provided up to date assessments of the situation along the healthcare pathway.

This PhD studentship has been unconventional in the sense that I have been fortunate to have six supervisors and have had joint supervision with another PhD student. Whilst this had the potential to be a chaotic approach, the reality is that this has formed a strong, cross-disciplinary collaboration which will be continuing on from these PhD studies.

Following on from this research, I plan to establish a health inequalities interest group at PHE Colindale for infectious diseases specifically to promote cross-learning and sharing of best practice for research and policy. I also plan to explore the role of socioeconomic inequalities in risk and consequences of *Listeria monocytogenes* for which PHE have an enhanced surveillance system and I hope to evaluate the impact of the rotavirus vaccine using syndromic surveillance systems, including assessing whether the vaccine resulted in any differential impact by SES. I will be advocating for the systematic collection of individual-level socioeconomic data through enhanced surveillance systems for GI infections in order to provide further information with which to assess the role of socioeconomic inequalities for specific GI infections.

References

ACHESON, D., BARKER, D., CHAMBERS, J., GRAHAM, H., MARMOT, M. & WHITEHEAD, M. 1998. Independent Inquiry into Inequalities in Health Report. London: The Stationery Office.

ACKERS, M. L., MAHON, B. E., LEAHY, E., GOODE, B., DAMROW, T., HAYES, P. S., BIBB, W. F., RICE, D. H., BARRETT, T. J., HUTWAGNER, L., GRIFFIN, P. M. & SLUTSKER, L. 1998. An outbreak of Escherichia coli O157:H7 infections associated with leaf lettuce consumption. *Journal of Infectious Diseases,* 177**,** 1588-93.

ADAK, B., LONG, S. M. & O'BRIEN, S. J. 2002. Trends in indigenous foodborne disease and deaths, England and Wales: 1992 to 2000. *Gut,* 51**,** 832-841.

ADAMS, N., ROSE, T., TAYLOR-ROBINSON, D., BARR, B., O'BRIEN S, J., VIOLATO, M., HAWKER, J. & WHITEHEAD, M. Does socioeconomic status influence risk of gastrointestinal infections in the community in the UK? European Public Health Conference, 2016a Vienna. European Journal of Public Health.

ADAMS, N., ROSE, T. C., TAYLOR-ROBINSON, D., BARR, B., HAWKER, J., O'BRIEN S, J., VIOLATO, M. & WHITEHEAD, M. 2015. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review. *PROSPERO International prospective register of systematic reviews*.

ADAMS, N. L., BYRNE, L., SMITH, G. A., ELSON, R., HARRIS, J. P., SALMON, R., SMITH, R., O'BRIEN, S. J., ADAK, G. K. & JENKINS, C. 2016b. Shiga Toxin-Producing Escherichia coli O157, England and Wales, 1983-2012. *Emerging Infectious Diseases,* 22**,** 590-7.

ADAMS, N. L., ROSE, T. C., HAWKER, J., VIOLATO, M., O'BRIEN S, J., WHITEHEAD, M., BARR, B. & TAYLOR-ROBINSON, D. 2017. Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 study). *European Journal of Public Health*.

ADLAM, S. B., PERERA, S., LAKE, R. J., CAMPBELL, D. M., WILLIMAN, J. A. & BAKER, M. G. 2011. Acute gastrointestinal illness in New Zealand: a community study. *Epidemiology and Infection,* 139**,** 302-8.

AL-JADER, L., SALMON, R. L., WALKER, A. M., WILLIAMS, H. M., WILLSHAW, G. A. & CHEASTY, T. 1999. Outbreak of Escherichia coli

O157 in a nursery: lessons for prevention. *Archives of Disease in Childhood,* 81**,** 60-3.

ARMON, K., STEPHENSON, T., GABRIEL, V., MACFAUL, R., ECCLESTON, P. & WERNEKE, U. 2001. Determining the common medical presenting problems to an accident and emergency department. *Archives of Disease in Childhood,* 84.

ASHTON, J. 1988. The History of Public Health in Liverpool: upwards and onwards, pendulum or helix? *Esmedune 2000: Vision or Dream (A Healthy Liverpool).* Department of Public Health, University of Liverpool.

ATCHISON, C. J., STOWE, J., ANDREWS, N., COLLINS, S., ALLEN, D. J., NAWAZ, S., BROWN, D., RAMSAY, M. E. & LADHANI, S. N. 2016. Rapid Declines in Age Group-Specific Rotavirus Infection and Acute Gastroenteritis Among Vaccinated and Unvaccinated Individuals Within 1 Year of Rotavirus Vaccine Introduction in England and Wales. *Journal of Infectious Diseases,* 213**,** 243-9.

BAKER, D., TAYLOR, H. & HENDERSON, J. 1998. Inequality in infant morbidity: Causes and consequences in England in the 1990s. *Journal of Epidemiology and Community Health,* 52**,** 451-458; 451.

BAKER, M. G., BARNARD, L. T., KVALSVIG, A., VERRALL, A., ZHANG, J., KEALL, M., WILSON, N., WALL, T. & HOWDEN-CHAPMAN, P. 2012. Increasing incidence of serious infectious diseases and inequalities in New Zealand: a national epidemiological study. *Lancet,* 379**,** 1112-1119.

BANATVALA, N., CRAMP, A., JONES, I. R. & FELDMAN, R. A. 1999. Salmonellosis in North Thames (East), UK: associated risk factors. *Epidemiology and Infection***,** 201-207.

BARNETT, K., MERCER, S. W., NORBURY, M., WATT, G., WYKE, S. & GUTHRIE, B. 2012. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet,* 380**,** 37-43.

BARROS, H. & LUNET, N. 2003. Association between child-care and acute diarrhea: A study in Portuguese children. *Revista de Saude Publica,* 37**,** 603-608; 603.

BAWA, Z., ELLIOT, A. J., MORBEY, R. A., LADHANI, S., CUNLIFFE, N. A., O'BRIEN S, J., REGAN, M. & SMITH, G. E. 2015. Assessing the Likely

Impact of a Rotavirus Vaccination Program in England: The Contribution of Syndromic Surveillance. *Clinical Infectious Diseases,* 61**,** 77-85.

BEALE, N., PEART, C., KAY, H., TAYLOR, G., BOYD, A. & HERRICK, D. 2010. 'ALSPAC' infant morbidity and council tax band: Doctor consultations are higher in lower bands. *European Journal of Public Health,* 20**,** 403-408.

BELL, B. P., GRIFFIN, P. M., LOZANO, P., CHRISTIE, D. L., KOBAYASHI, J. M. & TARR, P. I. 1997. Predictors of hemolytic uremic syndrome in children during a large outbreak of Escherichia coli O157:H7 infections. *Pediatrics,* 100**,** E12.

BELL, W. R. 1997. Thrombotic thrombocytopenic purpura/hemolytic uremic syndrome relapse: frequency, pathogenesis, and meaning. *Seminars in Hematology,* 34**,** 134-9.

BEMIS, K., MARCUS, R. & HADLER, J. L. 2014. Socioeconomic status and campylobacteriosis, Connecticut, USA, 1999-2009. *Emerging Infectious Diseases,* 20**,** 1240-2.

BENTANCOR, A. B., AMEAL, L. A., CALVINO, M. F., MARTINEZ, M. C., MICCIO, L. & DEGREGORIO, O. J. 2012. Risk factors for Shiga toxin-producing Escherichia coli infections in preadolescent schoolchildren in Buenos Aires, Argentina. *Journal of Infection in Developing Countries,* 6**,** 378-86.

BERKEY, C. S., HOAGLIN, D. C., MOSTELLER, F. & COLDITZ, G. A. 1995. A random-effects regression model for meta-analysis. *Statistics in Medicine,* 14**,** 395-411.

BESSELL, P. R., MATTHEWS, L., SMITH-PALMER, A., ROTARIU, O., STRACHAN, N. J., FORBES, K. J., COWDEN, J. M., REID, S. W. & INNOCENT, G. T. 2010. Geographic determinants of reported human Campylobacter infections in Scotland. *BMC Public Health,* 10**,** 423.

BIERING-SØRENSEN, S., SONDERGAARD, G., VITTING ANDERSEN, K., ANDERSEN, A. M. & MORTENSEN, L. H. 2012. Time trends in socio-economic factors and risk of hospitalisation with infectious diseases in pre-school children 1985-2004: a Danish register-based study. *Paediatric and Perinatal Epidemiology,* 26**,** 226-35.

BLACK, D., MORRIS, J. N., SMITH, C. & TOWNSEND, P. 1980. Inequalities in Health: Report of a Research Working Group (The Black Report). London: DHSS.

BORGNOLO, G., BARBONE, F., SCORNAVACCA, G., FRANCO, D., VINCI, A. & IUCULANO, F. 1996. A case-control study of Salmonella gastrointestinal infection in Italian children. *Acta Paediatrica, International Journal of Paediatrics,* 85**,** 804-8.

BRITISH PAEDIATRIC SURVEILLANCE UNIT. 2014. *BPSU - Haemolytic Uraemic Syndrome* [Online]. Available: http://www.rcpch.ac.uk/bpsu-haemolytic-uraemic-syndrome [Accessed 13/06/2017].

BRITISH PAEDIATRIC SURVEILLANCE UNIT. 2016. *BPSU - How it works* [Online]. Available: http://www.rcpch.ac.uk/what-we-do/bpsu/what-bpsu/how-bpsu-system-works/how-bpsu-system-works [Accessed 13/06/2017].

BRITISH PAEDIATRIC SURVEILLANCE UNIT. 2017. *BPSU - Studies* [Online]. Available: http://www.rcpch.ac.uk/bpsu/currentstudies [Accessed 13/06/2017].

BURT, J., HOOPER, R. & JESSOPP, L. 2003. The relationship between use of NHS Direct and deprivation in southeast London: an ecological analysis. *Journal of Public Health Medicine,* 25**,** 174-6.

BYRNE, L., ADAMS, N., GLEN, K., DALLMAN, T. J., KAR-PURKAYASTHA, I., BEASLEY, G., WILLIS, C., PADFIELD, S., ADAK, G. & JENKINS, C. 2016. Epidemiological and Microbiological Investigation of an Outbreak of Severe Disease from Shiga Toxin-Producing Escherichia coli O157 Infection Associated with Consumption of a Slaw Garnish. *Journal of Food Protection,* 79**,** 1161-8.

BYRNE, L., FISHER, I., PETERS, T., MATHER, A., THOMSON, N., ROSNER, B., BERNARD, H., MCKEOWN, P., CORMICAN, M., COWDEN, J., AIYEDUN, V., LANE, C. & INTERNATIONAL OUTBREAK CONTROL, T. 2014a. A multi-country outbreak of Salmonella Newport gastroenteritis in Europe associated with watermelon from Brazil, confirmed by whole genome sequencing: October 2011 to January 2012. *Euro Surveillance,* 19**,** 6-13.

BYRNE, L., JENKINS, C., LAUNDERS, N., ELSON, R. & ADAK, G. K. 2015. The epidemiology, microbiology and clinical impact of Shiga toxin-

producing Escherichia coli in England, 2009-2012. *Epidemiology and Infection*, 1-13.

BYRNE, L., ON BEHALF OF THE BPSU HUS STUDY TEAM, 2017. BPSU HUS Study Report.

BYRNE, L., VANSTONE, G. L., PERRY, N. T., LAUNDERS, N., ADAK, G. K., GODBOLE, G., GRANT, K. A., SMITH, R. & JENKINS, C. 2014b. Epidemiology and microbiology of Shiga toxin-producing Escherichia coli other than serogroup O157 in England, 2009-2013. *Journal of Medical Microbiology,* 63**,** 1181-8.

BYTZER, P., HOWELL, S., LEEMON, M., YOUNG, L., JONES, M. & TALLEY, N. 2001. Low socioeconomic class is a risk factor for upper and lower gastrointestinal symptoms: a population based study in 15 000 Australian adults. *Gut,* 49**,** 66-72.

CARLISLE, R., AVERY, A. J. & MARSH, P. 2002. Primary care teams work harder in deprived areas. *Journal of Public Health Medicine,* 24**,** 43-8.

CHANG, H.-G. H., TSERENPUNTSAG, B., KACICA, M., SMITH, P. F. & MORSE, D. L. 2004. Hemolytic uremic syndrome incidence in New York. *Emerging Infectious Diseases,* 10**,** 928-31.

CHANG, M., GROSECLOSE, S. L., ZAIDI, A. A. & BRADEN, C. R. 2009. An ecological analysis of sociodemographic factors associated with the incidence of salmonellosis, shigellosis, and E. coli O157:H7 infections in US counties. *Epidemiology and Infection,* 137**,** 810-20.

CHESHIRE, J. 2012. Lives on the Line: Mapping Life Expectancy Along the London Tube Network. *Environment and Planning,* 44.

CIMOLAI, N., BASALYGA, S., MAH, D. G., MORRISON, B. J. & CARTER, J. E. 1994. A continuing assessment of risk factors for the development of Escherichia coli O157:H7-associated hemolytic uremic syndrome. *Clinical Nephrology,* 42**,** 85-9.

CLEARY, T. G. & LOPEZ, E. L. 1989. The Shiga-like toxin-producing Escherichia coli and hemolytic-uremic syndrome. *Pediatric Infectious Disease Journal,* 8**,** 720-724.

COHEN, S. A., EGOROV, A. I., JAGAI, J. S., MATYAS, B. T., DEMARIA, A., JR., CHUI, K. K., GRIFFITHS, J. K. & NAUMOVA, E. N. 2008. The SEEDs of two gastrointestinal diseases: socioeconomic, environmental, and

demographic factors related to cryptosporidiosis and giardiasis in Massachusetts. *Environmental Research,* 108**,** 185-91.

COKER, A. O., ISOKPEHI, R. D., THOMAS, B. N., AMISU, K. O. & OBI, C. L. 2002. Human campylobacteriosis in developing countries. *Emerging Infectious Diseases,* 8**,** 237-44.

CONWAY, S. P., PHILLIPS, R. R. & PANDAY, S. 1990. Admission to hospital with gastroenteritis. *Archives of Disease in Childhood,* 65**,** 579-84.

COOK, E. J., RANDHAWA, G., LARGE, S., GUPPY, A., CHATER, A. M. & ALI, N. 2014. Barriers and facilitators to using NHS Direct: a qualitative study of 'users' and 'non-users'. *BMC Health Services Research,* 14**,** 487.

COOKSON, R., PROPPER, C., ASARIA, M. & RAINE, R. 2016. Socioeconomic inequalities in health care in England. *In:* CENTRE FOR HEALTH ECONOMICS (ed.). York: University of York,.

COOPER, D., ARNOLD, E., SMITH, G., HOLLYOAK, V., CHINEMANA, F., BAKER, M. & O'BRIEN, S. 2005. The effect of deprivation, age and sex on NHS Direct call rates. *British Journal of General Practice,* 55**,** 287-91.

COOPER, D., SMITH, G., O'BRIEN, S., HOLLYOAK, V. & BAKER, M. 2003. What can Analysis of Calls to NHS Direct Tell us about the Epidemiology of Gastrointestinal Infections in the Community? *Journal of Infection,* 46**,** 101-105.

DAHLGREN, G. & WHITEHEAD, M. Tackling inequalities in health: what can we learn from what has been tried? Working paper prepared for the King's Fund International Seminar on Tackling Inequalities in Health, 1993 Ditchley Park, Oxfordshire. London: King's Fund.

DAHLGREN, G. & WHITEHEAD, M. 2007a. Concepts and principles for tackling social inequalities in health: Levelling up Part 1. *Studies on social and economic determinants of population health.* WHO Collaborating Centre for Policy Research on Social Determinants of Health, University of Liverpool.

DAHLGREN, G. & WHITEHEAD, M. 2007b. European strategies for tackling social inequities in health: Levelling up Part 2. Copenhagen: WHO Regional office for Europe.

DALLMAN, T. J., ASHTON, P. M., BYRNE, L., PERRY, N. T., PETROVSKA, L., ELLIS, R., ALLISON, L., HANSON, M., HOLMES, A., GUNN, G. J., CHASE-TOPPING, M. E., WOOLHOUSE, M. E. J., GRANT, K. A.,

GALLY, D. L., WAIN, J. & JENKINS, C. 2015. Applying phylogenomics to understand the emergence of Shiga-toxin-producing Escherichia coli O157:H7 strains causing severe human disease in the UK. *Microbial Genomics*, 1.

DAVID, S. T., MACDOUGALL, L., LOUIE, K., MCINTYRE, L., PACCAGNELLA, A. M., SCHLEICHER, S. & HAMADE, A. 2004. Petting zoo-associated Escherichia coli 0157:h7--secondary transmission, asymptomatic infection, and prolonged shedding in the classroom. *Canadian Communicable Disease Report*, 30**,** 173-80.

DAVIES, H. T. O., CROMBIE, I. K. & TAVAKOLI, M. 1998. When can odds ratios mislead? *British Medical Journal*, 316**,** 989-991.

DAVIS, M. A., MOORE, D. L., BAKER, K. N., FRENCH, N. P., PATNODE, M., HENSLEY, J., MACDONALD, K. & BESSER, T. E. 2013. Risk factors for campylobacteriosis in two washington state counties with high numbers of dairy farms. *Journal of Clinical Microbiology*, 51**,** 3921-7.

DE WIT, M. A., KOOPMANS, M. P., KORTBEEK, L. M., WANNET, W. J., VINJE, J., VAN LEUSDEN, F., BARTELDS, A. I. & VAN DUYNHOVEN, Y. T. 2001a. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. *American Journal of Epidemiology*, 154**,** 666-74.

DE WIT, M. A., KORTBEEK, L. M., KOOPMANS, M. P., DE JAGER, C. J., WANNET, W. J., BARTELDS, A. I. & VAN DUYNHOVEN, Y. T. 2001b. A comparison of gastroenteritis in a general practice-based study and a community-based study. *Epidemiology and Infection*, 127**,** 389-97.

DE WIT, M. A. S., KOOPMANS M. P. G. & VAN DUYNHOVEN Y. T. H. P. 2003. Risk Factors for Norovirus, Sapporo-like Virus, and Group A Rotavirus Gastroenteritis. *Emerging Infectious Diseases*, 9**,** 1563-1570; 1563.

DEPARTMENT FOR COMMUNITIES AND LOCAL GOVERNMENT. 2011. *English Indices of Deprivation 2010* [Online]. Department for Communities and Local Government. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf [Accessed 16/09/2016].

DEPARTMENT FOR ENVIRONMENT, FOOD AND RURAL AFFAIRS,. 2016. *Rural Urban Classification* [Online]. Available:

https://www.gov.uk/government/collections/rural-urban-classification [Accessed].

DIDERICHSEN, F., EVANS, T. & WHITEHEAD, M. 2001. The Social Basis of Disparities in Health. *In:* TIMOTHY EVANS, M. W., FINN DIDERICHSEN, ABBAS BHUIYA, AND MEG WIRTH (ed.) *Challenging Inequities in Health: From Ethics to Action.* New York: Oxford University Press.

DOORDUYN, Y., VAN DEN BRANDHOF, W. E., VAN DUYNHOVEN, Y. T., BREUKINK, B. J., WAGENAAR, J. A. & VAN PELT, W. 2010. Risk factors for indigenous Campylobacter jejuni and Campylobacter coli infections in The Netherlands: a case-control study. *Epidemiology and Infection,* 138**,** 1391-404.

DOORDUYN, Y., VAN PELT, W. & HAVELAAR, A. H. 2012. The burden of infectious intestinal disease (IID) in the community: a survey of self-reported IID in The Netherlands. *Epidemiology and Infection,* 140**,** 1185-92.

DUNLOP, S., COYTE, P. C. & MCISAAC, W. 2000. Socio-economic status and the utilisation of physicians' services: results from the Canadian National Population Health Survey. *Social Science and Medicine,* 51**,** 123-133.

DUNN, P. M. 2002. Dr William Farr of Shropshire (1807-1883): obstetric mortality and training. *Archives of Disease in Childhood Fetal Neonatal Edition,* 87**,** F67-9.

DUNTEMAN, G. H. & HO, M.-H. R. 2006. An introduction to generalized linear models. *Quantitative Applications in the Social Sciences.* Thousand Oaks, CA: SAGE Publications Ltd.

DWORKIN, M. S., SHOEMAKER, P. C., GOLDOFT, M. J. & KOBAYASHI, J. M. 2001. Reactive arthritis and Reiter's syndrome following an outbreak of gastroenteritis caused by Salmonella enteritidis. *Clinical Infectious Diseases,* 33**,** 1010-4.

EATON-EVANS, J. & DUGDALE, A. E. 1987. Effects of feeding and social factors on diarrhoea and vomiting in infants. *Archives of Disease in Childhood,* 62**,** 445-448; 445.

EDWARDS, A. P. R. 1996. Patterns of help-seeking behaviour for toddlers from two contrasting socio-economic groups: New evidence on a neglected topic. *Family Practice,* 13**,** 377-381; 377.

ETHELBERG, S., OLSEN, K. E., SCHEUTZ, F., JENSEN, C., SCHIELLERUP, P., ENBERG, J., PETERSEN, A. M., OLESEN, B., GERNER-SMIDT, P. & MOLBAK, K. 2004. Virulence factors for hemolytic uremic syndrome, Denmark. *Emerging Infectious Diseases,* 10**,** 842-7.

EVANS , M. R., ROBERTS, R. J., RIBEIRO, C. D., GARDNER, D. & KEMBREY, D. 1996. A milk-borne campylobacter outbreak following an educational farm visit. *Epidemiology and Infection,* 117**,** 457-462.

EVANS, M. R., SARVOTHAM T., THOMAS D. R. & HOWARD A. J. 2006. Domestic and travel-related foodborne gastrointestinal illness in a population health survey. *Epidemiology and Infection,* 134**,** 686-693; 686.

FAHEY, T., MORGAN, D., GUNNEBURG, C., ADAK, G. K., MAJID, F. & KACZMARSKI, E. 1995. An outbreak of Campylobacter jejuni enteritis associated with failed milk pasteurisation. *Journal of Infection,* 31**,** 137-43.

FEIN, S. B., LIN, C. T. J. & LEVY, A. S. 1995. Foodborne illness: Perceptions, experience, and preventive behaviors in the United States. *Journal of Food Protection,* 58**,** 1405-1411.

FERNÁNDEZ, H., VERA, F., VILLANUEVA, M. P. & GARCIA, A. 2008. Occurrence of campylobacter species in healthy well-nourished and malnourished children. *Brazilian Journal of Microbiology,* 39**,** 56-58.

FISCHER WALKER, C. L., APPLEGATE, J. A. & BLACK, R. E. 2012. Haemolytic-Uraemic Syndrome as a Sequela of Diarrhoeal Disease. *Journal of Health, Population and Nutrition,* 30**,** 257-261.

FOOD STANDARDS AGENCY 2000. A report of infectious intestinal disease in England. The Stationary Office.

FOOD STANDARDS AGENCY. 2014. *Don't wash raw chicken* [Online]. Available: https://www.food.gov.uk/news-updates/campaigns/campylobacter/fsw-2014 [Accessed].

FOX, J. 2016. Generalized Linear Models. *In:* FOX, J. (ed.) *Applied Regression Analysis and Generalized Linear Models.* 3 ed. USA: SAGE Publications.

FREEMAN, R., DABRERA, G., LANE, C., ADAMS, N., BROWNING, L., FOWLER, T., GORTON, R., PETERS, T., MATHER, H., ASHTON, P., DALLMAN, T., GODBOLE, G., TUBIN-DELIC, D., CHARLETT, A., FISHER, I. & ADAK, G. K. 2016. Association between use of proton pump inhibitors and non-typhoidal salmonellosis identified following investigation

into an outbreak of Salmonella Mikawasima in the UK, 2013. *Epidemiology and Infection,* 144**,** 968-75.

FRIEDMAN, C. R., HOEKSTRA, R. M., SAMUEL, M., MARCUS, R., BENDER, J., SHIFERAW, B., REDDY, S., AHUJA, S. D., HELFRICK, D. L., HARDNETT, F., CARTER, M., ANDERSON, B., TAUXE, R. V. & EMERGING INFECTIONS PROGRAM FOODNET WORKING, G. 2004. Risk factors for sporadic Campylobacter infection in the United States: A case-control study in FoodNet sites. *Clinical Infectious Diseases,* 38 Suppl 3**,** S285-96.

FULLERTON, K. E., INGRAM, L. A., JONES, T. F., ANDERSON, B. J., MCCARTHY, P. V., HURD, S., SHIFERAW, B., VUGIA, D., HAUBERT, N., HAYES, T., WEDEL, S., SCALLAN, E., HENAO, O. & ANGULO, F. J. 2007. Sporadic campylobacter infection in infants: a population-based surveillance case-control study. *Pediatric Infectious Disease Journal,* 26**,** 19-24.

GIBNEY, K. B., CHENG, A. C., HALL, R. & LEDER, K. 2017. Sociodemographic and geographical inequalities in notifiable infectious diseases in Australia: a retrospective analysis of 21 years of national disease surveillance data. *Lancet Infectious Diseases,* 17**,** 86-97.

GILLESPIE, I., MOOK, P., LITTLE, C., GRANT, K. & MCLAUCHLIN, J. 2010a. Human listeriosis in England, 2001-2007: association with neighbourhood deprivation. *Euro Surveillance,* 15.

GILLESPIE, I. A., MOOK, P., LITTLE, C. L., GRANT, K. & ADAK, G. K. 2010b. Listeria monocytogenes infection in the over-60s in England between 2005 and 2008: a retrospective case-control study utilizing market research panel data. *Foodborne Pathogens and Disease,* 7**,** 1373-9.

GILLESPIE, I. A., O'BRIEN, S. J., ADAK, G. K., CHEASTY, T. & WILLSHAW, G. 2005. Foodborne general outbreaks of Shiga toxin-producing Escherichia coli O157 in England and Wales 1992-2002: where are the risks? *Epidemiology and Infection,* 133**,** 803-8.

GILLESPIE, I. A., O'BRIEN, S. J., ADAK, G. K., TAM, C. C., FROST, J. A., BOLTON, F. J., TOMPKINS, D. S. & CAMPYLOBACTER SENTINEL SURVEILLANCE SCHEME, C. 2003. Point source outbreaks of

Campylobacter jejuni infection--are they more common than we think and what might cause them? *Epidemiology and Infection,* 130**,** 367-75.

GILLESPIE, I. A., O'BRIEN, S. J., PENMAN, C., TOMPKINS, D., COWDEN, J. & HUMPHREY, T. J. 2008. Demographic determinants for Campylobacter infection in England and Wales: implications for future epidemiological studies. *Epidemiology and Infection,* 136**,** 1717-25.

GOH, S., NEWMAN, C., KNOWLES, M., BOLTON, F. J., HOLLYOAK, V., RICHARDS, S., DALEY, P., COUNTER, D., SMITH, H. R. & KEPPIE, N. 2002. E. coli O157 phage type 21/28 outbreak in North Cumbria associated with pasteurized milk. *Epidemiology and Infection,* 129**,** 451-7.

GOLDBLATT, P. & WHITEHEAD, M. 2000. Inequalities in health - development and change. *In:* STATISTICS, O. F. N. (ed.) *Population Trends 100.*

GOULD, L. H., DEMMA, L., JONES, T. F., HURD, S., VUGIA, D. J., SMITH, K., SHIFERAW, B., SEGLER, S., PALMER, A., ZANSKY, S. & GRIFFIN, P. M. 2009. Hemolytic uremic syndrome and death in persons with Escherichia coli O157:H7 infection, foodborne diseases active surveillance network sites, 2000-2006. *Clinical Infectious Diseases,* 49**,** 1480-5.

GOVERNMENT STATISTICAL SERVICE 2013. The 2011 Rural-Urban Classification For Small Area Geographies: A User Guide and Frequently Asked Questions (v1.0).

GRAFEN, A. & HAILS, R. 2002. *Modern statistics for the life sciences,* Oxford, Oxford University Press.

GRAHAM, H. 2009. *Understanding health inequalities,* Maidenhead, Open University Press, McGraw-Hill Education.

GREEN, C. G., KRAUSE, D. O. & WYLIE, J. L. 2006. Spatial analysis of campylobacter infection in the Canadian province of Manitoba. *International Journal of Health Geographics,* 5**,** 2.

GRIFFIN, G., ON BEHALF OF THE E. COLI O157 INDEPENDENT INVESTIGATION COMMITTEE, 2010. Review of the major outbreak of E. coli O157 in Surrey, 2009

HALL, G. V., KIRK, M. D., ASHBOLT, R., STAFFORD, R., LALOR, K. & AND THE OZFOODNET WORKING GROUP 2006. Frequency of infectious gastrointestinal illness in Australia, 2002: regional, seasonal and demographic variation. *Epidemiology and Infection,* 134**,** 111-8.

HALLIDAY, S. 2003. Duncan of Liverpool: Britain's first Medical Officer. *Journal of Medical Biography,* 11**,** 142-9.

HARKER, K. S., LANE, C., GORMLEY, F. J. & ADAK, G. K. 2014. National outbreaks of Salmonella infection in the UK, 2000-2011. *Epidemiology and Infection,* 142**,** 601-7.

HEMPEL, S., MILES, J. N., BOOTH, M. J., WANG, Z., MORTON, S. C. & SHEKELLE, P. G. 2013. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews,* 2**,** 107.

HERIKSTAD, H., YANG, S., VAN GILDER, T. J., VUGIA, D., HADLER, J., BLAKE, P., DENEEN, V., SHIFERAW, B. & ANGULO, F. J. 2002. A population-based estimate of the burden of diarrhoeal illness in the United States: FoodNet, 1996–7. *Epidemiology and Infection,* 129.

HIGGINS, J. & GREEN, S. E. 2011. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0.* [Online]. Available: www.handbook.cochrane.org [Accessed 18/10/2016].

HIGGINS, J. & THOMPSON, S. G. 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine,* 23**,** 1663-1682.

HIPPISLEY-COX, J., HARDY, C., PRINGLE, M., FIELDING, K., CARLISLE, R. & CHILVERS, C. 1997. The effect of deprivation on variations in general practitioners' referral rates: a cross sectional study of computerised data on new medical and surgical outpatient referrals in Nottinghamshire. *British Medical Journal,* 314**,** 1458-61.

HUGHES, G. J. & GORTON, R. 2015. Inequalities in the incidence of infectious disease in the North East of England: a population-based study. *Epidemiology and Infection,* 143**,** 189-201.

IACONO, G., MEROLLA, R., D'AMICO, D., BONCI, E., CAVATAIO, F., DI PRIMA, L., SCALICI, C., INDINNIMEO, L., AVERNA, M. R., CARROCCIO, A. & PAEDIATRIC STUDY GROUP ON GASTROINTESTINAL SYMPTOMS IN, I. 2005. Gastrointestinal symptoms in infancy: a population-based prospective study. *Digestive and Liver Disease,* 37**,** 432-8.

IHEKWEAZU, C., CARROLL, K., ADAK, B., SMITH, G., PRITCHARD, G. C., GILLESPIE, I. A., VERLANDER, N. Q., HARVEY-VINCE, L., REACHER, M., EDEGHERE, O., SULTAN, B., COOPER, R., MORGAN,

G., KINROSS, P. T., BOXALL, N. S., IVERSEN, A. & BICKLER, G. 2012. Large outbreak of verocytotoxin-producing Escherichia coli O157 infection in visitors to a petting farm in South East England, 2009. *Epidemiology and Infection,* 140**,** 1400-13.

INNS, T., BEASLEY, G., LANE, C., HOPPS, V., PETERS, T., PATHAK, K., PEREZ-MORENO, R., ADAK, G., SHANKAR, A. & OUTBREAK CONTROL, T. 2013. Outbreak of Salmonella enterica Goldcoast infection associated with whelk consumption, England, June to October 2013. *Euro Surveillance,* 18.

JACKSON, R., SMITH, D., TABNAK, F. & VUGIA, D. 2015. Disparities of shigellosis rates among California children by race/ethnicity and census tract poverty level, 2000-2010. *Pediatric Infectious Disease Journal,* 34**,** 843-847.

JALAVA, K., OLLGREN, J., EKLUND, M., SIITONEN, A. & KUUSI, M. 2011. Agricultural, socioeconomic and environmental variables as risks for human verotoxigenic Escherichia coli (VTEC) infection in Finland. *BMC Infectious Diseases,* 11**,** 275.

JENKINS, C., DALLMAN, T. J., LAUNDERS, N., WILLIS, C., BYRNE, L., JORGENSEN, F., EPPINGER, M., ADAK, G. K., AIRD, H., ELVISS, N., GRANT, K. A., MORGAN, D. & MCLAUCHLIN, J. 2015. Public Health Investigation of Two Outbreaks of Shiga Toxin-Producing Escherichia coli O157 Associated with Consumption of Watercress. *Applied Environmental Microbiology,* 81**,** 3946-52.

JENSEN, A. K., SIMONSEN, J. & ETHELBERG, S. 2017. Use of Proton Pump Inhibitors and the Risk of Listeriosis: A Nationwide Registry-based Case-Control Study. *Clinical Infectious Diseases,* 64.

JIVRAJ, S. & KHAN, O. 2013. Ethnicity and deprivation in England: How likely are ethnic minorities to live in deprived neighbourhoods? *The Dynamics of Diversity: evidence from the 2011 census.* Centre on Dynamics of Ethnicity (CoDE),.

JONES, T. F., MCMILLIAN, M. B., SCALLAN, E., FRENZEN, P. D., CRONQUIST, A. B., THOMAS, S. & ANGULO, F. J. 2007. A population-based estimate of the substantial burden of diarrhoeal disease in the United States; FoodNet, 1996-2003. *Epidemiology and Infection,* 135**,** 293-301.

KAKAI, R., WAMOLA, I. A., BWAYO, J. J. & NDINYA-ACHOLA, J. O. 1995. Enteric pathogens in malnourished children with diarrhoea. *East African Medical Journal,* 72**,** 288-9.

KAPPERUD, G., ESPELAND, G., WAHL, E., WALDE, A., HERIKSTAD, H., GUSTAVSEN, S., TVEITS, I., NATÅS, O., BEVANGER, L., DIGRANES, A. 2003. Factors associated with increased and decreased risk of Campylobacter infection: A prospective case-control study in Norway. *American Journal of Epidemiology,* 158**,** 234-242.

KARMALI, M. A. 1989. Infection by verocytotoxin-producing Escherichia coli. *Clinical Microbiology Reviews,* 2**,** 15-38.

KINNEY, J. S., GROSS, T. P., PORTER, C. C., ROGERS, M. F., SCHONBERGER, L. B. & HURWITZ, E. S. 1988. Hemolytic-uremic syndrome: a population-based study in Washington, DC and Baltimore, Maryland. *American Journal of Public Health,* 78**,** 64-5.

LAKE, I. R., HARRISON, F. C., CHALMERS, R. M., BENTHAM, G., NICHOLS, G., HUNTER, P. R., KOVATS, R. S. & GRUNDY, C. 2007. Case-control study of environmental and social factors influencing cryptosporidiosis. *European Journal of Epidemiology,* 22**,** 805-11.

LANE, C. R., LEBAIGUE, S., ESAN, O. B., AWOFISYO, A. A., ADAMS, N. L., FISHER, I. S., GRANT, K. A., PETERS, T. M., LARKIN, L., DAVIES, R. H. & ADAK, G. K. 2014. Salmonella enterica serovar Enteritidis, England and Wales, 1945-2011. *Emerging Infectious Diseases,* 20**,** 1097-104.

LAUNDERS, N., BYRNE, L., JENKINS, C., HARKER, K., CHARLETT, A. & ADAK, G. K. 2016. Disease severity of Shiga toxin-producing E. coli O157 and factors influencing the development of typical haemolytic uraemic syndrome: a retrospective cohort study, 2009–2012. *BMJ Open,* 6.

LAUNDERS, N., LOCKING, M. E., HANSON, M., WILLSHAW, G., CHARLETT, A., SALMON, R., COWDEN, J. & ADAK, G. K. 2015. A large Great Britain-wide outbreak of STEC O157 phage type 8 linked to handling of raw leeks and potatoes. *Epidemiology and Infection***,** 1-11.

LILIENFELD, D. E. 2007. Celebration: William Farr (1807–1883)—an appreciation on the 200th anniversary of his birth. *International Journal of Epidemiology,* 36**,** 985-987.

LITTLE, C. L., GORMLEY, F. J., RAWAL, N. & RICHARDSON, J. F. 2010. A recipe for disaster: outbreaks of campylobacteriosis associated with poultry liver pâté in England and Wales. *Epidemiology and Infection,* 138**,** 1691-1694.

LITTLE, C. L., RICHARDSON, J. F., OWEN, R. J., DE PINNA, E. & THRELFALL, E. J. 2008. Campylobacter and Salmonella in raw red meats in the United Kingdom: prevalence, characterization and antimicrobial resistance pattern, 2003-2005. *Food Microbiology,* 25**,** 538-43.

LLOYD-EVANS, N., DRASAR, B. S. & TOMKINS, A. M. 1983. A comparison of the prevalence of campylobacter, Shigellae and Salmonellae in faeces of malnourished and well nourished children in The Gambia and Northern Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 77**,** 245-7.

LOCKING, M. E., O'BRIEN, S. J., REILLY, W. J., WRIGHT, E. M., CAMPBELL, D. M., COIA, J. E., BROWNING, L. M. & RAMSAY, C. N. 2001. Risk factors for sporadic cases of Escherichia coli O157 infection: the importance of contact with animal excreta. *Epidemiology and Infection,* 127**,** 215-20.

LOCKING, M. E., POLLOCK, K. G., ALLISON, L. J., RAE, L., HANSON, M. F. & COWDEN, J. M. 2011. Escherichia coli O157 infection and secondary spread, Scotland, 1999-2008. *Emerging Infectious Diseases,* 17**,** 524-7.

LONG, S. M., ADAK, G. K., O'BRIEN, S. J. & GILLESPIE, I. A. 2002. General outbreaks of infectious intestinal disease linked with salad vegetables and fruit, England and Wales, 1992-2000. *Communicable Disease and Public Health,* 5**,** 101-5.

LUDVIGSSON, J. F., ON BEHALF OF THE ABIS STUDY GROUP, 2006. Epidemiological study of constipation and other gastrointestinal symptoms in 8000 children. *Acta Paediatrica, International Journal of Paediatrics,* 95**,** 573-580.

LYNN, R. M., O'BRIEN, S. J., TAYLOR, C. M., ADAK, G. K., CHART, H., CHEASTY, T., COIA, J. E., GILLESPIE, I. A., LOCKING, M. E., REILLY, W. J., SMITH, H. R., WATERS, A. & WILLSHAW, G. A. 2005. Childhood hemolytic uremic syndrome, United Kingdom and Ireland. *Emerging Infectious Diseases,* 11**,** 590-6.

MAJOWICZ, S. E., DOR, K., FLINT, J. A., EDGE, V. L., READ, S., BUFFETT, M. C., MCEWEN, S., MCNAB, W. B., STACEY, D., SOCKETT, P. & WILSON, J. B. 2004. Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community. *Epidemiology and Infection,* 132**,** 607-617.

MAJOWICZ, S. E., HORROCKS, J. & BOCKING, K. 2007. Demographic determinants of acute gastrointestinal illness in Canada: a population study. *BMC Public Health,* 7**,** 8.

MARMOT, M., ALLEN, J., GOLDBLATT, P., BOYCE, T., MCNEISH, D., GRADY, M. & GEDDES, I. 2010. Fair society, healthy lives. The Marmot Review, 2010.

MCATEER, A., ELLIOTT, A. M. & HANNAFORD, P. C. 2011. Ascertaining the size of the symptom iceberg in a UK-wide community-based survey. *British Journal of General Practice,* 61**,** e1-e11.

MCCARTHY, N. & GIESECKE, J. 2001. Incidence of Guillain-Barre syndrome following infection with Campylobacter jejuni. *American Journal of Epidemiology,* 153**,** 610-4.

MET OFFICE. 2015. *Climate Zones* [Online]. Available: http://www.metoffice.gov.uk/climate-guide/climate/zones [Accessed 18/10/2016].

MILFORD, D. V., TAYLOR, C. M., GUTTRIDGE, B., HALL, S. M., ROWE, B. & KLEANTHOUS, H. 1990. Haemolytic uraemic syndromes in the British Isles 1985-8: association with verocytotoxin producing Escherichia coli. Part 1: Clinical and epidemiological aspects. *Archives of Disease in Childhood,* 65**,** 716-21.

MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. G. & GROUP, P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine,* 6**,** e1000097.

MOLBAK, K. & NEIMANN, J. 2002. Risk factors for sporadic infection with Salmonella enteritidis, Denmark, 1997-1999. *American Journal of Epidemiology,* 156**,** 654-61.

MUSHER, D. M. & MUSHER, B. L. 2004. Contagious Acute Gastrointestinal Infections. *New England Journal of Medicine,* 351**,** 2417-2427.

NATIONAL HEALTH SERVICE. 2015. *Urgent and emergency care services in England - NHS 111 Service* [Online]. Available: http://www.nhs.uk/NHSEngland/AboutNHSservices/Emergencyandurgentcareservices/Pages/NHS-111.aspx [Accessed 23/03/2016].

NATIONAL INSTITUTE FOR HEALTH RESEARCH HEALTH PROTECTION UNIT IN GASTROINTESTINAL INFECTIONS. 2014. *About the Health Project Research Unit* [Online]. Available: http://hprugi.nihr.ac.uk/about-us/ [Accessed 09/11/2017].

NATIONAL INSTITUTE FOR HEALTH RESEARCH HEALTH PROTECTION UNIT IN GASTROINTESTINAL INFECTIONS. 2017. *Patient and Public Involvement* [Online]. Available: http://hprugi.nihr.ac.uk/patient-and-public-involvement/ [Accessed 09/11/2017].

NEAL, K. R., BARKER, L. & SPILLER, R. C. 2002. Prognosis in post-infective irritable bowel syndrome: a six year follow up study. *Gut,* 51**,** 410-413.

NEIGHBOURHOOD STATISTICS. n.d. *Super Output Areas Explained* [Online]. Available: http://neighbourhood.statistics.gov.uk/HTMLDocs/nessgeography/superoutputareasexplained/output-areas-explained.htm [Accessed 02/05/2017].

NEIMANN, J., ENGBERG, J., MOLBAK, K. & WEGENER, H. C. 2003. A case-control study of risk factors for sporadic campylobacter infections in Denmark. *Epidemiology and Infection,* 130**,** 353-366.

NEMATIAN, J., NEMATIAN, E., GHOLAMREZANEZHAD, A. & ASGARI, A. A. 2004. Prevalence of intestinal parasitic infections and their relation with socio-economic factors and hygienic habits in Tehran primary school students. *Acta Tropica,* 92**,** 179-86.

NEWMAN, K. L., LEON, J. S., REBOLLEDO, P. A. & SCALLAN, E. 2015. The impact of socioeconomic status on foodborne illness in high-income countries: a systematic review. *Epidemiology and Infection,* 143**,** 1-13.

NHS DIRECT. 2014. *What is NHS Direct? History* [Online]. Available: http://webarchive.nationalarchives.gov.uk/20140220132333/http://www.nhsdirect.nhs.uk/About/WhatIsNHSDirect/History [Accessed 02/05/2017].

NICHOLS, G. L., RICHARDSON, J. F., SHEPPARD, S. K., LANE, C. & SARRAN, C. 2012. Campylobacter epidemiology: a descriptive study

reviewing 1 million cases in England and Wales between 1989 and 2011. *BMJ Open,* 2.

NORDAHL, H., LANGE, T., OSLER, M., DIDERICHSEN, F., ANDERSEN, I., PRESCOTT, E., TJONNELAND, A., FREDERIKSEN, B. L. & ROD, N. H. 2014. Education and cause-specific mortality: the mediating role of differential exposure and vulnerability to behavioral risk factors. *Epidemiology,* 25**,** 389-96.

O'BRIEN, S. J., ADAK, G. K. & GILHAM, C. 2001. Contact with farming environment as a major risk factor for Shiga toxin (Vero cytotoxin)-producing Escherichia coli O157 infection in humans. *Emerging Infectious Diseases,* 7**,** 1049-51.

O'RYAN, M., PRADO, V. & PICKERING, L. K. 2005. A millennium update on pediatric diarrheal illness in the developing world. *Seminars in Pediatric Infectious Diseases,* 16**,** 125-36.

OFFICE FOR NATIONAL STATISTICS 2010. Standard Occupational Classification 2010 Volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010) User Manual. Office for National Statistics.

OFFICE FOR NATIONAL STATISTICS 2011. Postcode Headcounts and Household Estimates - 2011 Census.

OFFICE FOR NATIONAL STATISTICS [DATASET]. 2011. *UK population based on the 2011 Census: Key Statistics KS611UK - NS-SEC* [Online]. Office for National Statistics. Available: https://www.nomisweb.co.uk/query/select/getdatasetbytheme.asp?opt=3&theme=&subgrp= [Accessed 16/09/2016].

OGILVIE, D., FAYTER, D., PETTICREW, M., SOWDEN, A., THOMAS, S., WHITEHEAD, M. & WORTHY, G. 2008. The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Medical Research Methodology,* 8**,** 8.

OLOWOKURE, B., HAWKER, J., WEINBERG, J., GILL, N. & SUFI, F. 1999. Deprivation and hospital admission for infectious intestinal diseases. *Lancet,* 353**,** 807-8.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT 2015. List of OECD Member countries - Ratification of the Convention on the OECD.

ÖZKAN, S., TÜZÜN, H., GÖRER, N., CEYHAN, M., AYCAN, S., ALBAYRAK, S. & BUMIN, M. A. 2007. Water usage habits and the incidence of diarrhea in rural Ankara, Turkey. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 101**,** 1131-1135; 1131.

PARLIAMENTARY OFFICE OF SCIENCE AND TECHNOLOGY 2017. UK Trends in Infectious Diseases. *In:* HOUSES OF PARLIAMENT (ed.) *Postnote.*

PARRY, S. M., SALMON, R. L., WILLSHAW, G. A. & CHEASTY, T. 1998. Risk factors for and prevention of sporadic infections with vero cytotoxin (shiga toxin) producing Escherichia coli O157. *Lancet,* 351**,** 1019-22.

PATIL, S. R., CATES, S. & MORALES, R. 2005. Consumer food safety knowledge, practices, and demographic differences: findings from a meta-analysis. *Journal of Food Protection,* 68**,** 1884-94.

PAYNE, C. J., PETROVIC, M., ROBERTS, R. J., PAUL, A., LINNANE, E., WALKER, M., KIRBY, D., BURGESS, A., SMITH, R. M., CHEASTY, T., WILLSHAW, G. & SALMON, R. L. 2003. Vero cytotoxin-producing Escherichia coli O157 gastroenteritis in farm visitors, North Wales. *Emerging Infectious Diseases,* 9**,** 526-30.

PEARCE, M. S., CAMPBELL, D. I., MANN, K. D., PARKER, L. & THOMAS, J. E. 2013. Deprivation, timing of preschool infections and H. pylori seropositivity at age 49-51 years: the Newcastle Thousand Families birth cohort. *BMC Infectious Diseases,* 13**,** 422.

PEARL, D. L., LOUIE, M., CHUI, L., DORE, K., GRIMSRUD, K. M., MARTIN, S. W., MICHEL, P., SVENSON, L. W. & MCEWEN, S. A. 2009. A multi-level approach for investigating socio-economic and agricultural risk factors associated with rates of reported cases of Escherichia coli O157 in humans in Alberta, Canada. *Zoonoses and Public Health,* 56**,** 455-64.

PENNINGTON, H. 2009. The Public Inquiry into the September 2005 Outbreak of E.coli O157 in South Wales: Summary. Aberdeen: HMSO.

PERSSON, S., OLSEN, K. E., ETHELBERG, S. & SCHEUTZ, F. 2007. Subtyping method for Escherichia coli shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. *J Clin Microbiol,* 45**,** 2020-4.

PETERS, J. L., SUTTON, A. J., JONES, D. R., ABRAMS, K. R. & RUSHTON, L. 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology,* 61**,** 991-6.

PHILLIPS, G., TAM, C. C., RODRIGUES, L. C. & LOPMAN, B. 2011. Risk factors for symptomatic and asymptomatic norovirus infection in the community. *Epidemiology and Infection,* 139**,** 1676-1686.

POCKETT, R. D., ADLARD, N., CARROLL, S. & RAJORIYA, F. 2011. Paediatric hospital admissions for rotavirus gastroenteritis and infectious gastroenteritis of all causes in England: an analysis of correlation with deprivation. *Current Medical Research and Opinion,* 27**,** 777-84.

POLLARD, C. M., MENG, X., WILLIAMSON, S., DODDS, J. & BINNS, C. W. 2014. Eating out is associated with self-reported food poisoning: a Western Australia population perspective, 1998 to 2009. *Public Health Nutrition,* 17**,** 2270-2277.

POPE, D. 2015. *Introduction to systematic reviews [lecture],* Liverpool, University of Liverpool.

PUBLIC HEALTH ENGLAND. 2016. *Vero cytotoxin-producing Escherichia coli (VTEC) O157 data 2006 to 2015* [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/616284/Vero_cytotoxin-producing_Escherichia_coli_VTEC_O157_data_2006_to_2015.pdf [Accessed 09/09/2017].

PUBLIC HEALTH ENGLAND GASTROINTESTINAL BACTERIA REFERENCE UNIT. 2017. *RE: In-house data.*

PUZZOLO, E., STANISTREET, D., POPE, D., BRUCE, N. & REHFUESS, E. 2013. Factors influencing the large scale uptake by households of cleaner and more efficient household energy technologies: a systematic review. Evidence for Policy and Practice Information and Co-ordinating Centre.

PYRA, M., CONOVER, C., HOWLAND, J. & SOYEMI, K. 2012. Response to: Determinants of campylobacteriosis notifications in New Zealand. *Epidemiology and Infection,* 140**,** 2087-8.

QUETZ, J. D. S., LIMA, I. F. N., HAVT, A., DE CARVALHO, E. B., LIMA, N. L., SOARES, A. M., MOTA, R. M. S., GUERRANT, R. L. & LIMA, A. A. M. 2010. Campylobacter jejuni and Campylobacter coli in children from communities in Northeastern Brazil: molecular detection and relation to nutritional status. *Diagnostic microbiology and infectious disease,* 67**,** 220-227.

RANGEL, J. M., SPARLING, P. H., CROWE, C., GRIFFIN, P. M. & SWERDLOW, D. L. 2005. Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982–2002. *Emerging Infectious Diseases,* 11**,** 603-609.

REHFUESS, E., PUZZOLO, E., STANISTREET, S., POPE, D. & BRUCE, N. G. 2014. Enablers and barriers to large-scale uptake of improved solid fuel stoves: a systematic review. *Environmental Health Perspectives,* 122.

RIND, E. & PEARCE, J. 2010. The spatial distribution of campylobacteriosis in New Zealand, 1997-2005. *Epidemiology and Infection,* 138**,** 1359-71.

ROGERS, M. F., RUTHERFORD, G. W., ALEXANDER, S. R., DILIBERTI, J. H., FOSTER, L., SCHONBERGER, L. B. & HURWITZ, E. S. 1986. A population-based study of hemolytic-uremic syndrome in Oregon, 1979-1982. *American Journal of Epidemiology,* 123**,** 137-42.

ROSE, T. C. 2017. Impact of socioeconomic inequalities and neighbourhood characteristics on emergency hospitalisations for IID in England.

ROSE, T. C., ADAMS, N., TAYLOR-ROBINSON, D. C., BARR, B., HAWKER, J., O'BRIEN, S., VIOLATO, M. & WHITEHEAD, M. 2016. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review protocol. *Systematic Reviews,* 5**,** 13.

ROSE, T. C., ADAMS, N. L., BARR, B., HAWKER, J., O'BRIEN, S. J., VIOLATO, M., WHITEHEAD, M. & TAYLOR-ROBINSON, D. C. 2017. Socioeconomic status is associated with symptom severity and sickness absence in people with infectious intestinal disease in the UK. *BMC Infectious Diseases,* 17**,** 447.

ROWE, P. C., ORRBINE, E., LIOR, H., WELLS, G. A., YETISIR, E., CLULOW, M. & MCLAINE, P. N. 1998. Risk of hemolytic uremic syndrome after sporadic Escherichia coli O157:H7 infection: results of a Canadian collaborative study. Investigators of the Canadian Pediatric Kidney Disease Research Center. *Journal of Pediatrics,* 132**,** 777-82.

ROWE, P. C., ORRBINE, E., WELLS, G. A. & MCLAINE, P. N. 1991. Epidemiology of hemolytic-uremic syndrome in Canadian children from 1986 to 1988. The Canadian Pediatric Kidney Disease Reference Centre. *Journal of Pediatrics,* 119**,** 218-24.

ROY, S. L., DELONG, S. M., STENZEL, S. A., SHIFERAW, B., ROBERTS, J. M., KHALAKDINA, A., MARCUS, R., SEGLER, S. D., SHAH, D. D., THOMAS, S., VUGIA, D. J., ZANSKY, S. M., DIETZ, V., BEACH, M. J. & EMERGING INFECTIONS PROGRAM FOODNET WORKING, G. 2004. Risk factors for sporadic cryptosporidiosis among immunocompetent persons in the United States from 1999 to 2001. *Journal of Clinical Microbiology,* 42**,** 2944-51.

ROYAL COMMISSION ON THE HEALTH OF TOWNS 1844. First report of commissioners of inquiry into the state of large towns and populous districts. London.

SAKUMA, M., URASHIMA, M. & OKABE, N. 2006. Verocytotoxin-producing Escherichia coli, Japan, 1999-2004. *Emerging Infectious Diseases,* 12**,** 323-5.

SARGEANT, J. M., MAJOWICZ, S. E. & SNELGROVE, J. 2008. The burden of acute gastrointestinal illness in Ontario, Canada, 2005-2006. *Epidemiology and Infection,* 136**,** 451-60.

SAUNDERS, P., SAUNDERS, A. & MIDDLETON, J. 2015. Living in a 'fat swamp': exposure to multiple sources of accessible, cheap, energy-dense fast foods in a deprived community. *British Journal of Nutrition,* 113**,** 1828-34.

SCALLAN, E., FITZGERALD, M., COLLINS, C., CROWLEY, D., DALY, L., DEVINE, M., IGOE, D., QUIGLEY, T., ROBINSON, T. & SMYTH, B. 2004. Acute gastroenteritis in northern Ireland and the Republic of Ireland: a telephone survey. *Communicable Disease and Public Health,* 7**,** 61-7.

SCALLAN, E., JONES T. F., CRONQUIST A., THOMAS S., FRENZEN P., HOEFER D., MEDUS C. & ANGULO F. J. 2006. Factors associated with

seeking medical care and submitting a stool sample in estimating the burden of foodborne illness. *Foodborne Pathogens and Disease,* 3**,** 432-438; 432.

SEMENZA, J. 2010. Strategies to intervene on social determinants of infectious diseases. *Euro Surveillance,* 15.

SETHI, D., CUMBERLAND, P., HUDSON, M. J., RODRIGUES, L. C., WHEELER, J. G., ROBERTS, J. A., TOMPKINS, D. S., COWDEN, J. M. & RODERICK, P. J. 2001. A study of infectious intestinal disease in England: risk factors associated with group A rotavirus in children. *Epidemiology and Infection,* 126**,** 63-70.

SHAH, S. M. & COOK, D. G. 2008. Socio-economic determinants of casualty and NHS Direct use. *Journal of Public Health* 30**,** 75-81.

SIMONSEN, J., FRISCH, M. & ETHELBERG, S. 2008. Socioeconomic risk factors for bacterial gastrointestinal infections. *Epidemiology,* 19**,** 282-90.

SOUTHERN, J. P., SMITH, R. & PALMER, S. R. 1990. Bird attack on milk bottles: possible mode of transmission of Campylobacter jejuni to man. *Lancet,* 336**,** 1425-1427.

SPENCER, S. E., MARSHALL, J., PIRIE, R., CAMPBELL, D., BAKER, M. G. & FRENCH, N. P. 2012. The spatial and temporal determinants of campylobacteriosis notifications in New Zealand, 2001-2007. *Epidemiology and Infection,* 140**,** 1663-77.

STARE, J. & MAUCORT-BOULCH, D. 2016. Odds Ratio, Hazard Ratio and Relative Risk. *Metodoloski zvezki,* 13**,** 59-67.

STERNE, J. A., WHITE, I. R., CARLIN, J. B., SPRATT, M., ROYSTON, P., KENWARD, M. G., WOOD, A. M. & CARPENTER, J. R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal,* 338**,** b2393.

STERNE, J. A. C. & EGGER, M. 2001. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology,* 54**,** 1046-1055.

STIRLING, A. M., WILSON, P. & MCCONNACHIE, A. 2001. Deprivation, psychological distress, and consultation length in general practice. *British Journal of General Practice,* 51**,** 456-60.

STONE, D. H., MITCHELL S., PACKHAM B. & WILLIAMS J. 1994. Prevalence and first-line treatment of diarrhoealsymptoms in the community. *Public Health,* 108**,** 61-68; 61.

SUTCLIFFE, P., PICARD, L., FORTIN, B., MALAVIARACHCHI, D., HOHENADEL, J. & O'DONNELL, B. 2004. Escherichia coli O157:H7 outbreak at a summer hockey camp, Sudbury 2004. *Canadian Communicable Disease Report,* 30**,** 189-94.

SWERDLOW, D. L. & GRIFFIN, P. M. 1997. Duration of faecal shedding of Escherichia coli O157:H7 among children in day-care centres. *Lancet,* 349**,** 745-6.

TAM, C., VIVANI, L., ADAK, B., BOLTON, E., DODDS, J., COWDEN, J., EVANS, M., GRAY, J., HUNTER, P., JACKSON, K., LETLEY, L., NEAL, K., RAIT, G., SMITH, G., SMYTH, B., TOMPKINS, D., VAN DER ES, M., RODRIGUES, L., O'BRIEN, S. & ON BEHALF OF THE IID2 STUDY EXECUTIVE COMMITTEE 2011a. The second study of infectious intestinal disease in the community (IID2 Study). Food Standards Agency.

TAM, C., VIVANI, L., RODRIGUES, L. & O'BRIEN, S. 2013. The second study of infectious intestinal disease (IID2): increased rates of recurrent diarrhoea in individuals aged 65 years and above. *BMC Public Health,* 13**,** 1-8.

TAM, C. C., RODRIGUES, L. C. & O'BRIEN, S. J. 2003. The study of infectious intestinal disease in England: what risk factors for presentation to general practice tell us about potential for selection bias in case-control studies of reported cases of diarrhoea. *International Journal of Epidemiology,* 32**,** 99-105.

TAM, C. C., RODRIGUES, L. C., VIVIANI, L., DODDS, J. P., EVANS, M. R., HUNTER, P. R., GRAY, J. J., LETLEY, L. H., RAIT, G., TOMPKINS, D. S. & O'BRIEN, S. J. 2011b. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut,* 61**,** 69-77.

TARR, P. I., GORDON, C. A. & CHANDLER, W. L. 2005. Shiga-toxin-producing Escherichia coli and haemolytic uraemic syndrome. *Lancet,* 365**,** 1073-86.

TARR, P. I. & HICKMAN, R. O. 1987. Hemolytic uremic syndrome epidemiology: a population-based study in King County, Washington, 1971 to 1980. *Pediatrics,* 80**,** 41-5.

TAYLOR-ROBINSON, D. 19/09/2017. Tackling health inequalities needs to start in early years. *In:* PUBLIC HEALTH ENGLAND, ed. Annual Conference: Day two round up, 2017.

TAYLOR, E. V., HOLT, K.G., MAHON, B.E., AYERS, T., NORTON, D., GOULD, L.H. 2012. Ground beef consumption patterns in the United States, FoodNet, 2006 through 2007. *Journal of Food Protection,* 75**,** 341-346.

TESCHKE, K., BELLACK, N., SHEN, H., ATWATER, J., CHU, R., KOEHOORN, M., MACNAB, Y. C., SCHREIER, H. & ISAAC-RENTON, J. L. 2010. Water and sewage systems, socio-demographics, and duration of residence associated with endemic intestinal infectious diseases: a cohort study. *BMC Public Health,* 10**,** 767.

TEUNIS, P., TAKUMI, K. & SHINAGAWA, K. 2004. Dose response for infection by Escherichia coli O157:H7 from outbreak data. *Risk Analysis,* 24**,** 401-7.

THE METAFOR PACKAGE: A META-ANALYSIS PACKAGE FOR R. 2004. *Contor-enhanced funnel plot* [Online]. Available: http://www.metafor-project.org/doku.php/plots:contour_enhanced_funnel_plot [Accessed 19/06/2017].

THE PENNINGTON GROUP 1997. Report on the circumstances leading to the 1996 outbreak of infection with E.coli O157 in Central Scotland, the implications for food safety and lessons to be learned. Edinburgh: The Stationery Office.

THE RAMBLERS' ASSOCIATION 2010. Walking facts and figures 2: Participation in walking Participation in walking. *In:* ASSOCIATION, T. R. (ed.). London.

THOMPSON, S. G. & SHARP, S. J. 1999. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine,* 18**,** 2693-708.

TSERENPUNTSAG, B., CHANG, H.-G., SMITH, P. F. & MORSE, D. L. 2005. Hemolytic uremic syndrome risk and Escherichia coli O157:H7. *Emerging Infectious Diseases,* 11**,** 1955-7.

TUDOR HART, J. 1971. The Inverse Care Law. *Lancet***,** 405-412.

TUTTLE, J., GOMEZ, T., DOYLE, M. P., WELLS, J. G., ZHAO, T., TAUXE, R. V. & GRIFFIN, P. M. 1999. Lessons from a large outbreak of Escherichia coli O157:H7 infections: insights into the infectious dose and method of widespread contamination of hamburger patties. *Epidemiology and Infection,* 122**,** 185-92.

UCLA INSTITUTE FOR DIGITAL RESEARCH AND EDUCATION. 2017a. *Multiple Imputation in STATA* [Online]. UCLA Institute for Digital Research and Education. Available: http://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/ [Accessed 24/04/2017].

UCLA INSTITUTE FOR DIGITAL RESEARCH AND EDUCATION. 2017b. *Survival analysis with STATA* [Online]. UCLA Institute for Digital Research and Education. Available: https://stats.idre.ucla.edu/stata/seminars/stata-survival/ [Accessed 08/11/2017].

UMAN, L. S. 2011. Systematic Reviews and Meta-Analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry,* 20**,** 57-59.

UNITED NATIONS DEVELOPMENT PROGRAMME. 2016. *Human Development Index (HDI)* [Online]. Available: http://hdr.undp.org/en/content/human-development-index-hdi [Accessed 18/10/2016].

UNIVERSITY OF OXFORD. n.d. *National Perinatal Epidemiology Unit IMD Tool* [Online]. University of Oxford,. Available: https://tools.npeu.ox.ac.uk/imd/ [Accessed 13/06/2017].

VALLY, H., HALL, G., DYDA, A., RAUPACH, J., KNOPE, K., COMBS, B. & DESMARCHELIER, P. 2012. Epidemiology of Shiga toxin producing Escherichia coli in Australia, 2000-2010. *BMC Public Health,* 12**,** 63.

VAN CAUTEREN, D., DE VALK, H., VAUX, S., LE STRAT, Y. & VAILLANT, V. 2012. Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study. *Epidemiology and Infection,* 140**,** 697-705.

VAN DAMME, P., GIAQUINTO, C., HUET, F., GOTHEFORS, L., MAXWELL, M., VAN DER WIELEN, M. & GROUP., O. B. O. T. R. S. 2007. Multicenter Prospective Study of the Burden of Rotavirus Acute Gastroenteritis in Europe, 2004–2005: The REVEAL Study. *Journal of Infectious Diseases,* 195**,** S4-S16.

VARGA, C., PEARL, D. L., MCEWEN, S. A., SARGEANT, J. M., POLLARI, F. & GUERIN, M. T. 2013. Evaluating area-level spatial clustering of Salmonella Enteritidis infections and their socioeconomic determinants in the greater Toronto area, Ontario, Canada (2007 - 2009): A retrospective population-based ecological study. *BMC Public Health,* 13**,** -.

VERMA, A., BOLTON, F. J., FIEFIELD, D., LAMB, P., WOLOSCHIN, E., SMITH, N. & MCCANN, R. 2007. An outbreak of E. coli O157 associated with a swimming pool: an unusual vehicle of transmission. *Epidemiology and Infection,* 135**,** 989-92.

VOSS, P. H. & REHFUESS, E. A. 2013. Quality appraisal in systematic reviews of public health interventions: an empirical study on the impact of choice of tool on meta-analysis. *Journal of Epidemiology and Community Health,* 67**,** 98-104.

WHITEHEAD, M. 1990. The concepts and principles of equity in health. Copenhagen: World Health Organisation Regional Office for Europe.

WHITEHEAD, M. 1992. The Health Divide. *In:* TOWNSEND, P., WHITEHEAD, M. & DAVIDSON, N. (eds.) *Inequalities in Health: The Black Report and The Health Divide*. 2 ed. London: Penguin Books.

WHITEHEAD, M. 1998. Life and death over the millennium. *In:* DREVER, F. & WHITEHEAD, M. (eds.) *Health Inequalities: Dicennial Supplement*. London: The Stationery Office.

WHITEHEAD, M., BAMBRA, C., BARR, B. & AL., E. 2014. Due North: Report of the Inquiry on Health Equity for the North of England. Liverpool: University of Liverpool and Centre for Local Economic Strategies.

WHITNEY, B. M., MAINERO, C., HUMES, E., HURD, S., NICCOLAI, L. & HADLER, J. L. 2015. Socioeconomic status and foodborne pathogens in Connecticut, USA, 2000–20111. *Emerging Infectious Diseases,* 21**,** 1617-1624.

WIENKE, A. 2003. Frailty Models. *In:* MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH (ed.) *Working Papers.*

WILKING, H., HOHLE, M., VELASCO, E., SUCKAU, M. & ECKMANS, T. 2012. Ecological analysis of social risk factors for Rotavirus infections in Berlin, Germany, 2007-2009. *International Journal of Health Geographics,* 11.

WILKING, H., SPITZNAGEL, H., WERBER, D., LANGE, C., JANSEN, A. & STARK, K. 2013. Acute gastrointestinal illness in adults in Germany: a population-based telephone survey. *Epidemiology and Infection,* 141**,** 2365-75.

WILKINSON, R. & MARMOT, M. (eds.) 2003. *Social determinants of health. The solid facts*: WHO Europe.

WORLD HEALTH ORGANISATION 2012. Global report for research on infectious diseases of poverty. *In:* WORLD HEALTH ORGANISATION (ed.).

WORLD HEALTH ORGANISATION. 2017a. *Diarrhoeal Disease* [Online]. Available: http://www.who.int/mediacentre/factsheets/fs330/en/ [Accessed 23/08/2017].

WORLD HEALTH ORGANISATION. 2017b. *What are social determinants of health?* [Online]. Available: http://www.who.int/social_determinants/sdh_definition/en/ [Accessed 11/04/2017].

YOUNUS, M., HARTWICK, E., SIDDIQI, A. A., WILKINS, M., DAVIES, H. D., RAHBAR, M., FUNK, J. & SAEED, M. 2007. The role of neighborhood level socioeconomic characteristics in Salmonella infections in Michigan (1997-2007): assessment using geographic information system. *International Journal of Health Geographics [Electronic Resource]*, 6**,** 56.

ZAPPE PASTUREL, B., CRUZ-CANO, R., ROSENBERG GOLDSTEIN, R. E., PALMER, A., BLYTHE, D., RYAN, P., HOGAN, B., JUNG, C., JOSEPH, S. W., WANG, M. Q., TING LEE, M. L., PUETT, R. & SAPKOTA, A. R. 2013. Impact of rurality, broiler operations, and community socioeconomic factors on the risk of campylobacteriosis in Maryland. *American Journal of Public Health*, 103**,** 2267-75.

Appendices

Appendix 1: (Supplementary material pertaining to Chapter 3 – Methods)

Appendix 1.1: Ethical Approval Documentation

NHS
Health Research Authority

South East Coast - Surrey Research Ethics Committee
Bristol Research Ethics Committee Centre
Whitefriars
Level 3, Block B
Lewins Mead
Bristol
BS1 2NT

Telephone: 0117 342 1380

04 December 2015

Miss Natalie Adams
Public Health England
61 Colindale Avenue
Colindale, London
NW9 5EQ

Dear Miss Adams

Study title:	**Exploring whether socioeconomic factors are associated with exposure to and consequences of common gastrointestinal infections**
REC reference:	**15/LO/2138**
IRAS project ID:	**182731**

The Proportionate Review Sub-committee of the South East Coast - Surrey Research Ethics Committee reviewed the above application on 01 December 2015.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager Miss Gemma Oakes, nrescommittee.secoast-surrey@nhs.net. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of

the study.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA Approval (England)/ NHS permission for research is available in the Integrated Research Application System, www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion").

Approved documents

The documents reviewed and approved were:

Document	*Version*	*Date*
Covering letter on headed paper		23 October 2015
IRAS Checklist XML [Checklist_26112015]		26 November 2015
Letter from sponsor [Sponsorship in Principle Letter 245]	1	15 July 2015
REC Application Form [REC_Form_25112015]		25 November 2015
Research protocol or project proposal [Protocol]	Version 1	29 October 2015
Summary CV for Chief Investigator (CI) [Natalie Adams CV]		01 June 2015
Summary CV for student [Tanith Rose CV]	V1	01 June 2015
Summary CV for supervisor (student research)		

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

There were no Declarations of Interest.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

15/LO/2138	Please quote this number on all correspondence

Yours sincerely

pp Dr Mark Atkins
Chair

Email: nrescommittee.secoast-surrey@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

"After ethical review – guidance for researchers" [SL-AR2]

Copy to: *Dr Elizabeth Coates, elizabeth.coates@phe.gov.uk*

South East Coast - Surrey Research Ethics Committee

Attendance at PRS Sub-Committee of the REC meeting on 01 December 2015

Committee Members:

Name	*Profession*	*Present*	*Notes*
Dr Mark Atkins (Chair)	Consultant Virologist	Yes	
Ms Wendy Joy	Retired Support Worker	Yes	
Mrs Chrissie Lawson	Nurse Specialist	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Miss Gemma Oakes	REC Manager

Appendix 2 (Supplementary material pertaining to Chapter 4 – Study 1)

Appendix 2.1: PROSPERO International prospective register of systematic reviews

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

PROSPERO International prospective register of systematic reviews

Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review

Natalie Adams, Tanith Rose, David Taylor-Robinson, Benjamin Barr, Jeremy Hawker, Sarah O'Brien, Mara Violato, Margaret Whitehead

Citation

Natalie Adams, Tanith Rose, David Taylor-Robinson, Benjamin Barr, Jeremy Hawker, Sarah O'Brien, Mara Violato, Margaret Whitehead. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review. PROSPERO 2015:CRD42015027231 Available from http://www.crd.york.ac.uk/PROSPERO_REBRANDING/display_record.asp?ID=CRD42015027231

Review question(s)

For individuals from developed countries, does lower compared to higher socioeconomic status influence the incidence or prevalence of gastrointestinal infection?

This systematic review aims to explore the current knowledge of the relationship between socioeconomic status (SES) and gastrointestinal (GI) infections in developed countries in order to identify gaps in the literature, and areas for further research. The results of this systematic review will help to inform the development of empirical research projects by providing evidence of the methods employed previously to investigate the relationship between SES and GI infections, including the potential confounding variables used.

Searches

To identify as much relevant literature as possible, we will incorporate three systematic methods into our search strategy. We will conduct electronic database searching as well as reference list and grey literature searching as described below.

Search strategy 1: Electronic database search

We will search three databases; MEDLINE, Web of Science Core Collection and Scopus using the search terms detailed in Appendix A. The decision about which databases to search was made using a list of health sciences databases, provided by the University of Liverpool (University of Liverpool, 2015). The recommended databases have been selected. These are considered most relevant to the research question and likely to yield a high number of relevant papers.

A combination of search terms will be used, selected to provide a comprehensive range of publications on the topic. These were piloted prior to selection and comprise specific GI infections and symptom-based terms, socioeconomic and inequality terms and developed countries of interest (detailed in Appendix A). A developed country is defined as being a member country of the Organisation for Economic Co-operation and Development (OECD). The GI infection terms listed have been selected because they represent the main GI pathogens known to cause the greatest burden to public health in the developed world. Whilst not exhaustive, the list is intended to provide a broad spectrum of bacterial, viral and protozoal infections.

The search terms for MEDLINE were developed initially. Where possible, terms were exploded to broaden the search. Terms were added as keywords if they could not be exploded or if the exploded terms were not relevant to the research question. Truncation and proximity operators were also applied as necessary to broaden the search.

For consistency, the exact same terms will be used for Scopus and Web of Science Core Collection, however as the functionality of each database is different, the terms have been adapted for correct use in each.

Search strategy 2: Reference list search

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

We will search the reference lists of any studies selected for inclusion.

Search strategy 3: Grey literature search

We will search for grey literature by entering the terms "gastrointestinal infection", "gastroenteritis", "diarrhoea", "diarrhea", "socioeconomic", "social class" and "deprivation" into the Google internet search engine and Google Scholar search application and assessing the first 100 results.

Restrictions:

We will restrict to publications using data from 1980 to present only. As social conditions change over time, this will ensure that publications are as relevant as possible to the present day. We will also restrict to publications available in English. Where available, filters for "human subjects" and "document type" will be applied to the database search results. These filters directly relate to the inclusion criteria.

Additional details relating to the search strategy can be found in the attached PDF document.

Link to search strategy
http://www.crd.york.ac.uk/PROSPEROFILES/27231_STRATEGY_20150916.pdf

Types of study to be included
Inclusion:

We will include observational studies (cross-sectional, ecological, case-control, cohort (prospective and retrospective)) reporting quantitative results and analysis of empirical data on the prevalence or incidence of any symptomatic GI infection by SES in a representative population sample. Socioeconomic status can be measured at individual or aggregate level by occupation, income, education, employment or deprivation. Only studies conducted in developed countries, defined as being a member country of the Organisation for Economic Co-operation and Development (OECD), written in or translated into English, reporting on human subjects and using data collected after 1980 will be included. For countries that joined the OECD after 1980, data collection must have occurred after the date the country became a member of the OECD.

Exclusion:

Studies not meeting the above criteria, including case studies, case series, literature reviews or studies reporting on outbreaks of GI infection, travel associated illness only or asymptomatic infections only will be excluded.

Condition or domain being studied
The impact of socioeconomic status (SES) on risk, incidence or prevalence of gastrointestinal (GI) infections in developed countries.

Socioeconomic inequalities are linked to both causes and consequences of ill health (Whitehead and Dahlgren, 2006), and diseases of both infectious and non-infectious natures such as respiratory infections (Biering-Sorensen et al, 2012; Hughes and Gorton, 2015), coronary heart disease and cancer (Graham, 2009). In general, life expectancies and the prevalence of most diseases display a social gradient whereby the poorest in society experience greater levels of illness and premature death than those further up the socioeconomic scale (Wilkinson and Marmot, 2003). Socioeconomic status is measured by a variety of individual and area-based parameters, including education, income, housing, occupation and home postcode. Many infections are known to vary by social group however the role of socioeconomic inequalities in GI infection is not well understood, and it is unclear whether certain groups of society are unequally vulnerable to GI infections.

Gastrointestinal infections, caused by organisms such as bacteria, viruses or protozoa, are a common source of disease in the UK leading to diarrhoea and vomiting as well as more serious health problems. Previous studies have estimated that around 25% of people in the UK will suffer an episode of infectious intestinal disease (IID) per year and that foodborne illness (a proportion of IID) in England and Wales costs around £1.5 billion per year (Food Standards Agency, 2011). It is reported that 10% of children present to healthcare services with gastroenteritis each

year, accounting for 16% of paediatric Accident and Emergency (A&E) presentations in one study (Armon et al, 2001). There are 8 million absences from school and at least 11 million working days lost to the economy each year due to GI infections (Food Standards Agency, 2011).

Interventions and strategies to reduce infection have been implemented over the years although have had seemingly limited impact. Many infections such as tuberculosis or human immunodeficiency virus (HIV) are known to vary by social group (Semenza, 2010) however the role of socioeconomic inequalities in risk, exposure and consequences of GI infection is not well understood, with studies presenting conflicting findings. Factors which could potentially mediate the relationship between SES and GI infection include age, gender, food consumption, animal or environmental contact, urban/rural environment, foreign travel and geography. This review aims to explore the current knowledge of the relationship between SES and GI infection in developed countries, and investigate possible explanations for any differences in the risk, incidence or prevalence of GI infection across socioeconomic groups.

Participants/ population

Any individual, of any age or gender from developed countries will be included. A developed country is defined as being a member country of the Organisation for Economic Co-operation and Development (OECD).

Intervention(s), exposure(s)

The exposure of interest is lower SES, measured at the individual or aggregate level by income, education, occupation, employment or deprivation.

Comparator(s)/ control

The comparator of interest is higher SES, measured at the individual or aggregate level by income, education, occupation, employment or deprivation.

Context

Studies conducted at individual or population level using a sample which is representative of the general population will be included. Studies conducted solely in a specific population sub-group without a general population comparator group, or studies conducted in institutional settings such as nurseries, hospitals and the military will be excluded, as will studies conducted in asymptomatic individuals only.

Outcome(s)

Primary outcomes

Gastrointestinal infection:

We will investigate the relationship between SES and GI infections in developed countries. The primary outcome of interest will be the incidence or prevalence of symptomatic GI infection measured using population level surveys, routine surveillance systems, laboratory data, hospitalisation data or mortality data by SES.

Secondary outcomes

None.

Data extraction, (selection and coding)

Two authors will run the searches and collect the information required from the selected studies. Titles and abstracts will be screened independently by two authors to ensure consistency in the application of inclusion and exclusion criteria. Any discrepancies will be discussed and reviewed until an agreement is reached between both reviewers.

The full text for studies deemed relevant after title and abstract review will be sought and reviewed in the same way. Where full texts are not available, they will be sought via institutional library sharing agreements.

All full text studies will be reviewed independently by two reviewers (T.R and N.A) to ensure that they conform to the inclusion and exclusion criteria.

Data to be extracted: aim/hypothesis; study design; level of analysis; country; sample size; age; age category; GI measure; GI way of measure and data source; socioeconomic status measure; socioeconomic status way of measure and data source; covariates; significant results; non-significant results; conclusions.

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

Risk of bias (quality) assessment
Risk of bias and quality assessment will be conducted by two reviewers (T.R and N.A). The Liverpool University Quality Assessment Tool (LQAT) will be used for this review, which will allow studies to be assessed using a validated tool specific to each study design. Any discrepancies identified will be discussed and reviewed.

Strategy for data synthesis
We will conduct a narrative data synthesis, including tabulation of the studies for comparison. We will analyse and synthesise the data by assigning codes to the data, linking concepts, core ideas or trends together (Machi and McEvoy, 2009). All measures of outcome will be considered. It is anticipated that there will be significant heterogeneity between studies however, where possible, meta-analysis will be conducted on combined results.

Analysis of subgroups or subsets
We will analyse the following subgroups using a narrative thematic approach. Where homogenous data allows, a meta-analysis will be conducted on a subset of identified studies. RevMan software will be used to conduct this analysis.

Pathogen type:

We will collect information on pathogen type where available to examine whether the relationship between GI infection and SES varies by pathogen type.

Age:

We will collect information on the age of the study subjects where available to examine whether the relationship between GI infection and SES varies by age.

Measurement(s) of gastrointestinal infection:

We will collect information on the methods used to measure GI infections such as surveillance systems, GP presentation data, hospital admissions or survey data to evaluate whether the methods used to measure GI infection have an impact on the results.

Measurement(s) of socioeconomic status:

We will collect information on the methods used to measure SES such as income, education, occupation or area-level deprivation to evaluate whether the methods used to measure SES have an influence on the results.

Country:

We will collect information on the country in order to take into account the healthcare systems in place and how they might influence the results.

Level of analysis:

We will collect information on the level at which the study is conducted, aggregate or individual, as this might influence the results.

Dissemination plans
The systematic review will be submitted for publication. Data will also be presented at conferences and will contribute to two PhD projects.

Contact details for further information
Tanith Rose

Department of Public Health and Policy,

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for
Health Research

Institute of Psychology, Health and Society,

University of Liverpool,

Office 314, Whelan Building,

Liverpool, L69 3GB,

United Kingdom

tanith.rose@liverpool.ac.uk

Organisational affiliation of the review
University of Liverpool, Department of Public Health and Policy; Public Health England; University of Oxford

Review team
Miss Natalie Adams, University of Liverpool, Public Health England
Mrs Tanith Rose, University of Liverpool
Dr David Taylor-Robinson, University of Liverpool
Mr Benjamin Barr, University of Liverpool
Professor Jeremy Hawker, Public Health England
Professor Sarah O'Brien, University of Liverpool
Dr Mara Violato, University of Oxford
Professor Margaret Whitehead, University of Liverpool

Anticipated or actual start date
01 May 2015

Anticipated completion date
01 July 2016

Funding sources/sponsors
National Institute for Health Research, Health Protection Research Unit in Gastrointestinal Infections

Conflicts of interest
None known

Language
English

Country
England

Subject index terms status
Subject indexing assigned by CRD

Subject index terms
Communicable Diseases; Developed Countries; Gastrointestinal Tract; Humans; Social Class

Stage of review
Ongoing

Date of registration in PROSPERO
19 October 2015

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

Date of publication of this revision
25 May 2016

DOI
10.15124/CRD42015027231

Stage of review at time of this submission	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	Yes
Formal screening of search results against eligibility criteria	Yes	Yes
Data extraction	Yes	Yes
Risk of bias (quality) assessment	Yes	Yes
Data analysis	Yes	No

Appendix 2.2: Systematic review protocol

Rose et al. Systematic Reviews (2016) 5:13
DOI 10.1186/s13643-016-0187-7

Systematic Reviews

PROTOCOL Open Access

Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review protocol

Tanith C. Rose[1,3,7*], Natalie Adams[1,3,5], David C. Taylor-Robinson[1,3], Benjamin Barr[1,3], Jeremy Hawker[1,5], Sarah O'Brien[1,4], Mara Violato[2,6] and Margaret Whitehead[1,3]

Abstract

Background: The association between low socioeconomic status (SES) and poor health is well documented in the existing literature. Nonetheless, evidence on the relationship between SES and gastrointestinal (GI) infections is limited, and the mechanisms underlying this relationship are not well understood with published studies pointing to conflicting results. This review aims to identify studies that investigate the relationship between SES and GI infections in developed countries, in order to assess the direction of the association and explore possible explanations for any differences in the risk, incidence or prevalence of GI infections across socioeconomic groups.

Methods: Three systematic methods will be used to identify relevant literature: electronic database, reference list and grey literature searching. The databases MEDLINE, Scopus and Web of Science Core Collection will be searched using a broad range of search terms. Screening of the results will be performed by two reviewers using pre-defined inclusion and exclusion criteria. The reference lists of included studies will be searched, and Google will be used to identify grey literature. Observational studies reporting quantitative results on the prevalence or incidence of any symptomatic GI infections by SES, in a representative population sample from a member country of the Organisation for Economic Co-operation and Development (OECD), will be included. Data will be extracted using a standardised form. Study quality will be assessed using the Liverpool University Quality Assessment Tools (LQAT). A narrative synthesis will be performed including tabulation of studies for comparison.

Discussion: This systematic review will consolidate the existing knowledge on the relationship between SES and GI infections. The results will help to identify gaps in the literature and will therefore provide an evidence base for future empirical studies to deepen the understanding of the relationship, including effective study design and appropriate data analysis methods. Ultimately, gaining insight into this relationship will help to inform policies to reduce any health inequalities identified.

Systematic review registration: PROSPERO CRD42015027231

Keywords: Socioeconomic factors, Income, Social class, Employment, Education, Gastrointestinal infection, Diarrhoea, Gastroenteritis, Foodborne diseases

* Correspondence: Tanith.Rose@liverpool.ac.uk
Tanith C. Rose and Natalie Adams are joint first authors.
[1]NIHR Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, Liverpool L69 7BE, UK
[3]Department of Public Health and Policy, University of Liverpool, Liverpool, UK
Full list of author information is available at the end of the article

Background

There is strong evidence of a social gradient in most health outcomes whereby the poorest in society experience greater levels of illness and premature death than those further up the socioeconomic scale [1]. Socioeconomic inequalities are linked to both causes and consequences of ill health [2] and have been well documented in diseases of a non-infectious nature, such as coronary heart disease and cancer [3]. Whilst there is evidence that the incidence of many infectious diseases, such as tuberculosis and human immunodeficiency virus [4–6], varies by social group, the association between socioeconomic status (SES) and gastrointestinal (GI) infections in particular is not well understood.

Gastrointestinal infections, caused by organisms such as bacteria, viruses or protozoa, are a common source of disease in the UK, leading to diarrhoea and vomiting and potentially more serious health problems, all of which can interfere with normal daily life. Previous studies have estimated that around 25 % of people in the UK will suffer an episode of infectious intestinal disease (IID) per year and that foodborne illness (a proportion of IID) in England and Wales costs around £1.5 billion per annum [7, 8]. It is reported that 10 % of children present to healthcare services with gastroenteritis each year, accounting for 16 % of paediatric accident and emergency presentations in one study [9]. There are eight million absences from school and at least 11 million working days lost to the economy each year due to GI infections [7].

The impact of SES on vulnerability to GI infections is unclear, and the limited existing evidence points to conflicting results. Higher prevalence of GI infections is often thought to be associated with more advantaged individuals. However, a recent systematic review looking at the impact of SES on laboratory-confirmed foodborne illness in developed countries suggests that this relationship is not so clear [10]. Newman et al. [10] identified 16 studies across four pathogens with mixed results, differing by pathogen. For example, in the most disadvantaged populations compared to the least disadvantaged, *Listeria* was more common, but *Campylobacter* was less common. In addition to the papers identified by Newman et al. [10], inconsistent results have also been observed among studies that have used syndromic definitions of GI infections, with some reporting higher rates of GI infections among those in lower socioeconomic groups [4, 11, 12] and others observing the opposite [13, 14]. These results clearly demonstrate the disagreements within this area of research.

A number of factors could explain these inconsistent results. The studies identified thus far cover a broad range of pathogens, and it may be that the relationship differs depending on whether the data are analysed at an all-GI-infection, pathogen-specific or species-specific level. This might suggest that the mode of transmission of an organism plays a role in the relationship and that this could be related to potentially socially patterned risk such as rural versus urban residency or exotic foreign travel. Furthermore, these studies have used different study designs, measured SES and GI infection in different ways and controlled for various confounders (such as age, labour market attachment, country of birth and agricultural occupation). Appropriate adjustment of confounding variables requires an understanding of the underlying mechanisms linking SES to GI infection risk, but there is little empirical evidence in this area.

A systematic review is warranted to summarise, organise and make sense of the contradictory findings observed in the literature. Our review aims to build on previous work by exploring the relationship between SES and a full range of GI infections. As it is possible that various socioeconomic or healthcare-seeking behavioural factors could influence whether an individual is diagnosed with a GI infection, we have also included syndromic definitions of GI infections. We aim to explore the current knowledge of the relationship in developed countries; assess the magnitude, statistical significance and direction of the association; and shed light into possible explanations for any observed differences in the risk, incidence or prevalence of GI infections across socioeconomic groups. The results of this review will help to inform the development of empirical research projects by identifying gaps in the literature and areas where further research is required. It will provide evidence of the methods employed previously to investigate the relationship between SES and GI infections, including information on the relevant confounding variables used.

Methods/design

To improve the transparency and completeness of the protocol, a completed copy of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols 2015 (PRISMA-P 2015) checklist [15] can be found in Additional file 1.

Research question

For individuals from developed countries, is lower compared to higher SES associated with the incidence or prevalence of GI infection?

Population

Any individual, of any age or gender, from a developed country will be included. A developed country is defined as being a member country of the Organisation for Economic Co-operation and Development (OECD). The OECD aims to continually monitor the economic

developments of its 34 member countries and provides policy recommendations to help governments tackle poverty through economic growth and stability [16].

Exposure

The exposure of interest is lower compared to higher SES, measured at the individual or aggregate level by income, education, occupation, employment or deprivation of area of residence.

Outcome

The primary outcome of interest will be the incidence or prevalence of any symptomatic GI infection measured using population level surveys, routine surveillance systems, laboratory data or hospitalisation data and includes syndromic definitions of GI infections without a laboratory diagnosis.

Inclusion/exclusion criteria

Observational studies (cross-sectional, ecological, case-control, cohort [prospective and retrospective]) reporting quantitative results and analysis of empirical data on the prevalence or incidence of any symptomatic GI infection by SES, in a representative population sample, will be included. Socioeconomic status can be measured by occupation, income, education, employment or deprivation at the individual or aggregate level. Only studies conducted in developed countries (defined as being a member country of the OECD), written in or translated into English, reporting on human subjects and using data collected after 1980, will be included. For countries that joined the OECD after 1980, data collection must have occurred after the date the country became a member of the OECD. Studies not meeting the above criteria, including case studies, case series or literature reviews, or studies reporting on outbreaks of GI infection, travel-associated illness only or asymptomatic infections only will be excluded. Studies conducted solely in a specific population subgroup without a general population comparator group or studies conducted in institutional settings such as nurseries, hospitals or the military will be excluded.

Search strategy

Three search strategies will be used to identify as much relevant literature as possible. Firstly, the electronic searching of three databases will be performed: MEDLINE (Ovid), Scopus and Web of Science Core Collection. The choice of database was discussed with a university librarian, and the three databases chosen were considered most relevant to the research question and likely to yield the highest number of relevant papers.

The search terms were piloted prior to selection and are comprised of specific GI infection and symptom-based terms, socioeconomic and inequality terms, and developed countries of interest (Additional file 2). Relevant synonyms for the SES and GI infection terms were identified using Roget's Thesaurus online [17] and the thesaurus in MEDLINE by mapping and inspecting the tree for each term. Relevant terms mentioned in articles identified in a pilot search of the literature were also added. Ultimately, the GI infection terms were selected because they represent the main GI pathogens known to cause the greatest burden to public health in the developed world. Whilst not exhaustive, the list is intended to provide a broad spectrum of bacterial, viral and protozoal infections.

The search terms for MEDLINE were developed initially. Where possible, terms were exploded to broaden the search. Terms were added as keywords if they could not be exploded or if the exploded terms were not relevant to the research question. Truncation and proximity operators were also applied as necessary to broaden the search. Terms were combined using Boolean operators.

For consistency, the exact same terms were used for Scopus and Web of Science Core Collection; however, as the functionality of each database is different, it was necessary to adapt the terms developed in MEDLINE for correct use in Scopus and Web of Science Core Collection. Specifically, the terms contained within the exploded terms in MEDLINE needed to be added as individual search terms for use in Scopus and Web of Science Core Collection, and it was necessary to indicate phrases with quotation marks. Additionally, the proximity operators differed for each database.

When the searches are run in Scopus and Web of Science Core Collection, each term will be searched for within the title, abstract and keywords of the documents contained in each database. Filters within the three databases will be applied to restrict the results to publications that have used data from 1980 to the present. As social conditions within countries change over time through development, and methods of classifying SES are also modified over time, restricting to publications using data from 1980 onwards will ensure that the results are as relevant as possible to the present day. Results will also be limited to publications available in the English language. Additionally, where available, filters for 'human subjects' and 'document type' will be applied to the database search results. All of these filters directly relate to the inclusion criteria. The publications remaining after the filters are applied will then be exported into reference managing software. In this software, the publications from the three databases will be combined and duplicates removed. The remaining publications will then be screened for relevance using the inclusion and exclusion criteria.

Titles and abstracts of the publications will be screened independently by two authors (NA and TR) to

ensure consistency in the application of the inclusion and exclusion criteria. Any discrepancies will be discussed and re-examined until an agreement is reached between both reviewers. The full text for studies deemed relevant after title and abstract screening will be retrieved and reviewed in the same way. Where full texts are not available, they will be sought via institutional library sharing agreements. All full-text studies will be screened independently by the same two reviewers to ensure that they conform to the inclusion and exclusion criteria.

The second strategy will consist of searching the reference lists of any studies selected for inclusion in the final review to identify potentially relevant articles that may have been missed by the electronic database searches. The abstracts of any references considered potentially relevant will be sought and screened for inclusion using the pre-defined inclusion and exclusion criteria. The full text for studies deemed relevant after title and abstract screening will be retrieved and reviewed in the same way. This reference list search will be conducted independently by two reviewers (NA and TR), and discrepancies will be discussed and eventually agreed upon at each stage.

The third method will be to conduct a search of the grey literature by entering the terms 'gastrointestinal infection', 'gastroenteritis', 'diarrhoea', 'diarrhea', 'socioeconomic', 'social class', 'income' and 'deprivation' into the Google internet search engine and the Google Scholar search application and assessing the first 100 results. Each result will be inspected for relevance using the inclusion an exclusion criteria. Again, this will be performed independently by the two reviewers (NA and TR), and disagreement will be resolved through discussion.

Quality assessment

Risk of bias and quality assessment of the identified studies will be conducted by the review team, independently and then reconciled. The Liverpool University Quality Assessment Tool (LQAT) will be used for this review, which will allow the methodological quality of the studies to be assessed using a tool specific to each study design [18]. It incorporates a star rating system to assess and qualify absence of bias, misclassification and confounding. The LQAT has been used in previous systematic reviews [19, 20] and has been independently evaluated against other quality assessment tools [21]. Any discrepancies between reviewers in the quality assessment of the studies will be discussed and re-examined.

Data analysis and synthesis

To organise these data and to facilitate comparison, tables will be created by extracting data from each study into a standardised Excel spreadsheet. Data to be extracted will include the following: aim/hypothesis, study design, level of analysis, country, sample size, age, age category, type of GI infection, GI infection method of measurement and data source, measure of SES, SES method of measurement and data source, covariates, statistically significant results, non-significant results, conclusions and quality assessment. Extracted data will be checked for accuracy by at least one other reviewer.

Due to the broad scope of this review, it is anticipated that there will be considerable heterogeneity between studies in terms of design, populations studied and the measurement of primary exposures and outcomes. The synthesis strategy will be driven by the data available; however, to explore the relationship between GI infections and SES, it is anticipated that a subgroup analysis will be performed on study design factors and potential moderating factors of the relationship, including but not limited to the following: pathogen type (based on mode of transmission); age; country (based on climate and relative level of development); methods used to measure GI infection; methods used to measure SES; and level of analysis (aggregate or individual). Separate tables will be created to compare and contrast the results of studies within and between the subgroups. If the data allow, further grouping of the studies within the subgroups will be performed to help summarise the study findings and answer the research question. The LQAT results will be used to determine the strength of the evidence from individual studies, and greater weight will be given to conclusions drawn from the most methodologically robust and reliable studies. A narrative synthesis will help to make sense of what is anticipated to be a diverse body of evidence and may lead to potential explanations for the contrasting findings observed in the literature. The methods used will be written up transparently, and the robustness of the synthesis will be assessed [22].

Where homogenous data allow, meta-analyses will be conducted on combined results. The synthesis strategy outlined above will assist in identifying data suitable for meta-analysis. Heterogeneity will be assessed by examining the forest plots to detect overlapping confidence intervals, using the chi^2 test with a P value of 0.10 to indicate statistical significance, and also applying the I^2 statistic with values of 30 to 60 %, 50 to 90 % and 75 to 100 % used to denote moderate, substantial and considerable levels of heterogeneity, respectively [23]. If the data allow, publication bias will be assessed using a funnel plot, and sensitivity analysis on the basis of study quality will be conducted to explore the robustness of the meta-analysis. RevMan software will be used to conduct these analyses [24]. A 'Summary of findings' table [25] will be used to present the results, and the Grading of Recommendations, Assessments, Development and Evaluation approach will be used to assess the quality of the body of evidence [26].

Dissemination

The systematic review will be submitted for publication. The findings of the review and data will be presented at conferences and will contribute to two PhD projects as part of the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections [27].

Discussion

Our systematic review aims to provide new insight into the understanding of the mixed results on the relationship between SES and GI infections as suggested by Newman et al. [10], by broadening the focus to a wider range of symptomatic GI infections and exploring whether a more conclusive pattern can be identified. This includes syndromic definitions of GI infections in the absence of laboratory confirmation. By including these definitions, we aim to identify literature on the burden of symptoms by SES and attempt to capture population groups who may not seek healthcare for their illness and consequently may not be included in studies which use laboratory data to identify cases only. This is particularly important for this review as the decision to seek healthcare may be related to SES.

In the UK, it is estimated that 17 million cases of infectious intestinal disease occur every year, resulting in approximately one million general practice consultations [7]. This, coupled with an increasingly overburdened National Health Service (NHS), highlights the importance of understanding the role of SES in GI infections in order to devise policies to target the strata of the population most at risk.

The results of this review will provide a more comprehensive evidence base of the relationship between symptomatic GI infections and SES to inform the development of empirical studies, including effective study design and appropriate data analysis methods, which will be used in two PhD projects.

Additional files

Additional file 1: PRISMA-P 2015 checklist. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols 2015 (PRISMA-P 2015) checklist was used to develop this protocol. Items 1b and 4 were not applicable.

Additional file 2: Search terms for MEDLINE, Scopus and Web of Science Core Collection. The search terms that will be used to identify relevant literature across three databases.

Abbreviations

GI: gastrointestinal; IID: infectious intestinal disease; LQAT: Liverpool University Quality Assessment Tools; NHS: National Health Service; NIHR HPRU: National Institute for Health Research Health Protection Research Unit; OECD: Organisation for Economic Co-operation and Development; PRISMA-P 2015: Preferred Reporting Items for Systematic Reviews and Meta-analyses for Protocols 2015; SES: socioeconomic status.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NA and TR wrote the protocol. DTR, BB, JH, SOB, MV and MW conceived the initial idea for the study, critically appraised the protocol and also contributed to its development by revising different versions. All authors approved the final version and take responsibility for its content.

Authors' information

NA and TR are PhD students funded by the National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections at the University of Liverpool.

Funding

This research is funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections at the University of Liverpool in partnership with Public Health England (PHE), University of East Anglia, University of Oxford and the Institute of Food Research. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or PHE.

Author details

[1]NIHR Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, Liverpool L69 7BE, UK. [2]NIHR Health Protections Research Unit in Gastrointestinal Infections, University of Oxford, Oxford, UK. [3]Department of Public Health and Policy, University of Liverpool, Liverpool, UK. [4]Institute of Infection and Global Health, University of Liverpool, Liverpool, UK. [5]Public Health England, London, UK. [6]Health Economics Research Centre, University of Oxford, Oxford, UK. [7]Department of Public Health and Policy, Institute of Psychology, Health and Society, University of Liverpool, Whelan Building, Liverpool L69 3GB, UK.

Received: 21 December 2015 Accepted: 12 January 2016
Published online: 21 January 2016

References

1. Wilkinson R, Marmot M. Social determinants of health: the solid facts. 2nd ed. Geneva: World Health Organisation; 2003. http://www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf. Accessed 17 Feb 2015.
2. Whitehead M, Dahlgren G. Concepts and principles for tackling social inequities in health: levelling up part 1. Copenhagen: World Health Organisation Regional Office for Europe; 2006. http://www.enothe.eu/cop/docs/concepts_and_principles.pdf. Accessed 17 Feb 2015.
3. Graham H. Understanding health inequalities. Maidenhead: Open University Press, McGraw-Hill Education; 2009.
4. Biering-Sorensen S, Sondergaard G, Vitting Andersen K, Andersen AM, Mortensen LH. Time trends in socio-economic factors and risk of hospitalisation with infectious diseases in pre-school children 1985–2004: a Danish register-based study. Paediatr Perinat Epidemiol. 2012;26(3):226–35.
5. Hughes GJ, Gorton R. Inequalities in the incidence of infectious disease in the North East of England: a population-based study. Epidemiol Infect. 2015;143(1):189–201.
6. Semenza JC. Strategies to intervene on social determinants of infectious diseases. Euro Surveill. 2010;15(27):32–9.
7. Food Standards Agency. The second study of infectious intestinal disease in the community (IID2 Study). https://www.food.gov.uk/science/research/foodborneillness/b14programme/b14projlist/b18021. Accessed 15 Dec 2015.
8. Food Standards Agency. Foodborne disease strategy 2010–15: an FSA programme for the reduction of foodborne disease in the UK 2011. http://www.food.gov.uk/sites/default/files/multimedia/pdfs/fds2015.pdf. Accessed 16 Dec 2015.
9. Armon K, Stephenson T, Gabriel V, MacFaul R, Eccleston P, Werneke U, et al. Determining the common medical presenting problems to an accident and emergency department. Arch Dis Child. 2001;84(5):390–2.
10. Newman KL, Leon JS, Rebolledo PA, Scallan E. The impact of socioeconomic status on foodborne illness in high-income countries: a systematic review. Epidemiol Infect. 2015;143(12):2473–85.
11. Olowokure B, Hawker J, Weinberg J, Gill N, Sufi F. Deprivation and hospital admission for infectious intestinal diseases. Lancet. 1999;353(9155):807–8.

12. Pockett RD, Adlard N, Carroll S, Rajoriya F. Paediatric hospital admissions for rotavirus gastroenteritis and infectious gastroenteritis of all causes in England: an analysis of correlation with deprivation. Curr Med Res Opin. 2011;27(4):777–84.
13. de Wit MAS, Koopmans MPG, Kortbeek LM, Wannet WJB, Vinje J, Van Leusden F, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol. 2001;154(7):666–74.
14. Pollard CM, Meng X, Williamson S, Dodds J, Binns CW. Eating out is associated with self-reported food poisoning: a Western Australia population perspective, 1998 to 2009. Public Health Nutr. 2014;17(10):2270–7.
15. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Prefered Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015: elaboration and explanation. Br Med J. 2015;2015:1–25.
16. Organisation for Economic Co-operation and Development. What we do and how. http://www.oecd.org/about/whatwedoandhow/. Accessed 26 Oct 2015.
17. Thesaurus.com. http://www.thesaurus.com/. Accessed 15 Dec 2015.
18. Pope D. Introduction to systematic reviews [lecture]. Liverpool: University of Liverpool; 2015.
19. Puzzolo E, Stanistreet S, Pope D, Bruce NG, Rehfuess EA. Factors influencing the large scale uptake by households of cleaner and more efficient household energy technologies. A systematic review. Evidence for Policy and Practice Information and Co-ordinating Centre, 2013. http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3426. Accessed 20 Oct 2015.
20. Rehfuess E, Puzzolo E, Stanistreet S, Pope D, Bruce NG. Enablers and barriers to large-scale uptake of improved solid fuel stoves: a systematic review. Environ Health Perspect. 2014;122(2):120–30.
21. Voss PH, Rehfuess EA. Quality appraisal in systematic reviews of public health interventions: an empirical study on the impact of choice of tool on meta-analysis. J Epidemiol Community Health. 2013;67(1):98–104.
22. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M. Guidance on the conduct of narrative synthesis in systematic reviews: a product of the ESRC Methods Programme. Version 1. Swindon: Economic and Social Research Council; 2006.
23. Identifying and measuring heterogeneity. http://handbook.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm. Accessed 17 Dec 2015.
24. Review Manager (RevMan) [Computer program]. Version 5.3. The Nordic Cochrane Centre, The Cochrane Collaboration; 2014
25. General template for 'Summary of findings' tables. http://handbook.cochrane.org/chapter_11/11_5_3_general_template_for_summary_of_findings_tables.htm. Accessed 17 Dec 2015.
26. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490.
27. National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections. http://www.hpru.ginihr.ac.uk/. Accessed 15 Dec 2015.

Appendix 2.3: Systematic review search terms

Medline (Ovid)

#1. Exp Socioeconomic Factors/
#2. Education*.mp.
#3. Exp Employment/
#4. Income*.mp.
#5. Occupation*.mp.
#6. Poverty.mp.
#7. Poorest.mp.
#8. exp Social Class/
#9. Inequalit*.mp.
#10. Socioeconomic*.mp.
#11. Depriv*.mp.
#12. Disadvantag*.mp.
#13. Salary.mp.
#14. Underprivileged.mp.
#15. Social determinant*.mp.
#16. (Social adj1 factor*).mp
#17. Socio*.mp

#18. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17

#19. exp Norovirus/
#20. Acute gastroenteritis.mp.
#21. infectious intestinal disease*.mp.
#22. gastrointestinal infection*.mp.
#23. exp Diarrhea/
#24. Rotavirus.mp.
#25. gastrointestinal pathogen*.mp.
#26. gastrointestinal bacteria.mp.
#27. enteric infection*.mp.
#28. diarrh*.mp.
#29. stomach flu.mp.
#30. gastric flu.mp.
#31. stomach bug*.mp.
#32. stomach virus*.mp.
#33. Exp Campylobacter/
#34. Exp Escherichia coli/
#35. Enterobacteriaceae Infection*.mp.
#36. Dysentery, Bacillary.mp
#37. Exp Escherichia coli Infections/
#38. Yersinia enterocolitica.mp.
#39. Exp Salmonella Infections/
#40. Exp Cryptosporidiidae/
#41. Exp Salmonella/
#42. Exp Shigella/
#43. Exp Giardia/
#44. Escherichia coli.mp.
#45. Exp Listeria/
#46. Small round structured virus*.mp.

#47. Winter vomiting disease*.mp.
#48. Sapovirus.mp.
#49. Caliciviridae.mp.
#50. VTEC.mp.
#51. STEC.mp.
#52. exp Foodborne Diseases/
#53. Food poisoning*.mp.
#54. Scombro*.mp.
#55. Clostridium perfringens.mp.
#56. Bacillus cereus.mp.
#57. Hepatitis A.mp.
#58. Hepatitis E.mp.

#59. 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58

#60. exp Australia/
#61. exp Austria/
#62. exp Belgium/
#63. exp Canada/
#64. exp Chile/
#65. exp Czech Republic/
#66. exp Denmark/
#67. exp Estonia/
#68. exp Finland/
#69. exp France/
#70. exp Germany/
#71. exp Greece/
#72. exp Hungary/
#73. exp Iceland/
#74. exp Ireland/
#75. exp Israel/
#76. exp Italy/
#77. exp Japan/
#78. exp Korea/
#79. exp Luxembourg/
#80. exp Mexico/
#81. exp Netherlands/
#82. exp New Zealand/
#83. exp Norway/
#84. exp Poland/
#85. exp Portugal/
#86. exp Slovak Republic/
#87. exp Slovenia/
#88. exp Spain/
#89. exp Sweden/
#90. exp Switzerland/
#91. exp Turkey/
#92. exp United Kingdom/
#93. exp United States/

#94. 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93

#95. 18 and 59 and 94

	Scopus - TITLE-ABS-KEY	WOS Core Collection - Topic
#1.	"Career mobility"	
#2.	Poverty	
#3.	"Social class*"	
#4.	"Social mobility"	
#5.	Education*	
#6.	Employment	
#7.	Unemployment	
#8.	Income*	
#9.	Occupation*	
#10.	Poor*	
#11.	Inequalit*	
#12.	Depriv*	
#13.	Disadvantag*	
#14.	Salary	
#15.	Underprivileged	
#16.	"Social determinant*"	
#17.	Social pre/1 factor*	Social near/1 factor*
#18.	Socio*	
#19.	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18	
#20.	Norovirus	
#21.	"Norwalk virus"	
#22.	"Acute gastroenteritis"	
#23.	"Infectious intestinal disease*"	
#24.	"Gastrointestinal infection*"	
#25.	Rotavirus	
#26.	"Gastrointestinal pathogen*"	
#27.	"Gastrointestinal bacteria"	
#28.	"Enteric infection*"	
#29.	Diarrh*	
#30.	"Stomach flu"	
#31.	"Gastric flu"	
#32.	"Stomach bug*"	
#33.	"Stomach virus*"	
#34.	"Escherichia coli"	
#35.	"Enterobacteriaceae Infection*"	
#36.	Dysentery Bacillary	
#37.	"Yersinia enterocolitica"	
#38.	"paratyphoid fever"	

#39. "typhoid fever"
#40. "Small round structured virus*"
#41. "Winter vomiting disease*"
#42. Sapovirus
#43. Caliciviridae
#44. Campylobacter*
#45. Cryptospor*
#46. Salmonell*
#47. Shigell*
#48. Giardia*
#49. Listeri*
#50. VTEC
#51. STEC
#52. "Foodborne Disease*"
#53. Botulism
#54. "Staphylococcal Food Poisoning*"
#55. "Food poisoning*"
#56. Scombro*
#57. "Clostridium perfringens"
#58. "Bacillus cereus"
#59. "Hepatitis A"
#60. "Hepatitis E"

#61. 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60

#62. Australia*
#63. "New South Wales"
#64. "Northern Territory"
#65. Queensland
#66. Tasmania
#67. Victoria
#68. Austria
#69. Belgium
#70. Canada
#71. Alberta
#72. "British Columbia"
#73. Manitoba
#74. "New Brunswick"
#75. "Newfoundland and Labrador"
#76. "Northwest Territories"
#77. "Nova Scotia"
#78. Nunavut
#79. Ontario
#80. "Prince Edward Island"
#81. Quebec
#82. Saskatchewan
#83. "Yukon Territory"

#84. Chile
#85. "Czech Republic"
#86. Denmark
#87. Greenland
#88. Estonia
#89. Finland
#90. France
#91. Paris
#92. Germany
#93. Berlin
#94. Greece
#95. Hungary
#96. Iceland
#97. Ireland
#98. Israel
#99. Italy
#100. Rome
#101. Sicily
#102. Japan
#103. Tokyo
#104. Korea
#105. Seoul
#106. Luxembourg
#107. Mexico
#108. Netherlands
#109. "New Zealand"
#110. Norway
#111. Svalbard
#112. Poland
#113. Portugal
#114. "Slovak Republic"
#115. Slovakia
#116. Slovenia
#117. Spain
#118. Sweden
#119. Switzerland
#120. Turkey
#121. "United Kingdom"
#122. "Great Britain"
#123. "Channel Islands"
#124. Guernsey
#125. England
#126. London
#127. Scotland
#128. Hebrides
#129. Wales
#130. "United States"
#131. "Appalachian Region"
#132. Alabama

#133. Georgia
#134. Kentucky
#135. Maryland
#136. “New York”
#137. Carolina
#138. Ohio
#139. Pennsylvania
#140. Tennessee
#141. Virginia
#142. “Great Lakes Region”
#143. Illinois
#144. Chicago
#145. Indiana
#146. Michigan
#147. Minnesota
#148. Wisconsin
#149. “Mid-Atlantic Region”
#150. Delaware
#151. “District of Columbia”
#152. Baltimore
#153. “New Jersey”
#154. Philadelphia
#155. Iowa
#156. Kansas
#157. Missouri
#158. Nebraska
#159. Dakota
#160. Oklahoma
#161. “New England”
#162. Connecticut
#163. Maine
#164. Massachusetts
#165. Boston
#166. “New Hampshire”
#167. “Rhode Island”
#168. Vermont
#169. Idaho
#170. Montana
#171. Oregon
#172. Washington
#173. Wyoming
#174. “Pacific States”
#175. Alaska
#176. California
#177. “Los Angeles”
#178. “San Francisco”
#179. Hawaii
#180. Arkansas
#181. Florida

#182. Louisiana
#183. "New Orleans"
#184. Mississippi
#185. Arizona
#186. Colorado
#187. Nevada
#188. "New Mexico"
#189. Texas
#190. Utah

#191. 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114 or 115 or 116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124 or 125 or 126 or 127 or 128 or 129 or 130 or 131 or 132 or 133 or 134 or 135 or 136 or 137 or 138 or 139 or 140 or 141 or 142 or 143 or 144 or 145 or 146 or 147 or 148 or 149 or 150 or 151 or 152 or 153 or 154 or 155 or 156 or 157 or 158 or 159 or 160 or 161 or 162 or 163 or 164 or 165 or 166 or 167 or 168 or 169 or 170 or 171 or 172 or 173 or 174 or 175 or 176 or 177 or 178 or 179 or 180 or 181 or 182 or 183 or 184 or 185 or 186 or 187 or 188 or 189 or 190

#192. 19 and 61 and 191

Appendix 2.4: Included studies

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Adlam	2011	High	New Zealand	All	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Income	Y	
Arena	2014	Medium	France	Adults	<200	Individual	Case-control	Multi-pathogen	GP Presentation	Multiple measures	Y	
Arsenault	2012	Low	Canada	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Education	Y	
Baker	1998	Low	UK/Ireland	Children	5001-10,000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Education	Y	
Banatvala	1999	Low	UK/Ireland	All	>100,000	Area	Ecological	Salmonellosis	Laboratory records	Deprivation	Y	
Barros	2003	Low	Portugal	Children	200-1000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Education	Y	
Beale	2010	Low	UK/Ireland	Children	5001-10,000	Area	Cohort	Acute GI infection (syndromic)	Population based survey	Social class	N	Analysed same individuals as Baker 1998
Beaudry	1995	Low	Canada	Children	200-1000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Social class	Y	
Bemis	2014	Low	United States	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Deprivation	Y	
Bessell	2010	Low	UK/Ireland	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Deprivation	Y	
Biering-Sorensen	2012	High	Denmark	Children	>100,000	Individual	Cohort	Acute GI infection (syndromic)	Hospital admission	Education	Y	
Bless	2014	High	Switzerland	All	200-1000	Individual	Case-control	Campylobacteriosis	Laboratory records	Education	Y	
Borgnolo	1996	High	Italy	Children	200-1000	Individual	Case-control	Salmonellosis	Hospital admission	Occupation	Y	
Bozkurt	2003	Low	Turkey	Children	200-1000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Bozkurt	1999	Low	Turkey	Children	200-1000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Britton	2010	Low	New Zealand	Not specified	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Deprivation	Y	
Chang	2009	High	United States	All	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Multiple measures	N	Did not use a dichotomous outcome

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Cohen	2008	Low	United States	All	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Income	Y	
Danis	2009	Low	UK/Ireland	All	200-1000	Individual	Case-control	Campylobacteriosis	Laboratory records	Employment	Y	
de Wit	2001	Medium	Netherlands	All	1001-5000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Education	Y	
de Wit	2003	Low	Netherlands	All	200-1000	Individual	Case-control	Multi-pathogen	Population based survey	Education	N	Subset of de Wit 2001
Dennehy	2006	High	United States	Children	1001-5000	Individual	Case-control	Rotavirus	Hospital admission	Education	Y	
Doorduyn	2012	High	Netherlands	All	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Education	Y	
Doré	2004	High	Canada	All	200-1000	Individual	Case-control	Salmonellosis	Laboratory records	Education	Y	
Duggirala	2005	Low	United States	All	200-1000	Individual	Case-control	Hepatitis A	Laboratory records	Education	Y	
Eaton-Evans	1987	Low	Australia	Children	<200	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Occupation	N	Insufficient quantitative data
Ethelberg	2006	Medium	Denmark	Children	1001-5000	Individual	Case-control	Acute GI infection (syndromic)	Laboratory records	Multiple measures	Y	
Etiler	2004	Low	Turkey	Children	200-1000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Evans	2006	Medium	UK/Ireland	Adults	10,001-100,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Employment	Y	
Faustini	2006	Medium	Italy	All	<200	Individual	Case-control	Giardiasis	Laboratory records	Education	Y	
Fein	1995	High	United States	Adults	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Education	Y	
Fewtrell	1997	Low	UK/Ireland	All	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Employment	N	Did not use a dichotomous outcome
Friedman	2004	Low	United States	All	1001-5000	Individual	Case-control	Campylobacteriosis	Laboratory records	Multiple measures	Y	
Fullerton	2007	Medium	United States	Children	1001-5000	Individual	Case-control	Campylobacteriosis	Laboratory records	Multiple measures	Y	
Gillespie	2010	Low	UK/Ireland	All	>100,000	Area	Ecological	Listeriosis	Laboratory records	Deprivation	Y	

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Gillespie	2008	Low	UK/Ireland	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Occupation	Y	
Green	2006	High	Canada	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Social class	Y	
Gupta	2004	Low	United States	All	>100,000	Area	Ecological	Shigellosis	Laboratory records	Deprivation	N	Insufficient quantitative data
Hall	2006	High	Australia	All	5001-10,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Herikstad	2002	Medium	United States	All	5001-10,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	N	Analysed same individuals as Jones 2007
Hu	2009	Low	Australia	Not specified	>100,000	Area	Ecological	Cryptosporidiosis	Laboratory records	Multiple measures	Y	
Hu	2010	Low	Australia	Not specified	>100,000	Area	Ecological	Cryptosporidiosis	Laboratory records	Social class	N	Did not use a dichotomous outcome
Hughes	2015	Low	UK/Ireland	All	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Deprivation	Y	
Iacono	2005	High	Italy	Children	1001-5000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Education	N	Insufficient quantitative data
Jackson	2015	Low	United States	Children	>100,000	Area	Ecological	Shigellosis	Laboratory records	Deprivation	Y	
Jalava	2011	Low	Finland	All	>100,000	Area	Ecological	STEC	Laboratory records	Multiple measures	Y	
Jones	2007	Medium	United States	All	10,001-100,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Education	Y	
Kass	1992	Low	United States	All	200-1000	Individual	Case-control	Salmonellosis	Laboratory records	Income	Y	
Kotloff	1988	Medium	United States	Children	200-1000	Individual	Case-control	Acute GI infection (syndromic)	Hospital admission	Education	N	Insufficient quantitative data
Kum-Nji	2009	High	United States	Children	200-1000	Individual	Cohort	Acute GI infection (syndromic)	Hospital admission	Employment	Y	
Kyle	2011	Low	UK/Ireland	Children	>100,000	Area	Ecological	Acute GI infection (syndromic)	Hospital admission	Deprivation	N	Did not use a dichotomous outcome
Lake	2007	Low	UK/Ireland	All	5001-10,000	Area	Ecological	Cryptosporidiosis	Laboratory records	Occupation	Y	
Lake	2009	Low	UK/Ireland	All	>100,000	Area	Ecological	Cryptosporidiosis	Laboratory records	Social class	N	Insufficient quantitative data
Lal	2012	Medium	New Zealand	All	>100,000	Area	Ecological	Salmonellosis	Multiple measures	Deprivation	Y	

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Lee	1991	Low	United States	All	>100,000	Area	Ecological	Shigellosis	Laboratory records	Deprivation	N	Insufficient quantitative data
Ludvigsson	2006	Low	Sweden	Children	5001-10,000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Education	Y	
MacRitchie	2013	Low	UK/Ireland	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Deprivation	N	Insufficient quantitative data
Majowicz	2007	Medium	Canada	All	5001-10,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
McAteer	2011	High	UK/Ireland	Adults	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	N	Insufficient quantitative data
McPherson	2009	Medium	Australia	All	200-1000	Individual	Case-control	STEC	Laboratory records	Education	Y	
Moorin	2010	Medium	Australia	All	>100,000	Area	Cohort	Acute GI infection (syndromic)	Hospital admission	Social class	Y	
Neal	1997	Low	UK/Ireland	Adults	200-1000	Individual	Case-control	Campylobacteriosis	Laboratory records	Social class	Y	
Nichols	2012	Low	UK/Ireland	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Deprivation	N	Insufficient quantitative data
Odoi	2004	Medium	Canada	Not specified	>100,000	Area	Ecological	Giardiasis	Laboratory records	Income	Y	
Olowokure	1999	Low	UK/Ireland	All	>100,000	Area	Ecological	Acute GI infection (syndromic)	Hospital admission	Deprivation	Y	
Özkan	2007	Low	Turkey	All	200-1000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Income	Y	
Özmert	2008	Low	Turkey	Children	200-1000	Individual	Case-control	Acute GI infection (syndromic)	Hospital admission	Education	Y	
Pardhan-Ali	2013	Low	Canada	All	10,001-100,000	Area	Ecological	Multi-pathogen	Laboratory records	Multiple measures	N	Insufficient quantitative data
Pearl	2009	Low	Canada	All	>100,000	Area	Ecological	STEC	Laboratory records	Income	N	Insufficient quantitative data
Penrose	2007	Low	United States	All	>100,000	Area	Ecological	Giardiasis	Laboratory records	Income	N	Insufficient quantitative data
Phillips	2011	Low	UK/Ireland	Children	200-1000	Individual	case-control	Norovirus	GP Presentation	Occupation	Y	
Pockett	2011	Medium	UK/Ireland	Children	>100,000	Area	Ecological	Acute GI infection (syndromic)	Hospital admission	Deprivation	N	Did not use a dichotomous outcome
Pollard	2014	High	Australia	Adults	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Education	Y	

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Pyra	2012	Low	United States	Adults	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Multiple measures	Y	
Quigley	2006	Medium	UK/Ireland	Children	200-1000	Individual	Case-control	Acute GI infection (syndromic)	GP Presentation	Occupation	N	Data analysed within Phillips 2011, Rodrigues 2001 and Sethi 2001
Rind	2010	High	New Zealand	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Social class	N	Did not use a dichotomous outcome
Rodrigues	2001	Medium	UK/Ireland	All	200-1000	Individual	Case-control	Campylobacteriosis	GP Presentation	Employment	Y	
Sakuma	2006	Low	Japan	All	>100,000	Area	Ecological	STEC	Laboratory records	Income	N	Insufficient quantitative data
Sargeant	2008	Medium	Canada	All	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Satterthwaite	1999	Medium	New Zealand	All	200-1000	Individual	Case-control	Yersinia enterocolitica	Laboratory records	Education	Y	
Scallan	2004	High	UK/Ireland	All	5001-10,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Occupation	Y	
Seo	2013	Medium	Korea	All	1001-5000	Individual	Case-control	Hepatitis A	Hospital admission	Multiple measures	Y	
Seo	2012	High	Korea	Not specified	>100,000	Area	Ecological	Hepatitis A	Laboratory records	Multiple measures	N	Insufficient quantitative data
Sethi	2001	Medium	UK/Ireland	Children	200-1000	Individual	Case-control	Rotavirus	GP Presentation	Occupation	Y	
Simonsen	2008	Medium	Denmark	All	>100,000	Individual	Cohort	Multi-pathogen	Laboratory records	Multiple measures	Y	
Spencer	2012	Low	New Zealand	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Deprivation	Y	
Stafford	1996	Low	Australia	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Social class	Y	
Stone	1994	Low	UK/Ireland	All	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Employment	N	Insufficient quantitative data
Tam	2013	High	UK/Ireland	All	5001-10,000	Individual	Cohort	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Teschke	2010	Medium	Canada	All	>100,000	Area	Cohort	Acute GI infection (syndromic)	Multiple measures	Income	Y	
Turkish Ministry of Health	1995	Low	Turkey	Children	1001-5000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Education	Y	

First Author	Year	Quality	Country/Region	Age Group	Sample Size	Level – individual/ area	Study Design	Pathogen/symptom	GI Measure	SES Measure	Included in meta-analysis	Reason for exclusion
Unicomb	2008	Medium	Australia	All	200-1000	Individual	Case-control	Campylobacteriosis	Laboratory records	Multiple measures	Y	
Van Cauteren	2012	Medium	France	All	10,001-100,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Multiple measures	Y	
Varga	2013	Medium	Canada	All	>100,000	Area	Ecological	Salmonellosis	Laboratory records	Multiple measures	Y	
Weisent	2012	Low	United States	Not specified	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Multiple measures	Y	
Whitney	2015	Low	United States	All	>100,000	Area	Ecological	Multi-pathogen	Laboratory records	Deprivation	Y	
Wilking	2013	Medium	Germany	Adults	10,001-100,000	Individual	Cross-sectional	Acute GI infection (syndromic)	Population based survey	Income	Y	
Wilking	2012	Low	Germany	All	>100,000	Area	Ecological	Rotavirus	Hospital admission	Employment	Y	
Xu	2015	Low	Australia	Children	>100,000	Area	Ecological	Acute GI infection (syndromic)	Hospital admission	Social class	Y	
Younus	2007	Low	United States	All	>100,000	Area	Ecological	Salmonellosis	Laboratory records	Multiple measures	Y	
Younus	2010	Low	United States	Children	200-1000	Individual	Case-control	Salmonellosis	Laboratory records	Multiple measures	Y	
Zappe Pasturel	2013	Low	United States	All	>100,000	Area	Ecological	Campylobacteriosis	Laboratory records	Multiple measures	Y	

Appendix 2.5: Excluded studies with rationale

First Author	Year	Reason for exclusion
Admoni	1995	Non OECD at time of data collection
Aksoy	2007	Asymptomatic participants potentially included
Allerberger	1996	No analysis by disadvantaged compared to advantaged
Alter	1982	Outcome not acute-GI infection
Alter	1989	Outcome not acute-GI infection
Alvarado-Esquivel	2014	Asymptomatic participants potentially included
Alvarez-Muñoz	1999	Asymptomatic participants potentially included
Andrews	2003	No analysis by disadvantaged compared to advantaged
Araya	1986	Non OECD at time of data collection
Art	1991	No analysis by disadvantaged compared to advantaged
Baaten	2007	Asymptomatic participants potentially included
Baker	2012	Outcome not acute-GI infection
Bell	2005	Asymptomatic participants potentially included
Bilenko	1999	Non OECD at time of data collection
Blackman	1989	No analysis by disadvantaged compared to advantaged
Bojalil	1994	Non OECD at time of data collection
Bollag	1980	No analysis by disadvantaged compared to advantaged
Bolumar	1995	Asymptomatic participants potentially included
Boreham	1981	Asymptomatic participants potentially included
Bozkurt	2001	No analysis by disadvantaged compared to advantaged
Brandt	1983	No analysis by disadvantaged compared to advantaged
Brieseman	1985	No analysis by disadvantaged compared to advantaged
Brieseman	1990	No analysis by disadvantaged compared to advantaged
Britt	2005	No analysis by disadvantaged compared to advantaged
Broner	2010	Outcome not acute-GI infection
Bryant	1989	No analysis by disadvantaged compared to advantaged
Bura	2012	Asymptomatic participants potentially included
Burström	2005	Data collected pre-1980
Buti	2006	Asymptomatic participants potentially included
Bytzer	2001	Outcome not acute-GI infection
Calderon	1990	Not available in English
Callery	2010	Outcome not acute-GI infection
Cantey	2011	No population comparison
Cedillo-Rivera	2009	Asymptomatic participants potentially included
Çeliksöz	2005a	Asymptomatic participants potentially included
Çeliksöz	2005b	Asymptomatic participants potentially included
Ceran	2012	Asymptomatic participants potentially included
Cevahir	2013	Asymptomatic participants potentially included
Ceyhan	2008	Asymptomatic participants potentially included
Charlett	2003	No analysis by disadvantaged compared to advantaged
Chiaramonte	1983	Asymptomatic participants potentially included
Cifuentes	1999	Outcome not acute-GI infection
Cifuentes	2002	Study participants duplicated elsewhere
Cilla	2010	No analysis by disadvantaged compared to advantaged
Colak	2002	Asymptomatic participants potentially included
Collins	2015	Outcome not acute-GI infection
Connelly	2014	Outcome not acute-GI infection
Conway	1990	Outcome not acute-GI infection
Cross	2009	No analysis by disadvantaged compared to advantaged
Davila	2009	No analysis by disadvantaged compared to advantaged
Davis	2013	No analysis by disadvantaged compared to advantaged

First Author	Year	Reason for exclusion
De Wit	2001b	No analysis by disadvantaged compared to advantaged
Deveci	2014	Asymptomatic participants potentially included
Diel	2001	No analysis by disadvantaged compared to advantaged
Ditah	2014	Asymptomatic participants potentially included
Domínguez	2007	Asymptomatic participants potentially included
Doni	2015	No analysis by disadvantaged compared to advantaged
Dostal	2001	No analysis by disadvantaged compared to advantaged
Drobeniuc	2013	No analysis by disadvantaged compared to advantaged
Dubnov	2004	Non OECD at time of data collection
Ekramul Hoque	2002	No analysis by disadvantaged compared to advantaged
Ellencweig	1986	Non OECD at time of data collection
elSaadany	2002	No analysis by disadvantaged compared to advantaged
Erdogan	2012	Asymptomatic participants potentially included
Erdogan	2004	Asymptomatic participants potentially included
Ersoy	1998	Asymptomatic participants potentially included
Escobedo	2011	No analysis by disadvantaged compared to advantaged
Esparza-Aguilar	2014	Outcome not acute-GI infection
Esparza-Aguilar	2013	Outcome not acute-GI infection
Faulkner	2003	Asymptomatic participants potentially included
Feeney	1998	Outcome not acute-GI infection
Fewtrell	1994	Study participants duplicated elsewhere
Finkelman	1994	Non OECD at time of data collection
Ford-Jones	2000	No analysis by disadvantaged compared to advantaged
Forman	1984	No analysis by disadvantaged compared to advantaged
Forsberg	2009	Review/case report/RCT
Francis	1984	No analysis by disadvantaged compared to advantaged
Fraser	1998	Non OECD at time of data collection
Frost	2004	Asymptomatic participants potentially included
Galanis	2014	No population comparison
Gangarosa	1992	Outcome not acute-GI infection
Gastañaduy	2013	No analysis by disadvantaged compared to advantaged
Gibson	1985	Outcome not acute-GI infection
Gomes	2011	Outcome not acute-GI infection
Graham	1988	No analysis by disadvantaged compared to advantaged
Grimwood	2006	No population comparison
Guerrero	2004	No analysis by disadvantaged compared to advantaged
Haley	2010	No analysis by disadvantaged compared to advantaged
Hannah	2005	Outcome not acute-GI infection
Harter	1982	No analysis by disadvantaged compared to advantaged
Hayes-Bautista	1994	No analysis by disadvantaged compared to advantaged
Hemmelgarn	1993	No analysis by disadvantaged compared to advantaged
Hepworth	2010	Study participants duplicated elsewhere
Hizo-Abes	2013	No population comparison
Ho	1988	Non OECD at time of data collection
Hoque	2002	No analysis by disadvantaged compared to advantaged
Howell	2006	Outcome not acute-GI infection
Huerta	2006	No analysis by disadvantaged compared to advantaged
Hughes	2013	Outcome not acute-GI infection
Ikram	1994	No analysis by disadvantaged compared to advantaged
Imhoff	2004	Study participants duplicated elsewhere
Jayasinghe	2013	No analysis by disadvantaged compared to advantaged
Jiménez-Moleón	2011	No analysis by disadvantaged compared to advantaged
Jones	2007	No analysis by disadvantaged compared to advantaged

First Author	Year	Reason for exclusion
Jones	1994	Review/case report/RCT
Julio	2012	Asymptomatic participants potentially included
Kanra	2002	Asymptomatic participants potentially included
Kapperud	2003	No analysis by disadvantaged compared to advantaged
Kapperud	1995	No analysis by disadvantaged compared to advantaged
Karaman	2015	No analysis by disadvantaged compared to advantaged
Kaya	2007	Asymptomatic participants potentially included
Koopman	1989	Asymptomatic participants potentially included
Krebs	2011	Non OECD at time of data collection
Kruszon-Moran	2005	Asymptomatic participants potentially included
Kuhls	1994	No analysis by disadvantaged compared to advantaged
Kurugöl	2003	No population comparison
Kurugöl	2011	Asymptomatic participants potentially included
Kurugöl	2009	Asymptomatic participants potentially included
Kyle	2012	Outcome not acute-GI infection
Lazcano-Ponce	2013	Asymptomatic participants potentially included
Leach	1999	Asymptomatic participants potentially included
Leach	2000	Asymptomatic participants potentially included
Lee	2013	No analysis by disadvantaged compared to advantaged
LeJeune	2010	Review/case report/RCT
Lerman	1999	No analysis by disadvantaged compared to advantaged
Leshem	2015	No analysis by disadvantaged compared to advantaged
Letaief	2005	Asymptomatic participants potentially included
Levesque	2013	No analysis by disadvantaged compared to advantaged
Levine	1993	No analysis by disadvantaged compared to advantaged
Levy	1998	No analysis by disadvantaged compared to advantaged
Liddle	1997	No analysis by disadvantaged compared to advantaged
Long	2006	Review/case report/RCT
Lopman	2012	Outcome not acute-GI infection
Lupo	1989	No analysis by disadvantaged compared to advantaged
Ma	2009	No analysis by disadvantaged compared to advantaged
Maguire	1995	No analysis by disadvantaged compared to advantaged
Majowicz	2004	Study participants duplicated elsewhere
Maltezou	2001	No population comparison
Manasek	2004	No analysis by disadvantaged compared to advantaged
Maral	2010	Asymptomatic participants potentially included
Markus	2011	Asymptomatic participants potentially included
Martínez-García	1989	Non OECD at time of data collection
Masia	2004	No analysis by disadvantaged compared to advantaged
McCann	2002	No analysis by disadvantaged compared to advantaged
McQuillan	2004	Asymptomatic participants potentially included
Medeiros	2006	Review/case report/RCT
Mehal	2012	Outcome not acute-GI infection
Mor	2015	Asymptomatic participants potentially included
Morales	1992	Asymptomatic participants potentially included
Morris	1983	Data collected pre-1980
Moyo	2014	Non OECD at time of data collection
Muhsen	2014	Non OECD at time of data collection
Muhsen	2010	Non OECD at time of data collection
Mullner	2010	No analysis by disadvantaged compared to advantaged
Naess	2012	No population comparison
Najnin	2014	No analysis by disadvantaged compared to advantaged
Nathwani	1995	Review/case report/RCT

First Author	Year	Reason for exclusion
Nelson	1985	Review/case report/RCT
Newbold	2013	No analysis by disadvantaged compared to advantaged
Newman	1999	No quantitative results
Nicoll	2000	No analysis by disadvantaged compared to advantaged
Noone	2000	No analysis by disadvantaged compared to advantaged
North	1999	No analysis by disadvantaged compared to advantaged
Novotny	1990	Asymptomatic participants potentially included
Ochnio	2005	Asymptomatic participants potentially included
Okur	2011	Not available in English
Okyay	2004	Asymptomatic participants potentially included
Olcay	2003	Asymptomatic participants potentially included
Olsen	2006	No analysis by disadvantaged compared to advantaged
Omurtag	2013	No analysis by disadvantaged compared to advantaged
Ong	2012	No analysis by disadvantaged compared to advantaged
Ostan	2007	Asymptomatic participants potentially included
Ozturk	2004	No analysis by disadvantaged compared to advantaged
Ozturk	1996	Outcome not acute-GI infection
Painter	2015	No analysis by disadvantaged compared to advantaged
Palti	1984	Non OECD at time of data collection
Parashar	1998	Outcome not acute-GI infection
Pasquini	1984	Asymptomatic participants potentially included
Patel	2015	Asymptomatic participants potentially included
Peasey	2004	Asymptomatic participants potentially included
Pérez-Rubio	2011	No analysis by disadvantaged compared to advantaged
Pollock	2006	No analysis by disadvantaged compared to advantaged
Potter	2003	No analysis by disadvantaged compared to advantaged
Psichogiou	1995	No analysis by disadvantaged compared to advantaged
Quihui	2010	Asymptomatic participants potentially included
Quihui	2006	Asymptomatic participants potentially included
Quihui-Cota	2015	Asymptomatic participants potentially included
Quinlan	2013	Review/case report/RCT
Rajan	1998	Asymptomatic participants potentially included
Redlinger	1998	No analysis by disadvantaged compared to advantaged
Redlinger	1997	Asymptomatic participants potentially included
Redlinger	2002	No analysis by disadvantaged compared to advantaged
Rees	1995	No analysis by disadvantaged compared to advantaged
Ricotta	2014	No analysis by disadvantaged compared to advantaged
Rishpon	1984	Non OECD at time of data collection
Roos	2005	Outcome not acute-GI infection
Russo	1997	Outcome not acute-GI infection
Sandberg	2006	No analysis by disadvantaged compared to advantaged
Scavia	2012	No analysis by disadvantaged compared to advantaged
Schmeer	2009	No analysis by disadvantaged compared to advantaged
Sénécal	2008	No population comparison
Sepulveda	1988	Non OECD at time of data collection
Silk	2012	Review/case report/RCT
Snel	2009a	Study participants duplicated elsewhere
Snel	2009b	No analysis by disadvantaged compared to advantaged
Spencer	1988	Non OECD at time of data collection
Stafford	2007	No analysis by disadvantaged compared to advantaged
Standeart	2015	No analysis by disadvantaged compared to advantaged
Stone	1993	No population comparison
Strauss	2001	No analysis by disadvantaged compared to advantaged

First Author	Year	Reason for exclusion
Stroffolini	1990	Asymptomatic participants potentially included
Stroffolini	1991	Asymptomatic participants potentially included
Stroffolini	1989	Asymptomatic participants potentially included
Stroffolini	1996	Outcome not acute-GI infection
Studahl	2000	No analysis by disadvantaged compared to advantaged
Talbot-Smith	2002	No analysis by disadvantaged compared to advantaged
Tappero	1995	No analysis by disadvantaged compared to advantaged
Termorshuizen	2000	Asymptomatic participants potentially included
Teshale	2015	Asymptomatic participants potentially included
Thomas	1993	Asymptomatic participants potentially included
Thomas	2008	No analysis by disadvantaged compared to advantaged
Thomas	2011	Non OECD at time of data collection
Thoren	1995	No analysis by disadvantaged compared to advantaged
Thoren	1988	No analysis by disadvantaged compared to advantaged
Thornley	2002	No analysis by disadvantaged compared to advantaged
Thrane	2005	No quantitative results
To	1996	Outcome not acute-GI infection
Tosun	2004	Asymptomatic participants potentially included
Uhlig	2014	No analysis by disadvantaged compared to advantaged
Vancelik	2006	Asymptomatic participants potentially included
Vasquez-Garibay	2015	Asymptomatic participants potentially included
Verma	2014	Non OECD at time of data collection
Vranckx	1990	Data collected pre-1980
Vranckx	1984	Asymptomatic participants potentially included
Vrbova	2012	No analysis by disadvantaged compared to advantaged
Vulcano	2007	Asymptomatic participants potentially included
Warburton	1994	No analysis by disadvantaged compared to advantaged
Withers	2002	No analysis by disadvantaged compared to advantaged
Yapicioglu	2002	Asymptomatic participants potentially included
Yoon	2014	Asymptomatic participants potentially included
Zaidi	2012	No analysis by disadvantaged compared to advantaged
Ziv	2011	Non OECD at time of data collection

Appendix 2.6: Data extraction fields

Field
Search origin
First author
Year of publication
Aim/hypothesis
Study design
Level of analysis
Country
Sample size
Age (years)
Age category
GI measured
Method of sampling cases
GI way of measurement and data source
General SES
SES measure
SES data source
Covariates controlled for
Significant results GI infection & SES ($p<0.05$)
Nonsignificant results GI infection & SES
Conclusion1: lower SES is associated with increased risk, incidence or prevalence of GI
Conclusion2: no significant relation GI & SES
Conclusion3: higher SES is associated with increased risk, incidence or prevalence of GI
Conclusion4: This hypothesis was not tested
Conclusion5: other
Notes
Quality

Appendix 2.7: Liverpool Quality Assessment Tools

Source: Pope (2015)

STUDY ID		EXTRACTED BY		EXTRACTION DATE	DD	MM	YY

GAPP: METHODOLOGICAL QUALITY APPRAISAL

PART 3: CROSS SECTIONAL STUDIES

PART A: Study Sample

QUALITY 1 – SELECTION BIAS

Is there evidence of selection bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 2 – RESPONSE BIAS

Is there evidence of response bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART B: Exposure Assessment SES

QUALITY 3 – Direct exposure

Ranking of exposure measurement?	
SES measured and analysed per group	
SES measured per individual based on aggregate SES data e.g. IMD	*
SES measured per individual directly e.g. income, education	**

QUALITY 4 – RECALL BIAS

Is there evidence of recall bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 5 – MEASUREMENT BIAS

Is there evidence of measurement bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART C: Outcome Assessment

QUALITY 6 – Assessment of GI infection

Ranking of outcome assessment?	
Self-reported recall	
Self-reported prospectively	*
Health record/physician diagnosed via symptoms	**
Laboratory confirmation	***

QUALITY 7 – BIAS IN ASCERTAINMENT

Is there evidence of ascertainment bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART D: Analysis of Results

QUALITY 8 – ADJUSTMENT FOR CONFOUNDING

Is there adjustment for confounding?	
No	Limited or no adjustment for confounding.
Yes - adequate	The main confounders adjusted for: age, sex (*)
Yes - good	Majority of known confounders in model (**)

PART E: Qualifying comments

QUALITY 9: indicate overall assessment and specific issues that you would like to draw attention to:

Total Stars	/12

STUDY ID		EXTRACTED BY		EXTRACTION DATE	DD	MM	YY

GAPP: METHODOLOGICAL QUALITY APPRAISAL

PART 4: CASE-CONTROL STUDIES

PART A: Study Sample

QUALITY 1 – CASE SELECTION BIAS

Is there evidence of selection bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 2 – CONTROL SELECTION BIAS

Is there evidence of response bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART B: Exposure Assessment SES

QUALITY 3 – Direct exposure

Ranking of exposure measurement?	
SES measured and analysed per group	
SES measured per individual based on aggregate SES data e.g. IMD	*
SES measured per individual directly e.g. income, education	**

QUALITY 4 – RECALL BIAS

Is there evidence of recall bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 5 – MEASUREMENT BIAS

Is there evidence of measurement bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART C: Outcome Assessment

QUALITY 6 – Assessment of GI infection

Ranking of outcome assessment?	
Self-reported symptoms recall	
Self-reported symptoms prospectively	*
Health record/physician diagnosed via symptoms	**
Laboratory confirmation	***

PART D: Analysis of Results

QUALITY 7 – ADJUSTMENT FOR CONFOUNDING
(Including matching at design stage)

Is there adjustment for confounding?	
No	Limited or no adjustment for confounding.
Yes - adequate	The main confounders adjusted for: age, sex (*)
Yes - good	Majority of known confounders in model (**)

PART E: Qualifying comments

QUALITY 8: indicate overall assessment and specific issues that you would like to draw attention to:

Total Stars	/11

STUDY ID		EXTRACTED BY		EXTRACTION DATE	DD	MM	YY

GAPP: METHODOLOGICAL QUALITY APPRAISAL

PART 5: COHORT STUDIES

PART A: Study Sample

QUALITY 1 – SELECTION BIAS

Is there evidence of selection bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 2 – RESPONSE BIAS

Is there evidence of response bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

QUALITY 3 – BIAS IN FOLLOW-UP

Is there evidence of bias in follow up?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART B: Exposure Assessment SES

QUALITY 4 – Direct exposure

Ranking of exposure measurement?	
SES measured and analysed per group	
SES measured per individual based on aggregate SES data e.g. IMD	*
SES measured per individual directly e.g. income, education	**

QUALITY 5 – MEASUREMENT BIAS

Is there evidence of measurement bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART C: Outcome Assessment

QUALITY 6 – Assessment of GI infection

Ranking of outcome assessment?	
Self-reported recall	
Self-reported prospectively	*
Health record/physician diagnosed via symptoms	**
Laboratory confirmation	***

QUALITY 7 – BIAS IN ASCERTAINMENT

Is there evidence of ascertainment bias?	
Yes	
Possible	
No	*
If Yes/possible, provide details:	

PART D: Analysis of Results

QUALITY 8 – ADJUSTMENT FOR CONFOUNDING

Is there adjustment for confounding?	
No	Limited or no adjustment for confounding.
Yes - adequate	The main confounders adjusted for: age, sex (*)
Yes - good	Majority of known confounders in model (**)

PART E: Qualifying comments

QUALITY 9: indicate overall assessment and specific issues that you would like to draw attention to:

Total Stars	/12

Appendix 2.8: Summary data for studies included in meta-analysis (n=77)

Study	Estimate	Lower CI	Upper CI	SE	Measure	Adjusted
Adlam 2011	1.11	0.67	2	0.278986	Relative Risk	Yes
Arena 2014	0.42	0.15	1.16	0.521821	Odds Ratio	Yes
Arsenault 2012	0.85	0.72	1	0.084203	Relative Risk	Yes
Baker 1998	1.32	1.19	1.45	0.050411	Odds Ratio	Yes
Banatvala 1999	0.79	0.7	0.9	0.064311	Odds Ratio	No
Barros 2003	1.59	0.85	2.94	0.316563	Odds Ratio	No
Beaudry 1995	0.68	0.29	1.64	0.441982	Rate Ratio	No
Bemis 2014	0.6	0.57	0.63	0.027352	Relative Risk	Yes
Bessell 2010	0.78	0.75	0.81	0.022206	Relative Risk	Yes
Biering-Sørensen 2012	1.52	1.44	1.61	0.028467	Hazard Ratio	Yes
Bless 2014	2	1.25	3.33	0.249956	Odds Ratio	Yes
Borgnolo 1996	1.86	1.03	3.38	0.303142	Odds Ratio	Yes
Bozkurt 1999	1.12	0.5	2.46	0.40464	Odds Ratio	No
Bozkurt 2003	2.07	1.59	2.71	0.136766	Relative Risk	No
Britton 2010	0.76	0.7	0.82	0.040995	Rate Ratio	Yes
Cohen 2008	0.44	0.24	0.8	0.305204	Odds Ratio	Yes
Cohen 2008	0.35	0.11	1.08	0.57852	Odds Ratio	Yes
Danis 2009	0.89	0.37	2.16	0.451019	Odds Ratio	No
de Wit 2001	0.65	0.56	0.75	0.074525	Odds Ratio	Yes
Dennehy 2006	1.5	1	2.3	0.212477	Odds Ratio	Yes
Doorduyn 2012	1.52	0.91	2.55	0.264194	Relative Risk	Yes
Doré 2004	0.6	0.3	0.9	0.280258	Odds Ratio	Yes
Duggirala 2005	5.26	1.67	16.67	0.586935	Odds Ratio	Yes
Ethelberg 2006	1.39	1.11	1.81	0.124736	Odds Ratio	Yes
Etiler 2004	1.89	1.2	2.96	0.230323	Relative Risk	No
Evans 2006	0.97	0.9	1.05	0.039324	Odds Ratio	Yes
Faustini 2006	0.26	0.08	0.91	0.62026	Odds Ratio	Yes
Faustini 2006	0.37	0.05	2.5	0.997965	Odds Ratio	No
Fein 1995	0.52	0.43	0.63	0.095954	Odds Ratio	No
Friedman 2004	0.56	0.43	0.71	0.127929	Odds Ratio	No
Fullerton 2007	4.69	3.17	6.93	0.199784	Odds Ratio	No
Gillespie 2008	0.94	0.9	0.99	0.024314	Relative Risk	No
Gillespie 2010	1.38	1.16	1.65	0.089887	Rate Ratio	No
Green 2006	0.67	0.57	0.79	0.083264	Rate Ratio	Yes
Hall 2006	1.11	0.45	2.5	0.437449	Odds Ratio	Yes
Hu 2009	1.15	1.02	1.3	0.0608	Relative Risk	Yes
Hughes 2015	0.82	0.76	0.88	0.038597	Rate Ratio	Yes
Jackson 2015	4	3.6	4.6	0.062531	Rate Ratio	Yes
Jalava 2011	0.58	0.41	0.81	0.171663	Relative Risk	Yes
Jones 2007	0.8	0.6	1	0.130313	Relative Risk	No
Kass 1992	1.25	0.49	3.17	0.475587	Odds Ratio	No
Kum-Nji 2009	0.84	0.46	1.54	0.308243	Relative Risk	Yes
Lake 2007	0.83	0.79	0.88	0.027523	Odds Ratio	Yes

Study	Estimate	Lower CI	Upper CI	SE	Measure	Adjusted
Lal 2012	0.67	0.66	0.68	0.007616	Rate Ratio	No
Ludvigsson 2006	1.93	1.21	3.06	0.236682	Odds Ratio	No
Majowicz 2007	0.83	0.5	1.37	0.257132	Odds Ratio	Yes
McPherson 2009	0.78	0.31	1.95	0.469301	Odds Ratio	No
Moorin 2010	1.61	1.54	1.67	0.020556	Odds Ratio	No
Neal 1997	1.07	0.33	3.44	0.595801	Odds Ratio	No
Odoi 2004	1.61	1.09	2.38	0.200025	Rate Ratio	Yes
Olowokure 1999	1.82	1.71	1.94	0.03182	Rate Ratio	No
Olowokure 1999	2.25	2.15	2.35	0.022107	Rate Ratio	No
Özkan 2007	3.54	1.24	10.16	0.536568	Odds Ratio	Yes
Özmert 2008	2.7	1.2	5.9	0.406283	Odds Ratio	Yes
Phillips 2011	2.3	1.4	3.9	0.261353	Odds Ratio	Yes
Pollard 2014	0.78	0.63	0.95	0.104781	Odds Ratio	Yes
Pyra 2012	0.66	0.56	0.77	0.079617	Rate Ratio	Yes
Rodrigues 2001	0.9	0.5	1.6	0.296722	Odds Ratio	No
Sargeant 2008	1.27	0.78	2.05	0.245748	Odds Ratio	No
Satterthwaite 1999	2	0.93	4.35	0.393558	Odds Ratio	Yes
Scallan 2004	0.77	0.61	0.97	0.118326	Odds Ratio	Yes
Seo 2013	0.39	0.17	0.93	0.435138	Odds Ratio	No
Sethi 2001	1.2	0.71	2.04	0.269245	Odds Ratio	No
Simonsen 2008	0.92	0.9	0.94	0.012321	Rate Ratio	Yes
Simonsen 2008	1.03	1	1.07	0.017628	Rate Ratio	Yes
Spencer 2012	0.75	0.64	0.89	0.083929	Relative Risk	Yes
Stafford 1996	1.09	1	1.19	0.045284	Rate Ratio	No
Tam 2013	0.82	0.64	1.03	0.121389	Hazard Ratio	Yes
Teschke 2010	1.18	1.03	1.34	0.06712	Odds Ratio	No
Turkish Ministry of Health 1995	1.82	1.4	2.3601	0.133861	Odds Ratio	No
Unicomb 2008	1.67	1.25	2	0.119899	Odds Ratio	No
Van Cauteren 2012	0.45	0.29	0.71	0.228414	Odds Ratio	No
Varga 2013	0.98	0.81	1.18	0.096343	Rate Ratio	Yes
Weisent 2012	1.42	1.15	1.75	0.10542	Rate Ratio	Yes
Whitney 2015	0.25	0.14	0.46	0.306296	Rate Ratio	No
Whitney 2015	1.09	0.92	1.27	0.081684	Rate Ratio	No
Wilking 2012	1.57	0.89	2.75	0.286373	Relative Risk	Yes
Wilking 2012	1.62	1.36	1.93	0.088512	Relative Risk	Yes
Wilking 2013	1.15	0.92	1.44	0.114292	Odds Ratio	Yes
Xu 2015	0.86	0.69	1.08	0.113668	Relative Risk	Yes
Younus 2007	0.71	0.65	1	0.109894	Rate Ratio	Yes
Younus 2010	0.86	0.38	1.94	0.414516	Odds Ratio	No
Zappe Pasturel 2013	1.38	1.21	1.57	0.066443	Rate Ratio	Yes

CI - Confidence Interval; SE - Standard Error

Appendix 2.9: PRISMA 2009 checklist

Source: Moher et al. (2009)

Section/topic	#	Checklist item	Reported on page #
		TITLE	
Title	1	Identify the report as a systematic review, meta-analysis, or both.	78
		ABSTRACT	
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	79
		INTRODUCTION	
Rationale	3	Describe the rationale for the review in the context of what is already known.	81
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	81
		METHODS	
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	83
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	83
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	83
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	83
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	84
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	84
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	84
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	85
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	85

Section/topic	#	Checklist item	Reported on page #
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	85
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	85
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	86
		RESULTS	
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	87
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	89
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	92
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	92
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	92
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	98
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; see Item 16).	100

		DISCUSSION	
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	102
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	102
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	102
		FUNDING	
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	ii

Appendix 3 (Supplementary material pertaining to Chapter 5 – Study 2)

Appendix 3.1: IID2 Cohort Study baseline questionnaire (Adults)

Source: Tam et al. (2011a)

ProsStu_Cohort_Base Questionnaire_Adult_06

iid2 The Second Infectious Intestinal Disease Study

Baseline Questionnaire (Weekly Follow-up Study)

The Second Study of Diarrhoea and Vomiting in the Community

Name of Research Nurse:	*For office use only*
Participant's study number:	
Mode of Contact:	

We want to know how often people in the UK suffer from diarrhoea or vomiting and the germs that cause this.

We need to collect some basic information before you take part in the Study.
The information that you give us will be treated in strict confidence

1. What is your surname: .....
   forename(s): .....
2. What is your date of birth (dd/mm/yyyy)? ____/____/____
3. Are you? Male ☐ Female ☐
4. Please give your address: .....
   .....
   .....
5. What is your postcode? ☐☐☐☐ ☐☐☐
6. What is your email address? .....
7. Which ethnic group do you belong to?

		Please tick one
White	British or Irish	
	Other	
Mixed	White & Black Caribbean	
	White and Black African	
	White and Asian	
	Other Mixed	
Asian or Asian British	Indian	
	Pakistani	
	Bangladeshi	
	Other Asian	
Black or Black British	Black Caribbean	
	Black African	
	Other Black	
Another Group	Chinese	
	Other ethnic group	

PLEASE TURN OVER

19th December 2007

8. Please tell us what the job title is of the **main earner** in your household:.....

9. Please tick one box to show which **best** describes the sort of work the **main earner** in your household does. (If the main earner is not working now, please tick a box to show what they did in their last job).

Please tick one box.

Modern professional occupations *such as:* teacher - nurse - physiotherapist - social worker - welfare officer - artist - musician - police officer (sergeant or above) - software designer	
Clerical and intermediate occupations *such as:* secretary - personal assistant - clerical worker - office clerk - call centre agent - nursing auxiliary - nursery nurse	
Senior managers or administrators (usually responsible for planning, organising and co-ordinating work and for finance) *such as:* finance manager - chief executive	
Technical and craft occupations *such as:* motor mechanic - fitter - inspector - plumber - printer - tool maker - electrician - gardener - train driver	
Semi-routine manual and service occupations *such as:* postal worker - machine operative - security guard - caretaker - farm worker - catering assistant - receptionist - sales assistant	
Routine manual and service occupations *such as:* HGV driver - van driver - cleaner - porter - packer - sewing machinist - messenger - labourer - waiter / waitress - bar staff	
Middle or junior managers *such as:* office manager - retail manager - bank manager, restaurant manager - warehouse manager - publican	
Traditional professional occupations *such as:* accountant - solicitor - medical practitioner - scientist - civil / mechanical engineer	

10. Last week, was the **main earner** in your home any of the following?

Please tick one box.

Retired	
Student	
Looking after home/family	
Currently sick/disabled	
None of the above	

11. Does (did) the **main earner** work as an employee or are (were) they self-employed?

Please tick one box.

Employee	
Self-employed with employees	
Self-employed/freelance without employees (please skip questions 12 and 13)	

12. **For employees**: indicate below how many people work (worked) for the **main earner's** employer at the place where they work (worked).
For self-employed: indicate below how many people the main earner employs (employed).

Please tick one box.

1 to 24	
25 or more	

13. Does (did) the **main earner** supervise any other employees?
A supervisor or foreman is responsible for overseeing the work of other employees on a day-to-day basis

Please tick one box.

Yes	
No	

14. The nurse may need to contact you at some time during the study.
What is the best telephone number to contact you on?

.....

15. Do you have a landline at your home? Yes ☐ No ☐

Thank you for agreeing to fill in this questionnaire.

Appendix 3.2: IID2 Cohort Study baseline questionnaire (Children)

Source: Tam et al. (2011a)

ProsStu_Cohort_Base Questionnaire_Child_06

iid2 The Second Infectious Intestinal Disease Study

Baseline Questionnaire (Weekly Follow-up Study)

The Second Study of Diarrhoea and Vomiting in the Community

Name of Research Nurse:	*For office use only*
Participant's Study Number:	
Mode of Contact:	

We want to know how often people in the UK suffer from diarrhoea or vomiting and the germs that cause this.

We need to collect some basic information about your child before they take part in the Study.

The information that you give us will be treated in strict confidence

1. What is your child's surname: .....

   child's forename(s): .....

2. What is your child's date of birth (dd/mm/yyyy)? ____/____/____

3. Are they? Male ☐ Female ☐

4. Please give your child's address: .....

   .....

   .....

5. What is your child's postcode? ☐☐☐☐ ☐☐☐

6. What is your email address? .....

7. Which ethnic group does your child belong to? Please tick one box.

White	British or Irish	
	Other	
Mixed	White & Black Caribbean	
	White and Black African	
	White and Asian	
	Other Mixed	
Asian or Asian British	Indian	
	Pakistani	
	Bangladeshi	
	Other Asian	
Black or Black British	Black Caribbean	
	Black African	
	Other Black	
Another Group	Chinese	
	Other ethnic group	

19th December 2007 PLEASE TURN OVER

8. Please tell us what the job title is of the **main earner** in your child's household:.....

9. Please tick one box to show which **best** describes the sort of work the **main earner** in your child's household does. (If the main earner is not working now, please tick a box to show what they did in their last job).

Please tick one box.

Modern professional occupations *such as:* teacher - nurse - physiotherapist - social worker - welfare officer - artist - musician - police officer (sergeant or above) - software designer	
Clerical and intermediate occupations *such as:* secretary - personal assistant - clerical worker - office clerk - call centre agent - nursing auxiliary - nursery nurse	
Senior managers or administrators (usually responsible for planning, organising and co-ordinating work and for finance) *such as:* finance manager - chief executive	
Technical and craft occupations *such as:* motor mechanic - fitter - inspector - plumber - printer - tool maker - electrician - gardener - train driver	
Semi-routine manual and service occupations *such as:* postal worker - machine operative - security guard - caretaker - farm worker - catering assistant - receptionist - sales assistant	
Routine manual and service occupations *such as:* HGV driver - van driver - cleaner - porter - packer - sewing machinist - messenger - labourer - waiter / waitress - bar staff	
Middle or junior managers *such as:* office manager - retail manager - bank manager, restaurant manager - warehouse manager - publican	
Traditional professional occupations *such as:* accountant - solicitor - medical practitioner - scientist - civil / mechanical engineer	

10. Last week, was the **main earner** in your home any of the following?

Please tick one box.

Retired	
Student	
Looking after home/family	
Currently sick/disabled	
None of the above	

11. Does (did) the **main earner** work as an employee or are (were) they self-employed?

Please tick one box.

Employee	
Self-employed with employees	
Self-employed/freelance without employees (please skip questions 12 and 13)	

12. **For employees**: indicate below how many people work (worked) for the **main earner's** employer at the place where they work (worked).
For self-employed: indicate below how many people the main earner employs (employed).

Please tick one box.

1 to 24	
25 or more	

13. Does (did) the **main earner** supervise any other employees?
A supervisor or foreman is responsible for overseeing the work of other employees on a day-to-day basis

Please tick one box.

Yes	
No	

14. The nurse may need to contact you at some time during the study. What is the best telephone number to contact you on?

.....

15. Do you have a landline at your home? Yes ☐ No ☐

Thank you for agreeing to fill in this questionnaire.

Appendix 3.3: IID2 Cohort Study case questionnaire (Adults)

Source: Tam et al. (2011a)

ProsStu_Cohort Study_Questionnaire_Adult__09

iid2 The Second Infectious Intestinal Disease Study

Questionnaire (Weekly Follow-up Study)

The Second Study of Diarrhoea and Vomiting in the Community

Name of Research Nurse: *For Official Use Only*

Participant's Study Number:

We want to know how often people in the UK suffer from diarrhoea or vomiting and the germs that cause this.

Thank you for agreeing to fill in this questionnaire.

Please read each question carefully before you answer it, and try to answer every question. Please either tick the appropriate box or write your answer in the space provided.

The information that you give us will be treated in strict confidence.

Part 1: This section asks about your age and sex

Please tell us:

1.1 Today's date (dd/mm/yyyy): _____/_____/_____

1.2 Your date of birth (dd/mm/yyyy): _____/_____/_____

1.3 Your sex: Male ☐ Female ☐

Part 2: This section asks about the symptoms you had during your recent illness

2.1 Do you have any of the following symptoms? For EACH symptom please tick Yes, No or Not sure.

Diarrhoea: (loose watery bowel movements)

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Diarrhoea with blood in it:

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Nausea (feeling sick):

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Vomiting (being sick):

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Abdominal cramps (colic):

Yes ☐ No ☐ Not sure ☐

Loss of appetite:

Yes ☐ No ☐ Not sure ☐

Fever (high temperature):

Yes ☐ No ☐ Not sure ☐

Cough or runny/blocked nose or sore throat:

Yes ☐ No ☐ Not sure ☐

Headache:

Yes ☐ No ☐ Not sure ☐

2.2 What was the date (dd/mm/yyyy) on which you first had diarrhoea and/or vomiting?

_____/_____/_____

2.3 If you answered "yes" to having diarrhoea, roughly how many times did you go to the toilet on the worst day (24 hours) of your illness?

Number of times ☐☐

2.4 If you answered "yes" to vomiting, roughly how many times did you vomit on the worst day (24 hours) of your illness?

Number of times ☐☐

2.5 Have you phoned NHS Direct/NHS 24 about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first phone NHS Direct/NHS 24 about these symptoms?

_____/_____/_____

2.6 Have you contacted the out-of-hours doctor service about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first contact the out-of-hours doctor service about these symptoms?

_____/_____/_____

2.7 Have you visited a Walk-in centre about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first contact the walk-in-centre about these symptoms?

_____/_____/_____

2.8 Have you spoken to your nurse or doctor on the 'phone for advice about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first phone for advice about these symptoms?

_____/_____/_____

2.9 Have you been to see a doctor or nurse in your practice about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first see your doctor about these symptoms?

_____/_____/_____

2.10 Did you go to hospital, Accident and Emergency (A&E) or casualty with this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you go to hospital, Accident and Emergency (A&E) or casualty about these symptoms?

_____/_____/_____

PLEASE TURN OVER

2.11 Were you admitted to hospital overnight or longer with this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) were you admitted to hospital with this illness?

_____/____/_____

If "yes", how many nights did you spend in hospital with this illness?

☐☐

2.12 Did your illness stop you from going to work or to school or carrying out your daily activities?

Yes ☐ No ☐ Not sure ☐

If "yes", how many days? ☐☐

Part 3: This section asks about your travel in the ten days before you became ill.

3.1 Did you travel outside the UK in the ten days before you became ill?

Yes ☐ No ☐ Not sure ☐

3.2 If you answered "yes", what dates (dd/mm/yy) were you away?

From: ____/____/____ To: _____/____/_____

3.3 If you were abroad, please tell us which country or countries you visited:

_____

Have you sent a faeces (stool) specimen?

Yes ☐ No ☐

If no, please do so as soon as possible, as this is really important for the study. You can get another specimen pot from your practice nurse if you do not have one.

Thank you for taking the time to fill in this questionnaire.
Please return this questionnaire to the research nurse at your doctor's surgery using the reply paid envelope

Appendix 3.4: IID2 Cohort Study case questionnaire (Children)

Source: Tam et al. (2011a)

ProsStu_Cohort Study_Questionnaire_Child__09

iid2 The Second Infectious Intestinal Disease Study

Questionnaire (Weekly Follow-up Study)

The Second Study of Diarrhoea and Vomiting in the Community

The "Tummy Bug" Study

Name of Research Nurse:	*For Official Use Only*
Participant's Study Number:	

We want to know how often people in the UK suffer from diarrhoea or vomiting and the germs that cause this.

Thank you for agreeing to fill in this questionnaire.

Please read each question carefully before you answer it, and try to answer every question. Please either tick the appropriate box or write your answer in the space provided.

The information that you give us will be treated in strict confidence.

Part 1: This section asks about your child's age and sex

Please tell us:

1.1 Today's date (dd/mm/yyyy): _____/_____/_____

1.2 Child's date of birth (dd/mm/yyyy): _____/_____/_____

1.3 Child's sex: Male ☐ Female ☐

Part 2: This section asks about the symptoms your child had during your recent illness

2.1 Did they have any of the following symptoms? For EACH symptom please tick Yes, No or Not sure.

Diarrhoea: (loose watery bowel movements)

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Diarrhoea with blood in it:

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

17th January 2008 PLEASE TURN OVER

Nausea (feeling sick):

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Vomiting (being sick):

Yes ☐ No ☐ Not sure ☐

Number of days: ☐☐

Still Present: Yes ☐ No ☐ Not sure ☐

Abdominal cramps (colic):

Yes ☐ No ☐ Not sure ☐

Loss of appetite:

Yes ☐ No ☐ Not sure ☐

Fever (high temperature):

Yes ☐ No ☐ Not sure ☐

Cough or runny/blocked nose or sore throat:

Yes ☐ No ☐ Not sure ☐

Headache:

Yes ☐ No ☐ Not sure ☐

2.2 What was the date (dd/mm/yyyy) on which your child first had diarrhoea and/or vomiting?

_____/_____/_____

2.3 If you answered "yes" to having diarrhoea, roughly how many times did your child go to the toilet on the worst day (24 hours) of their illness?

Number of times ☐☐

2.4 If you answered "yes" to having vomiting, roughly how many times did your child go to the toilet on the worst day (24 hours) of their illness?

Number of times ☐☐

2.5 Have you phoned NHS Direct/NHS 24 about your child's illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first phone NHS Direct/NHS 24 about your child's symptoms?

_____/_____/_____

2.6 Have you contacted the out-of-hours doctor service about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first contact the out-of-hours doctor service about your child's symptoms?

_____/_____/_____

2.7 Have you visited a Walk-in centre about your child's illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first contact the walk-in-centre about your child's symptoms?

_____/_____/_____

2.8 Have you spoken to your child's nurse or doctor on the 'phone for advice about their illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you first phone for advice about these symptoms?

_____/_____/_____

2.9 Have you been to your child's doctor or nurse in your practice about this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did your child first see their doctor about these symptoms?

_____/_____/_____

2.10 Did you take your child to hospital, Accident and Emergency (A&E) or casualty with this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) did you go to hospital, Accident and Emergency (A&E) or casualty about these symptoms?

_____/_____/_____

PLEASE TURN OVER

2.11 Was your child admitted to hospital overnight or longer with this illness?

Yes ☐ No ☐ Not sure ☐

If "yes", on what date (dd/mm/yyyy) were you admitted to hospital with this illness?

_____/_____/_____

If "yes", how many nights did you spend in hospital with this illness?

☐☐

2.12 Did your child's illness stop them from going to school or day care?

Yes ☐ No ☐ Not sure ☐

If "yes", how many days? ☐☐

Part 3: This section asks about your travel in the ten days before you became ill.

3.1 Did your child travel outside the UK in the ten days before you became ill?

Yes ☐ No ☐ Not sure ☐

3.2 If you answered "yes", what dates (dd/mm/yy) were you away?

From: _____/_____/_____ To: _____/_____/_____

3.3 If you were abroad, please tell us which country or countries your child visited:

_____

Have you sent a faeces (stool) specimen?

Yes ☐ No ☐

If no, please do so as soon as possible, as this is really important for the study. You can get another specimen pot from your practice nurse if you do not have one.

Thank you for taking the time to fill in this questionnaire.
Please return this questionnaire to the research nurse at your doctor's surgery using the reply paid envelope

Appendix 3.5: *Sensitivity analysis – multiple spells of follow up (n subjects=9332; n failures=1010)*

Variable	Category	Unadjusted		Adjusted[a]		p value
		Hazard ratio	**(95% CI)**	**Hazard ratio**	**(95% CI)**	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.04	(0.88-1.22)	1.03	(0.86-1.22)	0.76
	Routine/manual	0.71	(0.60-0.85)	**0.74**	**(0.61-0.90)**	**0.002**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.12	(0.98-1.28)	1.08	(0.94-1.25)	0.28
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.79	(0.69-0.91)	1.00	(0.83-1.21)	0.98

NS-SEC: National Statistics- Socioeconomic Classification; CI: confidence interval

[a] Adjusted for all other covariates in the model

Baseline hazard stratified by age group and sex

Missing data: NS-SEC was not classifiable for 1,112 individuals. Employment status was missing for 12 records. Rural-urban classification was missing for four records.

Appendix 3.6: IID2 study population compared to UK population (Office for National Statistics [dataset], 2011).

NS-SEC*	**IID2**	**UK population**
Managerial/professional	52%	31%
Intermediate	15%	22%
Routine/manual	17%	33%

*Recoded from five class self-coded version to three-class version (Office for National Statistics, 2010)

Appendix 4: (Supplementary material pertaining to Chapter 6 – Study 3)

Appendix 4.1: Crude rates of calls per 10,000 person-months by system

	1 (Least Disadvantaged)	2	3	4	5 (Most Disadvantaged)
NHS Direct					
GI-calls overall	2.7	2.7	2.8	3.0	2.9
Female	3.0	3.0	3.2	3.4	3.3
Male	2.4	2.4	2.5	2.6	2.5
0-4	17.7	17.5	17.3	16.5	14.1
5-9	2.1	2.1	2.2	2.2	2.0
10-14	0.9	1.0	1.1	1.1	0.9
15-19	2.0	2.1	2.3	2.4	2.3
20-59	1.9	1.9	2.0	2.1	2.1
60+	1.8	1.8	1.9	2.0	2.0
Non GI-calls overall	36.1	36.6	38.5	41.4	43.3
Female	41.3	41.9	44.5	48.2	50.7
Male	30.7	31.0	32.4	34.3	35.7
0-4	146.3	142.7	138.7	129.7	112.4
5-9	26.7	26.5	26.9	26.1	23.7
10-14	17.5	17.5	18.2	17.9	16.7
15-19	34.1	36.0	39.4	43.3	46.0
20-59	32.4	33.5	35.9	39.3	42.7
60+	25.7	26.1	27.3	28.8	29.7
NHS 111					
GI-calls overall	5.4	6.4	7.5	7.8	8.1
Female	6.2	7.4	8.6	9.0	9.3
Male	4.5	5.3	6.3	6.6	6.9
0-4	34.9	42.1	49.0	48.7	46.7
5-9	3.7	4.3	5.2	5.6	5.6
10-14	1.7	2.0	2.3	2.4	2.4
15-19	3.7	4.6	5.6	6.0	6.3
20-59	2.9	3.6	4.2	4.4	4.5
60+	5.4	6.1	6.9	6.7	6.4
Non GI-calls overall	118.3	121.6	129.6	134.1	125.0
Female	133.6	137.8	147.0	153.1	143.0
Male	102.4	104.8	111.4	114.5	106.7
0-4	381.8	384.5	383.7	355.2	289.9
5-9	91.5	91.8	94.8	90.9	76.0
10-14	55.0	55.2	55.8	55.4	49.1
15-19	104.5	110.9	123.7	133.6	131.2
20-59	91.5	95.9	105.4	113.6	113.3
60+	137.7	139.8	147.2	152.4	133.9

Appendix 4.2: Characteristics of callers (n= 24,214,879)

		Q1 (Least Disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (Most Disadvantaged) n (%)
Total (all calls)		1,582,603 (6.5)	5,157,046 (21.3)	6,193,779 (25.6)	7,851,831 (32.4)	3,429,620 (14.2)
Call category	GI-calls	83,410 (5.3)	294,772 (5.7)	374,595 (6.0)	481,341 (6.1)	216,725 (6.3)
	Non GI-calls	1,499,193 (94.7)	4,862,274 (94.3)	5,819,184 (94.0)	7,370,490 (93.9)	3,212,895 (93.7)
Age Group	0-4	316,842 (20.4)	976,409 (19.3)	1,134,535 (18.6)	1,449,207 (18.8)	633,286 (18.8)
	5-9	71,038 (4.6)	208,061 (4.1)	238,879 (3.9)	298,289 (3.9)	130,220 (3.9)
	10-14	46,874 (3.0)	137,961 (2.7)	155,231 (2.6)	189,834 (2.5)	82,547 (2.5)
	15-19	83,971 (5.4)	287,795 (5.7)	364,344 (6.0)	494,412 (6.4)	241,483 (7.2)
	20-59	637,843 (41.0)	2,150,324 (42.5)	2,703,485 (44.4)	3,715,283 (48.2)	1,751,116 (52.1)
	60+	397,873 (25.6)	1,303,990 (25.7)	1,489,286 (24.5)	1,568,590 (20.3)	524,632 (15.6)
Sex	Female	903,786 (57.9)	2,949,741 (58.0)	3,555,030 (58.1)	4,514,634 (58.2)	1,961,619 (58.0)
	Male	658,179 (42.1)	2,139,443 (42.0)	2,560,104 (41.9)	3,238,100 (41.8)	1,419,261 (42.0)
Urban decile (%)	<10	379,964 (24.0)	1,040,760 (20.2)	649,928 (10.5)	254,580 (3.2)	21,784 (0.6)
	10-19	98,951 (6.3)	242,213 (4.7)	132,824 (2.1)	53,224 (0.7)	12,864 (0.4)
	20-29	46,181 (2.9)	236,478 (4.6)	204,197 (3.3)	81,038 (1.0)	0 (0.0)
	30-39	8,956 (0.6)	293,384 (5.7)	178,433 (2.9)	118,953 (1.5)	15,650 (0.5)
	40-49	79,828 (5.0)	304,555 (5.9)	198,914 (3.2)	176,707 (2.3)	40,342 (1.2)
	50-59	97,216 (6.1)	198,534 (3.9)	359,697 (5.8)	114,244 (1.5)	19,266 (0.6)
	60-69	75,426 (4.8)	336,418 (6.5)	325,678 (5.3)	167,293 (2.1)	6,720 (0.2)
	70-79	49,940 (3.2)	194,519 (3.8)	323,195 (5.2)	390,363 (5.0)	97,025 (2.8)
	80-89	107,092 (6.8)	359,769 (7.0)	487,752 (7.9)	370,970 (4.7)	25,158 (0.7)
	90-100	639,049 (40.4)	1,950,416 (37.8)	3,333,161 (53.8)	6,124,459 (78.0)	3,190,811 (93.0)
Source	NHS Direct	510,962 (32.3)	1,652,336 (32.0)	1,963,470 (31.7)	2,559,152 (32.6)	1,188,337 (34.6)
	NHS 111	1,071,641 (67.7)	3,504,710 (68.0)	4,230,309 (68.3)	5,292,679 (67.4)	2,241,283 (65.4)
Rotavirus vaccine	Pre-	506,861 (32.0)	1,638,506 (31.8)	1,947,615 (31.4)	2,539,415 (32.3)	1,180,305 (34.4)
	Post-	1,075,742 (68.0)	3,518,540 (68.2)	4,246,164 (68.6)	5,312,416 (67.7)	2,249,315 (65.6)

Missing data: age - 431,239; sex - 314,982; diarrhoea - 2

Appendix 4.3: Characteristics of callers – GI-calls (n=1,450,843)

		Q1 (Least Disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (Most Disadvantaged) n (%)
Total (all calls)		83,410 (5.7)	294,772 (20.3)	374,595 (25.8)	481,341 (33.2)	216,725 (14.9)
Age Group	0-4	31,077 (37.3)	107,209 (36.4)	137,569 (36.7)	186,504 (38.8)	88,349 (40.8)
	5-9	3,564 (4.3)	11,325 (3.8)	14,584 (3.9)	19,590 (4.1)	9,550 (4.4)
	10-14	1,719 (2.1)	5,744 (2.0)	7,043 (1.9)	8,972 (1.9)	4,109 (1.9)
	15-19	3,523 (4.2)	13,353 (4.5)	17,891 (4.8)	23,899 (5.0)	11,735 (5.4)
	20-59	25,501 (30.6)	93,775 (31.8)	120,765 (32.3)	163,269 (34.0)	75,755 (35.0)
	60+	17,982 (4.3)	63,199 (3.8)	76,528 (3.9)	78,470 (4.1)	27,092 (4.4)
Sex	Female	48,260 (57.9)	171,212 (58.1)	218,216 (58.3)	279,450 (58.1)	125,050 (57.7)
	Male	35,125 (42.1)	123,471 (41.9)	156,286 (41.7)	201,740 (41.9)	91,586 (42.3)
Urban decile (%)	<10	17,791 (21.3)	55,961 (19.0)	36,874 (9.8)	14,044 (2.9)	1,096 (0.5)
	10-19	4,892 (5.9)	14,074 (4.8)	7,678 (2.1)	2,564(0.5)	739 (0.3)
	20-29	2,378 (2.9)	12,885 (4.4)	12,333 (3.3)	4,979 (1.0)	0 (0.0)
	30-39	364 (0.4)	16,915 (5.7)	10,474 (2.8)	7,351 (1.5)	924 (0.4)
	40-49	4,508 (5.4)	17,350 (5.9)	11,953 (3.2)	10,163 (2.1)	1,604 (0.7)
	50-59	4,537 (5.4)	10,948 (3.7)	20,946 (5.6)	6,655 (1.4)	1,186 (0.5)
	60-69	4,759 (5.7)	17,147 (5.8)	18,594 (5.0)	9,598 (2.0)	431 (1.0)
	70-79	3,160 (3.8)	11,963 (4.1)	18,058 (4.8)	20,585 (4.3)	6,126 (2.8)
	80-89	6,384 (7.7)	20,561 (7.0)	29,443 (7.9)	21,065 (4.4)	1,391 (0.6)
	90-100	34,637 (41.5)	116,968 (39.7)	208,242 (55.6)	384,337 (79.8)	203,228 (93.8)
Source	NHS Direct	34,396 (41.2)	109,773 (37.2)	129,624 (34.6)	167,240 (34.7)	72,330 (33.4)
	NHS 111	49,014 (58.8)	184,999 (62.8)	244,971 (65.4)	314,101 (65.3)	144,395 (66.6)
Rotavirus vaccine	Pre-	34,219 (41.0)	109,082 (37.0)	128,794 (34.4)	166,315 (34.6)	71,977 (33.2)
	Post-	49,191 (59.0)	185,690 (63.0)	245,801 (65.6)	315,026 (65.4)	144,748 (66.8)

Missing data: age - 838; sex - 447; diarrhoea - 2

Appendix 4.4: Characteristics of callers – non GI-calls (n=22,764,036)

		Q1 (Least Disadvantaged) n (%)	Q2 n (%)	Q3 n (%)	Q4 n (%)	Q5 (Most Disadvantaged) n (%)
Total (all calls)		1,499,193 (6.6)	4,862,274 (21.4)	5,819,184 (25.6)	7,370,490 (32.4)	3,212,895 (14.1)
Age Group	0-4	285,265 (19.4)	869,200 (18.2)	996,966 (17.5)	1,262,703 (17.5)	544,937 (17.3)
	5-9	67,474 (4.6)	197,736 (4.1)	224,295 (3.9)	278,699 (3.9)	120,670 (3.8)
	10-14	45,155 (3.1)	132,217 (2.8)	148,188 (2.6)	180,862 (2.5)	78,438 (2.5)
	15-19	80,448 (5.5)	274,442 (5.8)	346,453 (6.1)	470,513 (6.5)	229,748 (7.3)
	20-59	612,342 (41.6)	2,056,549 (43.1)	2,582,720 (45.2)	3,551,654 (49.1)	1,675,361 (53.2)
	60+	379,891 (7.6)	1,240,791 (26.0)	1,412,758 (24.7)	1,490,120 (20.6)	497,540 (15.8)
Sex	Female	855,526 (57.9)	2,778,529 (58.0)	3,336,814 (58.1)	4,235,184 (58.2)	1,836,569 (58.0)
	Male	623,054 (42.1)	2,015,972 (42.0)	2,403,818 (41.9)	3,036,360 (41.8)	1,327,675 (42.0)
Urban decile (%)	<10	362,173 (24.2)	984,799 (20.3)	613,054 (10.5)	240,536 (3.3)	20,688 (0.6)
	10-19	94,059 (6.3)	228,139 (4.7)	125,146 (2.2)	50,660 (0.7)	12,125 (0.4)
	20-29	43,803 (2.9)	223,593 (4.6)	191,864 (3.3)	76,059 (1.0)	0 (0.0)
	30-39	8,592 (0.6)	276,469 (5.7)	167,959 (2.9)	111,602 (1.5)	14,726 (0.5)
	40-49	75,320 (5.0)	287,205 (5.9)	186,961 (3.2)	166,544 (2.3)	38,738 (1.2)
	50-59	92,679 (6.2)	187,586 (3.9)	338,751 (5.8)	107,589 (1.5)	18,080 (0.6)
	60-69	70,667 (4.7)	319,271 (6.6)	307,084 (5.3)	157,695 (2.1)	6,289 (0.2)
	70-79	46,780 (3.1)	182,556 (3.8)	305,137 (5.2)	369,778 (5.0)	90,899 (2.8)
	80-89	100,708 (6.7)	339,208 (7.0)	458,309 (7.9)	349,905 (4.7)	23,767 (0.7)
	90-100	604,412 (40.3)	1,83,448 (37.7)	3,124,919 (53.7)	5,740,122 (77.9)	2,987,583 (93.0)
Source	NHS Direct	476,566 (31.8)	1,542,563 (31.7)	1,833,846 (31.5)	2,391,912 (32.5)	1,116,007 (34.7)
	NHS 111	1,022,627 (68.2)	3,319,711 (68.3)	3,985,338 (68.5)	4,978,578 (67.5)	2,096,888 (65.3)
Rotavirus vaccine	Pre-	472,642 (31.5)	1,529,424 (31.5)	1,818,821 (31.3)	2,373,100 (32.2)	1,108,328 (34.5)
	Post-	1,026,551 (68.5)	3,332,850 (68.5)	4,000,363 (68.7)	4,997,390 (67.8)	2,104,567 (65.5)

Missing data: age - 430,401; sex - 314,535

Appendix 4.5: Univariate and multivariable regression analysis GI calls full model – NHS Direct (n=24,985)

Variable	Category	Univariate	Multivariable[a]	p value
		RR (95% CI)	RR (95% CI)	
IMD Quintile	1 (least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.00 (0.99-1.01)	1.02 (1.00-1.04)	0.05
	3	1.04 (1.03-1.05)	1.02 (1.00-1.04)	0.10
	4	1.10 (1.09-1.11)	1.04 (1.02-1.06)	<0.001
	5 (most disadvantaged)	1.07 (1.05-1.08)	1.01 (0.99-1.03)	1.00
Age Group	0-4	8.17 (8.12-8.22)	1.25 (1.22-1.29)	<0.001
	5-9	1.06 (1.06-1.07)	1.47 (1.40-1.56)	<0.001
	10-14	0.51 (0.50-0.52)	0.63 (0.58-0.68)	<0.001
	15-19	1.12 (1.11-1.14)	0.92 (0.87-0.97)	<0.01
	20-59	1.00 (reference)	1.00 (reference)	
	60+	0.95 (0.94-0.96)	1.02 (0.99-1.05)	1.00
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.28 (1.27-1.28)	1.63 (1.62-1.65)	<0.001
Urban decile (%)	<10	0.72 (0.71-0.72)	0.76 (0.75-0.76)	<0.001
	10-19	0.77 (0.76-0.79)	0.80 (0.78-0.81)	<0.001
	20-29	0.82 (0.81-0.84)	0.85 (0.83-0.86)	<0.001
	30-39	0.83 (0.82-0.85)	0.85 (0.83-0.87)	<0.001
	40-49	0.87 (0.85-0.88)	0.90 (0.88-0.91)	<0.001
	50-59	0.86 (0.84-0.87)	0.87 (0.85-0.88)	<0.001
	60-69	0.91 (0.90-0.93)	0.92 (0.91-0.93)	<0.001
	70-79	0.97 (0.95-0.98)	0.98 (0.97-1.00)	0.05
	80-89	0.95 (0.93-0.96)	0.95 (0.94-0.96)	<0.001
	90-100	1.00 (reference)	1.00 (reference)	
Age Group * Sex	0-4: Female	-	0.60 (0.59-0.61)	<0.001
	5-9: Female	-	0.62 (0.60-0.64)	<0.001
	10-14: Female	-	0.65 (0.63-0.68)	<0.001
	15-19: Female	-	1.26 (1.23-1.30)	<0.001
	60+: Female	-	0.90 (0.89-0.92)	<0.001
Age Group * IMD Quintile	0-4: 2	-	0.96 (0.93-0.99)	<0.001
	0-4: 3	-	0.92 (0.89-0.94)	<0.001
	0-4: 4	-	0.82 (0.80-0.84)	<0.001
	0-4: 5	-	0.72 (0.69-0.74)	<0.001
	5-9: 2	-	0.96 (0.91-1.02)	1.00
	5-9: 3	-	1.00 (0.94-1.06)	1.00
	5-9: 4	-	0.92 (0.87-0.98)	<0.01
	5-9: 5	-	0.84 (0.79-0.90)	<0.001
	10-14: 2	-	1.06 (0.98-1.16)	1.00
	10-14: 3	-	1.09 (1.01-1.19)	0.05
	10-14: 4	-	1.03 (0.95-1.12)	1.00
	10-14: 5	-	0.92 (0.84-1.00)	0.10
	15-19: 2	-	1.04 (0.98-1.11)	1.00
	15-19: 3	-	1.10 (1.04-1.17)	<0.01
	15-19: 4	-	1.07 (1.01-1.14)	0.05
	15-19: 5	-	1.04 (0.98-1.11)	1.00
	60+: 2	-	0.97 (0.94-1.01)	1.00
	60+: 3	-	1.01 (0.98-1.05)	1.00
	60+: 4	-	0.98 (0.95-1.02)	1.00
	60+: 5	-	0.98 (0.95-1.02)	1.00

RR – Risk Ratio; [a] Adjusted for all other covariates in the model

Appendix 4.6: Univariate and multivariable regression analysis GI calls full model – NHS 111 (n=24,985)

Variable	Category	Univariate	Multivariable[a]	p value
		RR (95% CI)	RR (95% CI)	
IMD Quintile	1 (least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.19 (1.18-1.20)	1.23 (1.21-1.26)	<0.001
	3	1.39 (1.38-1.41)	1.42 (1.40-1.45)	<0.001
	4	1.46 (1.45-1.47)	1.46 (1.44-1.49)	<0.001
	5 (most disadvantaged)	1.50 (1.49-1.52)	1.50 (1.47-1.53)	<0.001
Age Group	0-4	11.31 (11.26-11.36)	18.54 (18.12-18.97)	<0.001
	5-9	1.24 (1.23-1.25)	1.90 (1.81-2.00)	<0.001
	10-14	0.54 (0.54-0.55)	0.82 (0.77-0.88)	<0.001
	15-19	1.34 (1.33-1.35)	1.07 (1.02-1.12)	<0.01
	20-59	1.00 (reference)	1.00 (reference)	
	60+	1.58 (1.57-1.59)	2.05 (2.00-2.11)	<0.01
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.37 (1.36-1.38)	2.00 (1.98-2.01)	<0.001
Urban decile (%)	<10	0.73 (0.73-0.74)	0.84 (0.83-0.84)	<0.001
	10-19	0.84 (0.83-0.85)	0.96 (0.94-0.97)	<0.001
	20-29	0.84 (0.83-0.85)	0.93 (0.91-0.94)	<0.001
	30-39	0.91 (0.90-0.92)	0.97 (0.96-0.99)	<0.001
	40-49	0.93 (0.92-0.94)	1.02 (1.01-1.04)	<0.001
	50-59	0.88 (0.87-0.90)	0.94 (0.93-0.95)	<0.001
	60-69	0.83 (0.82-0.84)	0.89 (0.81-0.90)	<0.001
	70-79	0.95 (0.94-0.95)	0.97 (0.96-0.98)	<0.001
	80-89	0.94 (0.94-0.95)	0.99 (0.98-1.00)	0.05
	90-100	1.00 (reference)	1.00 (reference)	
Age Group * Sex	0-4: Female	-	0.48 (0.47-0.49)	<0.001
	5-9: Female	-	0.51 (0.50-0.53)	<0.001
	10-14: Female	-	0.56 (0.54-0.58)	<0.001
	15-19: Female	-	1.34 (1.31-1.37)	<0.001
	60+: Female	-	0.84 (0.83-0.85)	<0.001
Age Group * IMD Quintile	0-4: 2	-	0.97 (0.94-1.00)	0.05
	0-4: 3	-	0.96 (0.94-0.99)	<0.01
	0-4: 4	-	0.91 (0.89-0.93)	<0.001
	0-4: 5	-	0.85 (0.82-0.87)	<0.001
	5-9: 2	-	0.94 (0.89-0.99)	0.05
	5-9: 3	-	0.97 (0.92-1.02)	1.00
	5-9: 4	-	0.98 (0.93-1.03)	1.00
	5-9: 5	-	0.96 (0.90-1.01)	0.10
	10-14: 2	-	0.96 (0.90-1.03)	1.00
	10-14: 3	-	0.95 (0.89-1.02)	1.00
	10-14: 4	-	0.93 (0.87-1.00)	0.05
	10-14: 5	-	0.90 (0.84-0.98)	<0.01
	15-19: 2	-	1.00 (0.95-1.05)	1.00
	15-19: 3	-	1.04 (0.99-1.09)	1.00
	15-19: 4	-	1.05 (1.00-1.11)	0.05
	15-19: 5	-	1.06 (1.01-1.12)	0.05
	60+: 2	-	0.92 (0.89-0.94)	<0.001
	60+: 3	-	0.88 (0.86-0.91)	<0.001
	60+: 4	-	0.82 (0.79-0.84)	<0.001
	60+: 5	-	0.74 (0.72-0.77)	<0.001

RR – Risk Ratio; [a] Adjusted for all other covariates in the model

Appendix 4.7: Univariate and multivariable regression analysis GI calls – Sensitivity postcode districts with population=0 recoded to population=1 by system

Variable	Category	Univariate	Multivariable[a]	p value
		RR (95% CI)	RR (95% CI)	
		NHS Direct (n=25,008)		
IMD Quintile	1 (least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.00 (0.99-1.01)	1.00 (0.99-1.01)	1.00
	3	1.04 (1.03-1.05)	0.99 (0.98-1.01)	1.00
	4	1.10 (1.09-1.11)	0.97 (0.96-0.98)	<0.001
	5 (most disadvantaged)	1.07 (1.05-1.08)	0.88 (0.87-0.89)	<0.001
Age Group	0-4	8.17 (8.12-8.22)	8.19 (8.14-8.24)	<0.001
	5-9	1.06 (1.06-1.07)	1.06 (1.05-1.08)	<0.001
	10-14	0.51 (0.50-0.52)	0.51 (0.50-0.52)	<0.001
	15-19	1.12 (1.11-1.14)	1.13 (1.11-1.14)	<0.001
	20-59	1.00 (reference)	1.00 (reference)	
	60+	0.95 (0.94-0.96)	0.95 (0.94-0.96)	<0.001
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.28 (1.27-1.28)	1.31 (1.30-1.32)	<0.001
Urban decile (%)	<10	0.72 (0.71-0.72)	0.75 (0.75-0.76)	<0.001
	10-19	0.77 (0.76-0.79)	0.80 (0.78-0.81)	<0.001
	20-29	0.82 (0.81-0.84)	0.85 (0.83-0.86)	<0.001
	30-39	0.83 (0.82-0.85)	0.85 (0.83-0.86)	<0.001
	40-49	0.87 (0.85-0.88)	0.90 (0.88-0.91)	<0.001
	50-59	0.86 (0.84-0.87)	0.87 (0.85-0.88)	<0.001
	60-69	0.91 (0.90-0.93)	0.92 (0.91-0.93)	<0.001
	70-79	0.97 (0.95-0.98)	0.98 (0.97-1.00)	0.05
	80-89	0.95 (0.93-0.96)	0.95 (0.94-0.96)	<0.001
	90-100	1.00 (reference)	1.00 (reference)	
		NHS 111 (n=25,008)		
IMD Quintile	1 (least disadvantaged)	1.00 (reference)	1.00 (reference)	
	2	1.19 (1.18-1.20)	1.19 (1.18-1.20)	<0.001
	3	1.39 (1.38-1.41)	1.36 (1.35-1.37)	<0.001
	4	1.46 (1.45-1.47)	1.35 (1.34-1.36)	<0.001
	5 (most disadvantaged)	1.50 (1.49-1.52)	1.32 (1.31-1.33)	<0.001
Age Group	0-4	11.31 (11.26-11.36)	11.32 (11.34-11.38)	<0.001
	5-9	1.24 (1.23-1.25)	1.25 (1.24-1.26)	<0.001
	10-14	0.54 (0.54-0.55)	0.55 (0.54-0.59)	<0.001
	15-19	1.34 (1.33-1.35)	1.35 (1.33-1.36)	<0.001
	20-59	1.00 (reference)	1.00 (reference)	
	60+	1.58 (1.57-1.59)	1.58 (1.58-1.59)	<0.001
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.37 (1.36-1.38)	1.46 (1.45-1.46)	<0.001
Urban decile (%)	<10	0.73 (0.73-0.74)	0.84 (0.83-0.85)	<0.001
	10-19	0.84 (0.83-0.85)	0.96 (0.94-0.97)	<0.001
	20-29	0.84 (0.83-0.85)	0.92 (0.91-0.94)	<0.001
	30-39	0.91 (0.90-0.92)	0.97 (0.96-0.99)	<0.001
	40-49	0.93 (0.92-0.94)	1.02 (1.01-1.04)	<0.001
	50-59	0.88 (0.87-0.90)	0.94 (0.93-0.95)	<0.001
	60-69	0.83 (0.82-0.84)	0.89 (0.88-0.90)	<0.001
	70-79	0.95 (0.94-0.95)	0.97 (0.96-0.98)	<0.001
	80-89	0.94 (0.94-0.95)	0.99 (0.98-1.00)	0.05
	90-100	1.00 (reference)	1.00 (reference)	

RR – Risk Ratio; [a] Adjusted for all other covariates in the model

Appendix 4.8: Univariate and multivariable regression analysis GI-calls – Sensitivity IMD score and proportion of population classed as urban as continuous variables by system

Variable	Category	Univariate	Multivariable[a]	p value
		RR (95% CI)	**RR (95% CI)**	
		NHS Direct (n=24,985)		
IMD Score		1.00 (1.00-1.00)	1.00 (1.00-1.00)	<0.001
Age Group	0-4	8.17 (9.12-8.22)	8.19 (8.14-8.24)	<0.001
	5-9	1.06 (1.04-1.07)	1.06 (1.05-1.08)	<0.001
	10-14	0.51 (0.50-0.52)	0.51 (0.50-0.52)	<0.001
	15-19	1.12 (1.11-1.14)	1.13 (1.11-1.14)	<0.001
	20-59	1.00 (reference)	1.00 (reference)	
	60+	0.95 (0.94-0.96)	0.95 (0.94-0.96)	<0.001
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.28 (1.27-1.28)	1.31 (1.30-1.32)	<0.001
Urban (%)		1.00 (1.00-1.00)	1.00 (1.00-1.00)	<0.001
		NHS 111 (n=24,985)		
IMD Score		1.01 (1.01-1.01)	1.00 (1.00-1.00)	<0.001
Age Group	0-4	11.31 (11.26 -11.36)	11.30 (11.25-11.36)	<0.001
	5-9	1.24 (1.23-1.25)	1.25 (1.23-1.26)	<0.001
	10-14	0.54 (0.54-0.55)	0.55 (0.54-0.57)	<0.001
	15-19	1.34 (1.33-1.35)	1.35 (1.33-1.36)	<0.001
	20-59	1.00 (reference)	1.00 (reference)	
	60+	1.58 (1.57-1.59)	1.59 (1.58-1.60)	<0.001
Sex	Male	1.00 (reference)	1.00 (reference)	
	Female	1.37 (1.36-1.38)	1.40 (1.39-1.40)	<0.001
Urban (%)		1.00 (1.00-1.00)	1.00 (1.00-1.00)	<0.001

RR – Risk Ratio; [a] Adjusted for all other covariates in the model

Appendix 4.9a: Univariate and multivariable regression analysis GI calls – Sensitivity analysis <1 year olds excluded NHS Direct (n=24,985)

Age group	IMD Quintile	RR[a] (95% CI)	p value
0-4	**1** (Least disadvantaged)	1.00 (reference)	
	2	1.02 (1.00-1.04)	<0.001
	3	1.01 (0.99-1.03)	<0.001
	4	0.97 (0.95-0.98)	<0.001
	5 (Most disadvantaged)	0.83 (0.82-0.85)	<0.001

[a]Adjusted for sex, % urban

Appendix 4.9b: Rates per 10,000 person-months and incidence rate ratio for exposed (most disadvantaged) compared to unexposed (least disadvantaged) in NHS Direct by age group Sensitivity analysis <1 year olds included compared to <1 year olds excluded

	<1s Included	<1s Excluded
GI calls		
Overall rate 0-4	16.5	11.5
0-4 Rate/10,000* in most disadvantaged	14.1	9.8
0-4 Rate/10,000* in least disadvantaged	17.7	12.2
Incidence rate ratio (95% CI)	0.79 (0.78-0.81)	0.80 (0.78-0.82)
Overall rate 5+	2.0	2.0
5+ Rate/10,000* in most disadvantaged	2.0	2.0
5+ Rate/10,000* in least disadvantaged	1.8	1.8
Incidence rate ratio (95% CI)	1.10 (1.08-1.12)	1.10 (1.08-1.12)
Non-GI calls		
Overall rate 0-4	132.7	86.0
0-4 Rate/10,000* in most disadvantaged	112.4	71.2
0-4 Rate/10,000* in least disadvantaged	146.3	96.4
Incidence rate ratio (95% CI)	0.77 (0.76-0.77)	0.74 (0.73-0.75)
Overall rate 5+	33.3	33.3
5+ Rate/10,000* in most disadvantaged	37.6	37.6
5+ Rate/10,000* in least disadvantaged	29.3	29.3
Incidence rate ratio (95% CI)	1.28 (1.28-1.29)	1.28 (1.28-1.29)

*person-months

Appendix 5: (Supplementary material pertaining to Chapter 7 – Study 4)

Appendix 5.1: Vero cytotoxin-producing Escherichia coli Enhanced Surveillance Questionnaire

Public Health England

Vero cytotoxin-producing *Escherichia coli* Enhanced surveillance questionnaire

SECTION A:QUESTIONNAIRE DETAILS

Interviewer name: Interview date : / /

Interviewer office: Interviewer Telephone:

Person interviewed name:

Details from: ☐ Case ☐ Case's Parent ☐ Other (specify):

SECTION B: CASE CLASSIFICATION – SEE GUIDANCE NOTES FOR DEFINITIONS

Is this case: ☐ Primary ☐ Co-primary ☐ Secondary ☐ Not sure ☐ Asymptomatic

If secondary, name of primary case:

Investigation is: ☐ Ongoing ☐ Complete

Outbreak keyword or number:

Outcome:

Select all that apply ☐ Recovered ☐ Still ill ☐ HUS/TTP ☐ Died

if died, then → Date of death: / /

SECTION C:PERSONAL DETAILS

First name: Family name:

Address:

Postcode: Tel (h): Tel (m):

Email:

Sex: ☐ M ☐ F Date of birth (dd/mm/yyyy): / / Age: yrs

NHS No: GP name:

GP address: GP Tel:

Are there any children living in the household? (other than the case) ☐ Y ☐ N ☐ U

Occupation:

Tick if any of the below risk groups apply

☐ Foodhandler (e.g. handle food professionally)
☐ Work in/attend healthcare setting
☐ Work in/attend childcare setting
☐ Work in contact with faeces (e.g. lab, farm etc.)
☐ Have difficulty maintaining personal hygiene
☐ Other risk category

If yes to any of the above, details:

Work/School: Tel:

Address:

Postcode: Date of last attendance: / /

Ethnicity: ☐ White ☐ Mixed ☐ Asian/Asian British ☐ Black/Black British
☐ Chinese Other:

April 2013 v3

SECTION D: SYMPTOMS OF ILLNESS

Onset date: ☐☐☐ Still ill: ☐ Y ☐ N Duration of illness(days): ☐☐

Symptoms experienced:	Yes	No	Not Sure	Ongoing	Duration (d)	Date of onset
Diarrhoea (3 or more loose stools in 24hrs)	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐
Bloody stools	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐
Nausea	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐
Vomiting	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐
Abdominal pain	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐
Fever	☐	☐	☐	☐	☐	☐☐/☐☐/☐☐

Other (specify): ☐

Sought healthcare: ☐ NHS Direct ☐ GP visit ☐ A&E ☐ Other (specify): ☐

Submitted stool sample: ☐ Y ☐ N

Admitted to hospital for this illness: ☐ Y ☐ N Admission date: ☐☐/☐☐/☐☐

Hospital name: ☐ Duration of stay (d): ☐

Self/Medicated with antibiotics: ☐ Y ☐ N ☐ Not sure If Y, specify: ☐

Self/Medicated with antidiarrhoeals: ☐ Y ☐ N ☐ Not sure If Y, specify: ☐

Detail any other health concerns: ☐

SECTION E: TRAVEL IN THE WEEK PRIOR TO ILLNESS

Travelled OUTSIDE of the UK: ☐ Y ☐ N ☐ Not sure

Specify countries visited (from most recent to least recent)

Country/Region	Date arrived	Date departed	Details
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐

Type of accommodation used:

☐ Hotel ☐ Bed & Breakfast ☐ Guesthouse ☐ Friends/Family

☐ Tourist Campsite ☐ Holiday dwelling ☐ Other (specify): ☐

Name of accommodation: ☐

Travelled WITHIN the UK: ☐ Y ☐ N ☐ Not sure

Specify town/resort visited (from most recent to least recent)

Town/Resort	Date arrived	Date departed	Details
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐

Type of accommodation used:

☐ Hotel ☐ Bed & Breakfast ☐ Guesthouse ☐ Friends/Family

☐ Tourist Campsite ☐ Holiday dwelling ☐ Other (specify): ☐

Name of accommodation: ☐ Postcode: ☐

April 2013 v3

SECTION F: FOOD HISTORY IN THE WEEK PRIOR TO ILLNESS

Ate out: ☐ Y ☐ N

Venues	Yes	No	Name/Location of place	Dates	Eaten
Restaurant 1	☐	☐			
Restaurant 2	☐	☐			
Takeaway1	☐	☐			
Takeaway2	☐	☐			
Café/Canteen	☐	☐			
Party/BBQ/function	☐	☐			
Other1	☐	☐			
Other2	☐	☐			

Do you handle/prepare most of the food within the household: ☐ Always ☐ Mostly ☐ Occasionally ☐ Never

Ate or handled any of the following:

		Yes	No	Product (e.g. boneless, skinless chicken breast)	Where purchased-name & location (e.g. Asda Acton, London)
HANDLED	Raw beef	☐	☐		
	Raw poultry	☐	☐		
	Raw lamb	☐	☐		
	Raw pork/gammon	☐	☐		
	Other raw meat (e.g. game, goat, ostrich etc.)	☐	☐		
	Raw vegetables	☐	☐		
	Pet/animal feed	☐	☐		
CONSUMED	Any meat	☐	☐	*If yes, indicate below:*	
	Beef	☐	☐		
	Cooked poultry	☐	☐		
	Cooked lamb	☐	☐		

April 2013 v3

CONSUMED

Cooked pork/gammon	☐	☐		
Other cooked meat e.g. game, goat, ostrich etc.)	☐	☐		
Cured meats	☐	☐		
Other processed meat	☐	☐		
Fish	☐	☐		
Shellfish	☐	☐		
Pasteurised milk	☐	☐		
Unpasteurised milk	☐	☐		
Hard cheese	☐	☐		
Soft cheese	☐	☐		
Yoghurt/fromage frais	☐	☐		
Cream	☐	☐		
Ice Cream	☐	☐		
Unpasteurised dairy products	☐	☐		
Pre-packaged salad	☐	☐		
Other salad	☐	☐		
Raw vegetables	☐	☐		
Soft fruit/berries	☐	☐		
Pre cut fruits	☐	☐		
Other Raw fruit	☐	☐		
Sprouted seeds/ beansprouts	☐	☐		
Fresh herbs	☐	☐		
Fruit juices	☐	☐		
Pre-packaged sandwiches etc.	☐	☐		
Other foods (e.g.nuts, confectionery, etc.)	☐	☐		

SECTION G: WATER EXPOSURE IN THE WEEK PRIOR TO ILLNESS

Drank unboiled water from any of the following:

Water supply	Yes	No	Details
Mains (municipal) water	☐	☐	
Private water supply (spring/well/borehole)	☐	☐	
Bottled water	☐	☐	
Unboiled river/stream/lake water	☐	☐	

Exposed to floodwater: ☐ Y ☐ N ☐ Not sure

If yes, details:

Experienced any household drainage/plumbing problems: ☐ Y ☐ N ☐ Not sure

If yes, details:

Participated in any of the following activities – either recreationally or for occupation:

Activity	Fresh water	Sea Water	No	Details
Swimming/paddling	☐	☐	☐	
Other (e.g. canoeing, fishing, sailing, sufing).	☐	☐	☐	

Is it possible that water was accidentally swallowed during any of the above: ☐ Y ☐ N ☐ Not sure

If yes, details:

SECTION H: ANIMAL CONTACT IN THE WEEK PRIOR TO ILLNESS

Contact with domestic animals/pets: ☐ Y ☐ N

If yes, indicate which animals below:

☐ Dogs ☐ Cats ☐ Rabbits ☐ Rodents ☐ Reptiles ☐ Birds ☐ Fish Other:

Did any of the above animals have diarrhoea: ☐ Y ☐ N ☐ Not sure

If yes, specify:

Was there a veterinary investigation? ☐ Y ☐ N ☐ Not sure

Contact with non domestic animals: ☐ Y ☐ N ☐ Not sure

Indicate which animals below:

Cattle ☐ Calves ☐ Sheep ☐ Lambs ☐ Goats ☐ Horses ☐ Pigs ☐ Reptiles ☐

Poultry ☐ Rabbits ☐ Deer ☐ Birds ☐ Fish ☐ Rodents ☐ Other:

Lives on, works at or access to a private farm: ☐ Y ☐ N ☐ Not sure

Attended an agricultural event e.g horse show ☐ Y ☐ N ☐ Not sure Attended date: / /

Visited a farm/petting zoo/bird reserve or such: ☐ Y ☐ N ☐ Not sure Visit date: / /

If yes, specify:

If Yes, Handled the animals: ☐ Y ☐ N ☐ Not sure

If yes, specify:

Bottle fed any of the animals: ☐ Y ☐ N ☐ Not sure

If yes, washed hands before eating food or before leaving: ☐ Y ☐ N ☐ Not sure

If yes, specify:

April 2013 v3

Consumed any food whilst there: ☐ Y ☐ N ☐ Not sure

If yes, specify:

Was the food (tick all that apply):

☐ Purchased on premises ☐ Brought from home/elsewhere

☐ Eaten in a separate area ☐ Eaten whilst in contact with the animals

☐ Eaten at a picnic table ☐ Eaten while sat on the grass/soil

SECTION I: ENVIRONMENTAL EXPOSURE IN THE WEEK PRIOR TO ILLNESS

Walked in a paddock/field where farm animals graze: ☐ Y ☐ N ☐ Not sure

If yes, specify:

Taken any day trips (beach, countryside etc.): ☐ Y ☐ N ☐ Not sure

If yes, specify:

In contact with wildlife species or their droppings: ☐ Y ☐ N ☐ Not sure

If yes, specify:

Had contact with soil, manure or sewage: ☐ Y ☐ N ☐ Not sure

If yes, specify:

SECTION J: ANY OTHER COMMENTS RELEVANT TO THIS CASE

Can the case be contacted again if further details are required: ☐

Appendix 5.2: BPSU Case Notification Form - HUS

STUDY ID: BPSU ID:

Case notification form - Strictly Confidential

The Incidence of Haemolytic Uraemic Syndrome in UK and the Republic of Ireland

Reporting Instructions:

Any patient ≤15yrs in whom a clinician has made a working diagnosis of HUS

Case Definition: Any child (≤15yrs) diagnosed with HUS

Questionnaire instructions: Y = yes, N = no, U = unknown

Section A: Reporter Details

1.1 Date of completion of questionnaire: ☐☐/☐☐/☐☐

1.2 Hospital name:

1.3 Consultant responsible for patient:

1.4 Telephone number: Email:

1.5 Has the patient been referred to/from another centre? Y ☐ N ☐ U ☐

If yes: 1) date of original admission: ☐☐/☐☐/☐☐

2) please name centre:

3) please name consultant:

Section B: Case Details

2.1 NHS/CHI No: ☐☐☐ ☐☐☐ ☐☐☐☐ Unknown: ☐

2.3 Date of admission : ☐☐/☐☐/☐☐ 2.4 Date of discharge (if applicable): ☐☐/☐☐/☐☐

2.5 Initials ☐☐☐☐☐

2.5 First half of Postcode: ☐☐☐☐☐ Town of current residence (if ROI):

2.6 Sex: M ☐ F ☐ Year of birth: ☐☐☐☐

2.7 Ethnicity*: ☐☐ If other, specify:

**Please choose the correct ethnicity code from Appendix 4i overleaf*

This page will be stored separately from the rest of the questionnaire, personal identifying information for the case will be used only for linkage records.

STUDY ID: BPSU ID:

Coding for Ethnic Group (ONS 2001 for UK wide data collection)

		Ethnicity Code	
A	White	1	Any White background
B	Mixed	2	White and Black Caribbean
		3	White and Black African
		4	White and Asian
		5	Any Other Mixed background, *please write in section B*
C	Asian or Asian British	6	Indian
		7	Pakistani
		8	Bangladeshi
		9	Any Other Asian background, *please write in section B*
D	Black or British Black	10	Caribbean
		11	African
		12	Any Other African background, *please write in section B*
E	Chinese or other ethnic group	13	Chinese
		14	Any Other, *please write in section B*
F	Unknown	15	Ethnicity not known

STUDY ID: **BPSU ID:**

Section C: Patient History

*Date measurement taken**

3.1 Weight of the patient at admission: ☐☐☐ (kgs) ☐☐/☐☐/☐☐

3.2 Height of the patient at admission: ☐☐☐ (cm) ☐☐/☐☐/☐☐

3.3 Did the patient travel outside of the UK in the two weeks prior to this admission? Y ☐ N ☐ U ☐

Country/region visited	*Date arrived*	*Date departed*	*Exact dates not known*
	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐
	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐	☐

3.4 Did the patient have any of the following in the two weeks prior to this admission (i.e. either self-prescribed or by a GP)?

	Y	*N*	*U*	*Date prescribed*	U
NSAIDs	☐	☐	☐	☐☐/☐☐/☐☐	☐
Corticosteroids	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antidiarrhoeals	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antimotility agents (e.g. loperamide)	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antibiotics/antibacterials	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antihypertensives	☐	☐	☐	☐☐/☐☐/☐☐	☐

3.5 Has the patient had HUS before? Y ☐ N ☐ U ☐

If yes, details:

3.6 Were there any pre-existing kidney diseases? Y ☐ N ☐ U ☐

If yes, details:

3.7 Were there any pre-existing urinary tract infection related problems? Y ☐ N ☐ U ☐

If yes, details:

**Please provide the measurement closest to the date of admission*

STUDY ID: **BPSU ID:**

3.8 Did anyone else in the SAME HOUSEHOLD have **bloody diarrhoea** in the TWO WEEKS prior to onset of symptoms? Y ☐ N ☐ U ☐

Did anyone else in the SAME HOUSEHOLD have **non-bloody diarrhoea** in the TWO WEEKS prior to onset of symptoms? Y ☐ N ☐ U ☐

If yes, year of birth	*Sex*	*Relationship to patient*	*Date of onset*	*U*
☐☐☐☐	M ☐ F ☐		☐☐/☐☐/☐☐	☐
☐☐☐☐	M ☐ F ☐		☐☐/☐☐/☐☐	☐

3.9 Has any other household member had HUS in the month prior to onset of symptoms? Y ☐ N ☐ U ☐

If yes, year of birth	*Sex*	*Relationship to patient*	*Date of onset*	*U*
☐☐☐☐	M ☐ F ☐		☐☐/☐☐/☐☐	☐
☐☐☐☐	M ☐ F ☐		☐☐/☐☐/☐☐	☐

Section D: Clinical features

4.1 During the admission, what was the:

			Date of measurement
Lowest haemoglobin level		(g/dL)	☐☐/☐☐/☐☐
Lowest platelet count		(x 10^9/L)	☐☐/☐☐/☐☐
Highest plasma creatinine		(micromol/L)	☐☐/☐☐/☐☐
Red blood cell fragmentation	Present ☐	Absent ☐	☐☐/☐☐/☐☐

4.2 Was HUS specifically diagnosed? Y ☐ N ☐ U ☐ When*: ☐☐/☐☐/☐☐

4.3 Did the patient have any of the following conditions?

	Y	*N*	*U*	*Date of Onset*
Diarrhoea	☐	☐	☐	☐☐/☐☐/☐☐
Bloody diarrhoea	☐	☐	☐	☐☐/☐☐/☐☐
Fever	☐	☐	☐	☐☐/☐☐/☐☐
Abdominal pain	☐	☐	☐	☐☐/☐☐/☐☐

** The date that a clinician made the diagnosis of HUS*

STUDY ID: BPSU ID:

4.4 Has the patient developed any of the following:

	Y	N	U	Date of Onset
Oligo/anuria	☐	☐	☐	☐☐/☐☐/☐☐
Septicaemia	☐	☐	☐	☐☐/☐☐/☐☐
Malignant hypertension	☐	☐	☐	☐☐/☐☐/☐☐
Major haemorrhage	☐	☐	☐	☐☐/☐☐/☐☐
Seizures and other neurological involvement	☐	☐	☐	☐☐/☐☐/☐☐
Cardiomyopathy	☐	☐	☐	☐☐/☐☐/☐☐
Diabetes mellitus	☐	☐	☐	☐☐/☐☐/☐☐
Influenza-like illness	☐	☐	☐	☐☐/☐☐/☐☐
Pneumococcus infection	☐	☐	☐	☐☐/☐☐/☐☐

Section E: Microbiology Laboratory Investigations

5.1 Has a stool specimen been submitted? Y ☐ N ☐ U ☐

If yes, date of sample: ☐☐/☐☐/☐☐ Laboratory name: ☐

5.2 Was an organism identified? Y ☐ N ☐ U ☐

If yes, name: ☐ Serotype: ☐ Phage type: ☐

5.3 If *E. coli* O157 was not isolated, was a stool specimen sent to the reference laboratory? Y ☐ N ☐ U ☐

If yes, date of sending: ☐☐/☐☐/☐☐ Date unknown: ☐

5.4 Has a serum specimen been sent to the reference laboratory? Y ☐ N ☐ U ☐

If yes, date of sample: ☐☐/☐☐/☐☐ Date unknown: ☐

5.5 Were antibodies to O157 detected? Y ☐ N ☐ U ☐

STUDY ID: **BPSU ID:**

Section F: Management

6.1 Has the patient been treated with any of the following since onset of illness*?

	Y	N	U	Date prescribed	U
NSAIDs	☐	☐	☐	☐☐/☐☐/☐☐	☐
Corticosteroids	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antidiarrhoeals	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antimotility agents (e.g. loperamide)	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antibiotics/antibacterials	☐	☐	☐	☐☐/☐☐/☐☐	☐
Antihypertensives	☐	☐	☐	☐☐/☐☐/☐☐	☐

6.2 Did the patient undergo any of the following?

Treatment	Y	N	U	Date 1st session	Date last session
Peritoneal dialysis	☐	☐	☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐
Haemodialysis	☐	☐	☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐
Plasma therapy	☐	☐	☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐
Laparotomy	☐	☐	☐	☐☐/☐☐/☐☐	
Haemofiltration	☐	☐	☐	☐☐/☐☐/☐☐	☐☐/☐☐/☐☐

6.3 Has the patient been admitted to ITU? Y ☐ N ☐ U ☐

If yes, date of admission: ☐☐/☐☐/☐☐

If yes, is the patient still in ITU? Y ☐ N ☐ U ☐

Days on ITU: ☐☐

** Where onset of illness is the onset of HUS or any prior diarrhoeal illness perceived to be related to HUS*

STUDY ID: BPSU ID:

Section G: Outcome by discharge

7.1 When did you last see the child? ☐☐/☐☐/☐☐

7.2 As far as indicated, please select from the following outcome categories:

	Y	N	U	Details
Seemingly full recovery	☐	☐	☐	
Dialysis dependent	☐	☐	☐	
Renal impairment	☐	☐	☐	
Neurological impairment	☐	☐	☐	
Continued treatment for hypertension	☐	☐	☐	
Any other	☐	☐	☐	

7.3 Has the patient died? Y ☐ N ☐ U ☐

If yes, date of death: ☐☐/☐☐/☐☐ Date unknown: ☐

Cause of death:

Section H: Notification

HUS was added to the notifiable disease list Schedule 1 of the Health Protection (Notification) Regulations (2010) such that a registered medical practitioner must notify the proper officer of the relevant local authority where they have reasonable grounds for suspecting that a patient they are attending has a particular disease.

8.1 Has the patient been notified to the Proper officer, as per Y ☐ N ☐ U ☐

If yes, date of notification: ☐☐/☐☐/☐☐ Date unknown: ☐

Section I: Comments

9.1 Is there anything else to mention?

Thank you for taking the time to complete the Questionnaire

Appendix 5.3: Characteristics of cohort participants by HUS status (n=1,059)

		No HUS n (%)	HUS n (%)
Total		852 (80.5)	207 (19.6)
IMD Quintile	1 (Least Disadvantaged)	198 (80.8)	47 (19.2)
	2	186 (84.2)	35 (15.8)
	3	166 (75.8)	53 (24.2)
	4	142 (76.8)	43 (23.2)
	5 (Most Disadvantaged)	160 (84.7)	29 (15.3)
Age group	<1	64 (91.4)	6 (8.6)
	1-4	370 (76.1)	116 (23.9)
	5-9	239 (80.7)	57 (19.3)
	10-15	179 (86.5)	28 (13.5)
Sex	Female	400 (77.5)	116 (22.5)
	Male	452 (83.2)	91 (16.8)
Age and Sex	Female <1	24 (85.7)	4 (14.3)
	Female 1-4	171 (74.0)	60 (26.0)
	Female 5-9	117 (79.1)	31 (20.9)
	Female 10-15	88 (80.7)	21 (19.3)
	Male <1	40 (95.2)	2 (4.8)
	Male 1-4	199 (78.0)	56 (22.0)
	Male 5-9	122 (82.4)	26 (17.6)
	Male 10-15	91 (92.9)	7 (7.1)
Ethnicity	White	552 (80.5)	134 (19.5)
	Non-white	138 (88.5)	18 (11.5)
	Unknown	162 (74.7)	55 (23.4)
Travel	Yes	128 (85.3)	22 (14.7)
	No	724 (79.7)	185 (20.4)
Rurality	Rural	230 (80.4)	56 (19.6)
	Urban	622 (80.5)	151 (19.5)
Region	East Midlands	65 (81.3)	15 (18.8)
	East of England	57 (80.3)	14 (19.7)
	London	93 (81.6)	21 (18.4)
	North East	64 (77.1)	19 (22.9)
	North West	153 (77.7)	44 (22.3)
	South East	92 (78.6)	25 (21.4)
	South West	101 (75.9)	32 (24.1)
	West Midlands	96 (84.2)	18 (15.8)
	Yorkshire and Humber	131 (87.3)	19 (12.7)
Stx	*Stx*1	17 (94.4)	1 (5.6)
	*Stx*2	609 (81.7)	136 (18.3)
	*Stx*1+2	219 (96.9)	7 (3.1)
	Serology	7 (10.6)	59 (89.4)
	Unknown	0 (0.0)	4 (100.0)
Symptoms	Diarrhea	803 (80.3)	197 (19.7)
	Bloody diarrhea	432 (74.0)	152 (26.0)
	Nausea	278 (75.8)	89 (24.3)
	Vomiting	330 (66.1)	169 (33.9)
	Abdominal pain	574 (78.2)	160 (21.8)
	Fever	273 (76.7)	83 (23.3)
Healthcare contact	Antibiotics	53 (40.8)	77 (59.2)
	NHS Direct	67 (72.0)	26 (28.0)
	GP	570 (83.7)	111 (16.3)
	A&E	186 (66.9)	92 (33.1)
	Other healthcare contact	98 (74.8)	33 (25.2)
	Hospital	223 (52.4)	203 (47.6)

HUS – hemolytic uraemic syndrome; *stx* – shiga toxin; NHS Direct – National Health Service telephone advice line, now NHS 111; GP – General Practitioner; A&E – accident and emergency

Appendix 6: Publications from this thesis

Appendix 6.1: List of publications

Adams, N., Rose, T. C., Taylor-Robinson, D., Barr, B., Hawker, J., O'Brien S, J., Violato, M. & Whitehead, M. 2015. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review. *PROSPERO International prospective register of systematic reviews*.

Adams, N. L., Rose, T. C., Hawker, J., Violato, M., O'Brien, S. J., Whitehead, M., Barr, B. & Taylor-Robinson, D. C. 2017. Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 study). *Eur J Public Health*.

Rose, T. C., **Adams, N. L**., Barr, B., Hawker, J., O'Brien, S. J., Violato, M., Whitehead, M. & Taylor-Robinson, D. C. 2017. Socioeconomic status is associated with symptom severity and sickness absence in people with infectious intestinal disease in the UK. *BMC Infect Dis,* 17**,** 447.

Appendix 6.2: List of abstracts

N Adams, T Rose, D Taylor-Robinson, B Barr, S O'Brien, M Violato, J Hawker, M Whitehead; Does socioeconomic status influence risk of gastrointestinal infections in the community in the UK?: Natalie Adams, *European Journal of Public Health*, Volume 26, Issue suppl_1, 1 November 2016, ckw174.167, https://doi.org/10.1093/eurpub/ckw174.167

T Rose, **N Adams**, D Taylor-Robinson, B Barr, J Hawker, S O'Brien, M Violato, M Whitehead; Relationship between socioeconomic status and measures of infectious intestinal disease severity: Tanith Rose, *European Journal of Public Health*, Volume 26, Issue suppl_1, 1 November 2016, ckw166.060, https://doi.org/10.1093/eurpub/ckw166.060

N Adams, T Rose, D Taylor-Robinson, B Barr, S O'Brien, M Violato, J Hawker, M Whitehead; relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review and meta-analysis: Natalie Adams and Tanith Rose, Society of Social Medicine Annual Scientific Meeting, 8 September 2017

Appendix 6.3 Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 study)

European Journal of Public Health, 1–5
doi:10.1093/eurpub/ckx091

Socioeconomic status and infectious intestinal disease in the community: a longitudinal study (IID2 study)

Natalie L. Adams[1,2,3], Tanith C. Rose[1,2], Jeremy Hawker[1,3], Mara Violato[1,4], Sarah J. O'Brien[1,2], Margaret Whitehead[1,2], Benjamin Barr[1,2], David C. Taylor-Robinson[1,2]

1 NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, UK
2 Department of Public Health and Policy, University of Liverpool, Liverpool, UK
3 National Infection Service, Public Health England, London/Birmingham, UK
4 Health Economics Research Centre, University of Oxford, Oxford, UK

Benjamin Barr and David C. Taylor-Robinson are joint senior authors.

Correspondence: Natalie L. Adams, Gastrointestinal Infections Department, National Infection Service, Public Health England, 61 Colindale Ave, Colindale NW9 5EQ, UK, Tel: 0208 327 6616, e-mail: natalie.adams@phe.gov.uk

Background: Infectious intestinal diseases (IID) are common, affecting around 25% of people in UK each year at an estimated annual cost to the economy, individuals and the NHS of £1.5 billion. While there is evidence of higher IID hospital admissions in more disadvantaged groups, the association between socioeconomic status (SES) and risk of IID remains unclear. This study aims to investigate the relationship between SES and IID in a large community cohort. **Methods:** Longitudinal analysis of a prospective community cohort in the UK following 6836 participants of all ages was undertaken. Hazard ratios for IID by SES were estimated using Cox proportional hazard, adjusting for follow-up time and potential confounding factors. **Results:** In the fully adjusted analysis, hazard ratio of IID was significantly lower among routine/manual occupations compared with managerial/professional occupations (HR 0.74, 95% CI 0.61–0.90). **Conclusion:** In this large community cohort, lower SES was associated with lower IID risk. This may be partially explained by the low response rate which varied by SES. However, it may be related to differences in exposure or recognition of IID symptoms by SES. Higher hospital admissions associated with lower SES observed in some studies could relate to more severe consequences, rather than increased infection risk.

Introduction

Infectious intestinal disease (IID) is common, leading to diarrhoea, vomiting and, occasionally, more serious complications such as renal failure. Previous estimates suggest around 25% of people in UK suffer an episode of IID per year[1] and that foodborne illness in England and Wales costs individuals, the economy and NHS around £1.5 billion annually.[2] Many infections are socially patterned, however, the role of socioeconomic status (SES) in risk of IID in developed countries, such as UK, is not well understood.[3]

A large proportion of the burden of IID remains hidden; it is estimated that there are 147 cases in the community for every one case reported to national surveillance;[2] many individuals do not present to healthcare as most infections are self-limiting. Additionally, it is unclear whether socioeconomic patterns reported in hospital and laboratory-based surveillance reflect differences in risk of infection or in reporting and healthcare-seeking behaviour.[4] Longitudinal population-based survey data can provide better estimates of differences in risk of infection that may not be captured through routine surveillance. This study aims to explore whether different socioeconomic groups experience different risk of IID in the UK, through the analysis of a large prospective population cohort, to improve understanding of the role of SES in IID in the community and to inform policies to reduce health inequalities. In this study, we provide an up-to-date assessment of the association between IID and SES for all ages in UK.

Methods

Design, setting and data source

We undertook a longitudinal analysis of data collected through a large prospective community cohort in UK (IID2 study).[1,2] A cohort of 6836 randomly selected participants was recruited from 88 representative general practices in UK. Sociodemographic information including age, gender and occupation were obtained through a baseline survey upon entry to the cohort and details of IID symptoms were recorded on a weekly basis for up to 1 year, from October 2007 to August 2009, through the return of an email or postcard indicating whether symptoms of diarrhoea and/or vomiting had been experienced in the previous week. Individuals who reported symptoms completed a more in-depth questionnaire through which details of illness and healthcare contact were recorded.

Overall participation rate was low (9%) and individuals who declined to participate were younger, more deprived, living in urban rather than rural areas and employed in lower supervisory and technical occupations.[2] The 6836 participants contributed 4658 person-years of follow-up; median follow-up duration was 39 weeks.[2] Among participants, no differences in follow-up were identified by sex, SES or rural–urban classification.[2] Average follow-up time was similar for those who experienced an episode of IID and those who did not.[2] Managerial/professional occupations were over-represented in the study, while intermediate, and semi-routine and routine occupations were under-represented, in comparison to the UK population.[2] Those of White ethnicity were slightly over-represented.[2]

Ethical approval and informed consent were originally obtained for the main study (07/MRE08/5). This included the provision to use the data for future research. Approval for this secondary analysis of the fully anonymised datasets was not required.[2]

Infectious intestinal disease was defined as loose stools or clinically significant vomiting (vomiting occurring more than once in 24 h and if it incapacitated the case or was accompanied by other symptoms such as cramps or fever)[2] lasting <2 weeks, in the absence of a known non-infectious cause, preceded by a symptom-free

period of 3 weeks.[2] Cases experiencing illness considered to be travel-related were excluded.

The primary exposure of interest was an individual-level measure of SES, self-reported occupation, with each individual assigned a National-Statistics Socioeconomic Classification (NS-SEC) using the five-class self-coded version.[5] For participants aged less than 16 years, NS-SEC was assigned based on the occupation of the head of the household. For the purposes of this study, the NS-SEC variable was recoded into the three-class version to provide a hierarchy of SES, with routine/manual occupations assumed equivalent to low SES and managerial/professional occupations to high SES.[5]

Analysis strategy

Analyses were conducted in Stata 13.1 (Statacorp, TX). Rates of IID within the study population and by SES were calculated accounting for follow-up time, to produce rates of IID per 1000 person-years with associated 95% confidence intervals. The main analysis investigated the relationship between SES, as measured by NS-SEC, and time to first IID episode for each participant using Cox proportional hazard regression modelling, with subsequent episodes of IID for an individual being dropped. We first explored univariate relationships between SES and the covariates of interest [rurality and employment status (employed/not working)] before fitting a multivariate Cox proportional hazard regression model, adjusting for the potentially confounding covariates and stratifying the baseline hazard on age and sex. Kaplan–Meier survival curves were estimated to check the proportional hazards assumption. Interaction terms between the socioeconomic variable NS-SEC and each variable in turn were tested for inclusion to investigate whether the strength of any relationship was moderated by the inclusion of another variable.

We undertook a number of robustness tests, firstly allowing individuals with multiple episodes of IID to re-enter the cohort following a period of censoring (due to symptoms meeting the case definition and requiring a censored period of 3 weeks after cessation of symptoms; non-response; or symptoms not meeting the case definition), accounting for clustering within individuals by using a robust estimate of variance allowing for inter-person correlation.

We repeated the analysis using a less sensitive case definition, whereby individuals reporting symptoms which could not be verified against the case definition (due to a lack of further details about foreign travel or symptom duration) were also included as cases in the analysis. We repeated the analysis including those unclassifiable within NS-SEC to investigate whether this had an impact on the results. This NS-SEC group comprised individuals for whom it was not possible to classify their occupation or who did not respond to occupation questions.

Stratification by age group was conducted to determine whether there were differences in the rate of IID by SES for children, adults and older participants. We repeated the analysis using an area-level measure of SES, the Index of Multiple Deprivation (IMD),[6] assigned to each individual based on their postcode.

As there were missing NS-SEC data for a group of participants for whom it was not possible to classify their occupation or who did not respond to the occupation question, Multiple Imputation using chained equations (MICE)[7] was used in order to include these cases.

Results

Characteristics of participants

Of the 6836 participants in the cohort, 998 individuals reported an episode of IID during 4583.5 person-years of follow-up. Fifty-two percent ($n = 3557$) were from managerial/professional occupations, 15% ($n = 1002$) were from intermediate occupations and 17% ($n = 1165$) were from routine/manual occupations, compared with 31%, 22% and 33% respectively in the general population.[8] For 1112 individuals (16.3%), NS-SEC was missing either because they did not respond to occupation questions or, if they did, it was not possible to classify their occupation according to the NS-SEC categories. SES was associated with age group, sex, rurality, employment status and the method of follow-up that participants elected to use (email or postcard). It was independent of ethnicity (table 1). Mean follow-up time was similar between NS-SEC groups.

Incidence was lower among routine/manual occupations compared with managerial/professional occupations (166.3/1000 person-years, 95% CI 140–197; 235.4/1000 person-years, 95% CI 217–256 (figure 1).

Main analysis

Participants for whom NS-SEC was not classifiable were excluded from the main analysis; 5724 participants were included. All potentially confounding variables were retained in the fully adjusted model; ethnicity and follow-up type were excluded as these were not considered to be confounders (table 2). IID hazard was significantly lower in routine/manual occupations compared with managerial/professional occupations (HR 0.74, 95% CI 0.61–0.90). No significant interactions were identified.

Sensitivity analyses

The lower hazard in routine/manual occupations compared with managerial/professional occupations was a consistent finding across the sensitivity analyses accounting for multiple spells of follow-up; using a less sensitive case definition; including the not-classifiable NS-SEC group; and using multiple imputation for NS-SEC (Appendices B–E).

In the models stratifying for age (Appendices F.1–F.3), the Hazard Ratio for routine/manual occupations compared with managerial/professional occupations tended to decrease with increasing age (65 and over: 0.60, 95% CI 0.40–0.89, $P = 0.012$; 0–17 years: 0.89 (95% CI 0.61–1.29, $P = 0.54$), however, these differences were non-significant.

Using the area-level IMD as a measure of SES, the most deprived (IMD quintile 1) had lower incidence compared with the least deprived (IMD quintile 5) (171.9/1000 person-years, 95% CI 132.6–222.8; 234.1, 95% CI 206.9–264.8) in accordance with the main analysis results. However, no statistically significant relationship was identified in the adjusted analysis (Supplementary Appendix G). The distribution of SES by IMD differed compared with the general population, with those in the most deprived quintile being underrepresented (7% versus 20%) and in the least deprived quintile (24% versus 20%) compared with the distribution in the general population.[2] No significant interactions were identified in any of the sensitivity analyses.

Discussion

In this analysis of a large representative UK sample following a prospective community cohort to monitor the development of IID symptoms, we investigated the relationship between IID and SES using occupation as an individual-level measure of SES. Lower SES was associated with significantly lower risk of IID. There were no significant age-stratified differences in the relationship between IID and SES.

We undertook a novel analysis of an existing population-based community cohort assessing the association of both individual and area-based measures of SES with IID. Survival analysis explored the relationship between IID and SES accounting for censored observations and different time to event for participants. Multiple sensitivity analyses were conducted to assess the robustness of the main results. A key strength of this study is that it does not require an individual

Table 1 Characteristics of cohort participants ($n = 6836$)

	Managerial/professional *n* (%)	Intermediate *n* (%)	Routine/manual *n* (%)	Not classifiable *n* (%)	*P* value
Total	3557 (52.0)	1002 (14.7)	1165 (17.0)	1112 (16.3)	
Incidence rate/1000 PYs	235.4	243.9	166.3	194.0	
Follow-up time (mean days)	242.1	240.6	245.2	257.4	
Age (mean)	47.2	48.5	49.3	53.0	
Case					
Yes	555 (55.6)	161 (16.1)	130 (13.0)	152 (15.2)	0.001
No	3002 (51.4)	841 (14.4)	1035 (17.7)	960 (16.4)	
Age group					
<18	605 (55.5)	152 (13.9)	178 (16.3)	156 (14.3)	<0.001
18-64	2095 (54.6)	593 (15.4)	627 (16.3)	525 (13.7)	
65+	857 (45.0)	257 (13.5)	360 (18.9)	431 (22.6)	
Sex					
Female	2175 (52.3)	669 (16.1)	640 (15.4)	676 (16.3)	<0.001
Male	1382 (51.6)	333 (12.4)	525 (19.6)	436 (16.3)	
Ethnicity					
White	3464 (52.0)	981 (14.7)	1145 (17.2)	1077 (16.2)	0.125
Non-White	93 (55.0)	21 (12.4)	20 (11.8)	35 (20.7)	
Rurality					
Urban	2522 (50.8)	694 (14.0)	958 (19.3)	789 (15.9)	<0.001
Rural	1034 (66.8)	307 (19.8)	206 (13.3)	323 (17.3)	
Follow-up					
Email	2564 (60.3)	622 (14.6)	529 (12.4)	539 (12.7)	<0.001
Postcard	993 (38.5)	380 (14.7)	636 (24.6)	573 (22.2)	
Employment status					
Employed	2493 (56.7)	713 (16.2)	769 (17.5)	423 (9.6)	<0.001
Not working	1061 (44.1)	287 (11.9)	396 (1.5)	664 (27.6)	

Notes: PYs, person-years. Missing data: Employment status was missing for 30 individuals. Rural–urban classification was missing for three individuals.

Table 2 Adjusted and unadjusted Cox regression analysis (n subjects = 5716; n failures = 845)

Variable	Category	Unadjusted		Adjusted[a]		*P* value
		Hazard ratio	(95% CI)	Hazard ratio	(95% CI)	
NS-SEC	Managerial/professional	1.0 (reference)		1.0 (reference)		
	Intermediate	1.04	(8.87–1.23)	1.03	(0.86–1.23)	0.74
	Routine/manual	0.71	(0.58–0.86)	**0.74**	**(0.61–0.90)**	**0.002**
Rurality	Urban	1.0 (reference)		1.0 (reference)		
	Rural	1.17	(1.01–1.36)	1.13	(0.98–1.31)	0.09
Employment status	Employed	1.0 (reference)		1.0 (reference)		
	Not working	0.78	(0.67–0.91)	1.00	(0.82–1.22)	1.00

Notes: Baseline hazard stratified by age group and sex. Missing data: NS-SEC was not classifiable for 1112 individuals. Employment status was missing for five individuals. Rural–urban classification was missing for three individuals. NS-SEC, National Statistics-Socioeconomic Classification; CI: confidence interval.
a: Adjusted for all other covariates in the model.

to seek care or have a specimen taken in order to be included in the study, thus reducing potential bias if healthcare-seeking behaviour differs by SES.

However, participation in the cohort study was low; only around 9% of the original number recruited and screened for participation, lower than the first IID study (35%),[2] and this varied by SES. Participation bias within cohort studies, particularly by SES, is a recognised limitation. The characteristics of the cohort population differed from the UK population, as those who were most disadvantaged were underrepresented compared with the UK population, while those who were advantaged were over-represented,[8] and a large number of participants ($n = 1112$) could not be classified by NS-SEC. It is possible that those who agreed to participate had a different risk of IID compared with those who refused which may limit the generalisability of results. The lack of a significant difference in risk by SES for children could be related to small numbers in the stratified groups which means the study may lack power for detecting a difference, although the trend was of a lower risk for lower SES participants.

However, despite these limitations, this study represents an important analysis of a large prospective community cohort in UK which suggests differences in risk of IID by SES among the population within this study. To the best of our knowledge, this is the most comprehensive analysis of IID by SES conducted in UK. Our study differs from two earlier analyses of the IID2 data. Tam et al.[9] used data from the IID2 study and found no significant difference in risk of multiple-spells of IID in disadvantaged compared with advantaged individuals, while Tam et al.[1] found no significant difference in incidence by socioeconomic groups. The different findings between these papers could relate to differences in research questions which were answered using different and question-specific methods, as well as differences in the outcome; as our outcome was time to event, our paper used survival analysis to account for differing follow-up times.

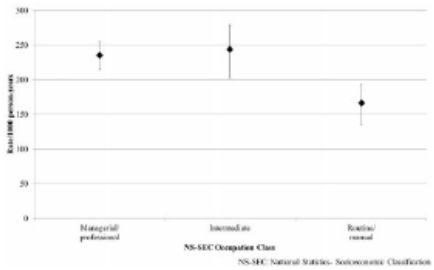


Figure 1 Incidence rates per 1000 person-years by NS-SEC classification. NS-SEC, National Statistics-Socioeconomic Classification

Despite potential issues with participation bias by SES, cohort studies are a robust method of assessing individual-level exposures. However, few population-based cohort studies have been conducted in developed countries to investigate differences in IID risk by SES; studies investigating this relationship between age groups are particularly limited.

In a Dutch cohort study, individuals with a low level of education had significantly lower odds of gastroenteritis compared with those with a high level of education (OR 0.65, 95% CI 0.56–0.75);[10] comparable with our adjusted estimate. Another cohort study,[11] in Denmark, which looked at specific bacterial pathogens as opposed to IID, found an increased risk in adults in higher SES groups for most pathogens (*Campylobacter*, *Salmonella* Enteritidis and *Shigella*), however, the pattern was less clear in children, with no association between risk and SES for most pathogens; these findings also concur with our results.

By contrast, a Canadian cohort study[12] found that individuals in neighbourhoods with low and medium household incomes had higher rates of IID compared with those living in neighbourhoods with high household incomes. In contrast to the other cohort studies above, including our study, the authors used physician visits rather than self-reported symptoms to define IID; when hospitalisation was used to define IID, the authors found no significant difference in rates by SES. Further, this study was designed to assess the association between environmental factors and IID incidence rather than SES specifically.

Several cohort studies which have focussed on children have found higher risk in more disadvantaged groups,[13–16] in contrast with our findings. However, two of these studies[13,16] were from the same survey, although used different SES measures to investigate the relationship, and specifically sampled very young children. Studies assessing SES specifically in children may be better powered or designed to investigate this relationship than studies looking at all ages, particularly as SES is more transient in children.

Many studies assessing the relationship between IID and SES in developed countries have used study designs other than population-based cohorts, such as cross-sectional population surveys, which have produced conflicting results. Some support our finding that lower SES is associated with lower risk of IID.[17–22] These studies looked at adults specifically or all ages combined and used mainly education as a measure of SES, with the exception of one study which used occupation.[17] Most cross-sectional population surveys, however, found no significant association,[19,20,22–31] including three studies which found significant associations with education but not with income and occupation,[19,20,22] suggesting that the association may vary with different measures of SES. The variability in these results also suggests that cross-sectional study designs may not provide the most robust estimates of the relationship between SES and IID.

There are several possible explanations for the finding of lower IID rates among individuals of lower SES. It may be artefactual and related to low response rate. The over-representation of advantaged individuals, or differential reporting by SES, may have resulted in a biased population. However, the sample was large, and the internal associations, which were the targets of inference within the sample population, are likely to be valid. Conversely, differences in the recognition or reporting of symptoms by SES or by healthcare seeking behaviour may partially explain the results. The results may also represent a real lower risk of IID among those who are disadvantaged through differences in exposures by SES (such as the consumption of less risky foods, reduced opportunity to eat meals outside of the home, reduced exposure to animal attractions, such as open farms, and reduced levels of foreign travel among those of a lower SES).[17,21]

There is some evidence from our study and others to suggest the existence of a relationship between IID and SES, with lower SES associated with lower rates of IID. Evidence from the literature, however, suggests that the consequences of IID are more severe for more disadvantaged population groups, with higher hospital admission rates for those of lower SES,[32–34] and that disadvantaged children may be at higher risk of IID infections.[13–16] Our results may underestimate the risk in disadvantaged groups and in children. While more disadvantaged individuals may be at a lower risk of, or vulnerability to, GI infections, the possibility of more severe consequences among these groups has implications for the clinical management of IID and for healthcare utilisation.

Further research is required to explore the role of symptom recognition, perception, healthcare seeking behaviour and other potentially mediating exposures to complement these results and help to explain the relationship between SES and GI infection. Focussing on children may clarify the inconsistent results seen across the literature, as would further research on the most appropriate SES measure to use to produce the most robust estimates of the association between IID and SES. Finally, a greater understanding of the individual behaviours and environmental risk factors by SES is crucial to understanding differential risk, vulnerability and consequences of IID. These results contribute to the evidence on community-level risk of GI infections. Alongside future planned analyses, this could ultimately be used to provide evidence to inform policies to address inequalities in risk, vulnerability and consequences of IID.

Acknowledgements

The authors would like to acknowledge Dr John Harris for his advice on the IID2 data, the IID2 study team for granting access to the data and all the participants, study nurses, general practitioners, practice staff, laboratory, research and administrative staff who took part in the IID2 study. This work was funded by the National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections.

Funding

This work was supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections [grant number NIHR HPRU 201210,038] at University of Liverpool in partnership with Public Health England (PHE), in collaboration with University of East Anglia, University of Oxford and the Quadram Institute.

Supplementary data

Supplementary data are available at *EURPUB* online.

Disclaimer

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

Conflict of interest: None declared.

Key points

- Gastrointestinal infections are common. There is evidence to suggest that some consequences of gastrointestinal infections, such as hospital admission, are greater in more disadvantaged individuals.
- However, the role of SES in risk of gastrointestinal infections is not well understood and studies that have investigated this relationship have presented conflicting findings.
- There is some evidence to suggest that, within the community, disadvantaged individuals are at lower risk of gastrointestinal infections compared with more advantaged individuals which could be a result of differences in exposures, healthcare seeking behaviours or symptom recognition.

References

1 Tam CC, Rodrigues LC, Viviani L, et al. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut* 2011;61:69–77.

2 Food Standards Agency. The second study of infectious intestinal disease in the community (IID2 Study). 2016. Available at: https://www.food.gov.uk/science/research/foodborneillness/b14programme/b14projlist/b18021 (8 November 2016, date last accessed).

3 Newman KL, Leon JS, Rebolledo PA, Scallan E. The impact of socioeconomic status on foodborne illness in high-income countries: a systematic review. *Epidemiol Infect* 2015;143:1–13.

4 Dunlop S, Coyte PC, McIsaac W. Socio-economic status and the utilisation of physicians' services: results from the Canadian National Population Health Survey. *Soc Sci Med* 2000;51:123–33.

5 Office for National Statistics. *Standard Occupational Classification 2010 Volume 3: The National Statistics Socio-Economic Classification (NS-SEC rebased on SOC2010) User Manual*. Office for National Statistics, 2010.

6 Department for Communities and Local Government. *English Indices of Deprivation*, 2010. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf (16 September 2016, date last accessed).

7 UCLA Institute for Digital Research and Education. *Multiple Imputation in STATA*. Available at: http://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/ (24 April 2017, date last accessed).

8 Office for National Statistics [dataset]. *UK Population Based on the 2011 Census: Key Statistics KS611UK – NS-SEC*. Available at: https://www.nomisweb.co.uk/query/select/getdatasetbytheme.asp?opt=3&theme=&subgrp=(16 September 2016, date last accessed).

9 Tam CC, Vivani L, Adak B, et al. The second study of infectious intestinal disease (IID2): increased rates of recurrent diarrhoea in individuals aged 65 years and above. *BMC Public Health* 2013;13:1–8.

10 de Wit MA, Koopmans MP, Kortbeek LM, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. *Am J Epidemiol* 2001;154:666–74.

11 Simonsen J, Frisch M, Ethelberg S. Socioeconomic risk factors for bacterial gastrointestinal infections. *Epidemiology* 2008;19:282–90.

12 Teschke K, Bellack N, Shen H, et al. Water and sewage systems, socio-demographics, and duration of residence associated with endemic intestinal infectious diseases: a cohort study. *BMC Public Health* 2010;10:767.

13 Baker D, Taylor H, Henderson J. Inequality in infant morbidity: causes and consequences in England in the 1990s. *J Epidemiol Community Health* 1998;52:451–8.

14 Ludvigsson JF. Epidemiological study of constipation and other gastrointestinal symptoms in 8000 children. *Acta Paediatr* 2006;95:573–80.

15 Eaton-Evans J, Dugdale AE. Effects of feeding and social factors on diarrhoea and vomiting in infants. *Arch Dis Child* 1987;62:445–8.

16 Beale N, Peart C, Kay H, et al. 'ALSPAC' infant morbidity and council tax band: doctor consultations are higher in lower bands. *Eur J Public Health* 2010;20:403–8.

17 Scallan E, Fitzgerald M, Collins C, et al. Acute gastroenteritis in northern Ireland and the Republic of Ireland: a telephone survey. *Commun Dis Public Health* 2004;7:61–7.

18 Fein SB, Lin CTJ, Levy AS. Foodborne illness: perceptions, experience, and preventive behaviors in the United States. *J Food Prot* 1995;58:1405–11.

19 Herikstad H, Yang S, Van Gilder TJ, et al. A population-based estimate of the burden of diarrhoeal illness in the United States: FoodNet, 1996-7. *Epidemiol Infect* 2002;129:9–17.

20 Majowicz SE, Horrocks J, Bocking K. Demographic determinants of acute gastrointestinal illness in Canada: a population study. *BMC Public Health* 2007;7:8.

21 Pollard CM, Meng X, Williamson S, et al. Eating out is associated with self-reported food poisoning a Western Australia population perspective, 1998 to 2009. *Public Health Nutr* 2014;17:2270–7.

22 Van Cauteren D, De Valk H, Vaux S, et al. Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study. *Epidemiol Infect* 2012;140:697–705.

23 Doorduyn Y, Van Pelt W, Havelaar AH. The burden of infectious intestinal disease (IID) in the community: a survey of self-reported IID in The Netherlands. *Epidemiol Infect* 2012;140:1185–92.

24 Evans MR, Sarvotham T, Thomas DR, Howard AJ. Domestic and travel-related foodborne gastrointestinal illness in a population health survey. *Epidemiol Infect* 2006;134:686–93.

25 Hall GV, Kirk MD, Ashbolt R, et al. Frequency of infectious gastrointestinal illness in Australia, 2002: regional, seasonal and demographic variation. *Epidemiol Infect* 2006;134:111–8.

26 Majowicz SE, Dore K, Flint JA, et al. Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community. *Epidemiol Infect* 2004;132:607–17.

27 McAteer A, Elliott AM, Hannaford PC. Ascertaining the size of the symptom iceberg in a UK-wide community-based survey. *Br J Gen Pract* 2011;61:e1–e11.

28 Wilking H, Spitznagel H, Werber D, et al. Acute gastrointestinal illness in adults in Germany: a population-based telephone survey. *Epidemiol Infect* 2013;141:2365–75.

29 Adlam SB, Perera S, Lake RJ, et al. Acute gastrointestinal illness in New Zealand: a community study. *Epidemiol Infect* 2011;139:302–8.

30 Sargeant JM, Majowicz SE, Snelgrove J. The burden of acute gastrointestinal illness in Ontario, Canada, 2005-2006. *Epidemiol Infect* 2008;136:451–60.

31 Jones TF, McMillian MB, Scallan E, et al. A population-based estimate of the substantial burden of diarrhoeal disease in the United States; FoodNet, 1996-2003. *Epidemiol Infect* 2007;135:293–301.

32 Pockett RD, Adlard N, Carroll S, Rajoriya F. Paediatric hospital admissions for rotavirus gastroenteritis and infectious gastroenteritis of all causes in England: an analysis of correlation with deprivation. *Curr Med Res Opin* 2011;27:777–84.

33 Olowokure B, Hawker J, Weinberg J, et al. Deprivation and hospital admission for infectious intestinal diseases. *Lancet* 1999;353:807–8.

34 Baker MG, Barnard LT, Kvalsvig A, et al. Increasing incidence of serious infectious diseases and inequalities in New Zealand: a national epidemiological study. *Lancet* 2012;379:1112–9.

Appendix 6.4: Socioeconomic status is associated with symptom severity and sickness absence in people with infectious intestinal disease in the UK

Rose et al. BMC Infectious Diseases (2017) 17:447
DOI 10.1186/s12879-017-2551-1

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Socioeconomic status is associated with symptom severity and sickness absence in people with infectious intestinal disease in the UK

Tanith C. Rose[1,2,5*], Natalie L. Adams[1,2,3], Benjamin Barr[1,2†], Jeremy Hawker[1,3], Sarah J. O'Brien[1,2], Mara Violato[1,4], Margaret Whitehead[1,2] and David C. Taylor-Robinson[1,2†]

Abstract

Background: The burden of infectious intestinal disease (IID) in the UK is substantial. Negative consequences including sickness absence are common, but little is known about the social patterning of these outcomes, or the extent to which they relate to disease severity.

Methods: We performed a cross-sectional analysis using IID cases identified from a large population-based survey, to explore the association between socioeconomic status (SES) and symptom severity and sickness absence; and to assess the role of symptom severity on the relationship between SES and absence. Regression modelling was used to investigate these associations, whilst controlling for potential confounders such as age, sex and ethnicity.

Results: Among 1164 cases, those of lower SES versus high had twice the odds of experiencing severe symptoms (OR 2.2, 95%CI;1.66–2.87). Lower SES was associated with higher odds of sickness absence (OR 1.8, 95%CI;1.26–2.69), however this association was attenuated after adjusting for symptom severity (OR 1.4, 95%CI;0.92–2.07).

Conclusions: In a large sample of IID cases, those of low SES versus high were more likely to report severe symptoms, and sickness absence; with greater severity largely explaining the higher absence. Public health interventions are needed to address the unequal consequences of IID identified.

Keywords: Socioeconomic factors, Occupation, Infectious intestinal disease, Diarrhoea, Sick leave, Symptom severity

Background

Infectious intestinal disease (IID) is extremely common, with an estimated 17 million sporadic cases occurring each year in the United Kingdom (UK) [1]. It also confers significant morbidity and associated healthcare costs. Around half of those who experience IID report absence from work or school which amounts to an estimated loss of nearly 19 million days per annum, with potential ramifications for adult earnings and child education [2]. Additionally, there are approximately one million general practice (GP) consultations for IID every year in the UK [1]. The burden of IID is clearly evident, yet relatively little is known about the extent of socioeconomic inequalities in the clinical, social and economic consequences of IID.

Studies conducted in developed countries, suggest individuals of low socioeconomic status (SES) compared to high, have higher rates of GP consultation [3, 4] and hospital admission due to IID [5–9]. For example, in the West Midlands in the UK, hospital admission rates for young children with IID were twice as high in the most deprived areas compared to the least [6]. However, the mechanisms explaining these apparent health inequalities are unknown. Contributing factors may include differential risk of infection, healthcare seeking behaviour, or disease severity across socioeconomic groups.

* Correspondence: Tanith.Rose@liverpool.ac.uk
† Equal contributors
[1]NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, UK
[2]Department of Public Health and Policy, University of Liverpool, Liverpool, UK
Full list of author information is available at the end of the article

Separating out the effects of these potential explanations is imperative to understand the role they play in generating the inequalities observed, and so that interventions and policies can be developed to tackle the problem. A cross-sectional analysis of IID cases identified in the English IID1 studies, showed that IID cases of lower SES (as measured by educational attainment) were more likely to present to their GP for an episode of IID, compared to those of higher SES [3]. In addition, disease severity was strongly predictive of GP presentation for IID, however numbers were insufficient to assess the relationship between SES and IID severity. These findings indicate that healthcare seeking behaviour for IID may be socially patterned, which potentially could be related to disease severity.

Negative consequences of IID also include sickness absence, which may or may not be related to IID severity. Rates of general (all cause) sickness absence, have been shown to be higher for those of lower SES compared to high [10], however some studies have demonstrated that this association can in part be explained by the increased levels of morbidity for those of lower SES [10, 11]. The few studies that have investigated the relationship between SES and sickness absence due to IID have produced conflicting results [12, 13]; and we are yet to find a study that has examined the role of IID severity on the relationship between SES and sickness absence. To gain a better understanding of inequalities in the consequences of IID, we analysed a large sample of IID cases to explore the association between SES and measures of self-reported IID symptom severity and sickness absence.

Methods

Study design and data source

We analysed cases of IID identified in the population-based IID2 study. The IID2 study was conducted across the UK in 2008–9 and contained several studies, the methods of which have been described in detail elsewhere [14]. The IID2 study was granted ethical approval by the North West Research Ethics Committee (07/MRE08/5) [15]. Participants gave written informed consent for their anonymised data to be used for future analyses.

The IID2 study contained two major components; a prospective cohort study and a GP presentation study. For the cohort study, patients were randomly selected from the registers of 88 general practices and invited to participate. Participants completed a baseline questionnaire containing questions on socio-demographic factors, and were followed-up weekly for one year to determine the incidence of IID. Incident cases completed symptom questionnaires including questions on symptom severity, absenteeism and recent foreign travel. For the GP presentation study, all patients who consulted their GP for an episode of IID across 37 of the 88 general practices, over a one year period were invited to participate in a survey which included the same socio-demographic and symptom questions as the former study.

Cases identified via both components of the IID2 study were combined for this analysis. Cases of IID were defined as people aged five years or older, with loose stools or clinically significant vomiting lasting less than two weeks, in the absence of a known non-infectious cause, preceded by a symptom-free period of three weeks [14]. We included cases aged five years or older, to limit potential misclassification of the more subjective symptoms, such as headache and nausea, in young children (see below details of symptom severity score). For cases meeting the case definition, all recurrent episodes of IID were removed regardless of the timeframe between episodes. If a case experienced more than one episode of IID during follow-up, only information related to the first episode was retained to create a sample of independent observations.

Outcomes and covariates

The outcomes of interest were symptom severity and sickness absence due to IID. The symptom severity score was derived from information on the presence/absence of nine symptoms, and the duration of four symptoms, which were self-reported by the cases, using previously published methods [3]. In brief, the presence and duration scores were multiplied, and the resulting product scores summed across the symptoms, creating an overall symptom severity score for each case (Additional file 1). The symptom severity variable was converted into tertiles, whereby three approximately equally sized groups were created according to the distribution of the severity score [3]. The second outcome of interest was sickness absence; a binary variable indicating whether the episode of IID prevented the case from going to work or school. Sickness absence was only defined for cases of school or working age (aged five years or older, and up to 60 years for women and 65 years for men, as older age groups were unlikely to be in work or education [16]).

The main exposure of interest was SES measured at the individual-level using the National Statistics Socioeconomic Classification (NS-SEC) [17]. The NS-SEC was designed to take into account the nature of modern inequalities, by measuring conditions of occupations and also employment relations [17, 18]. To derive the NS-SEC, participants answered via self-completion questionnaire, questions relating to the occupation and employment status of the main-earner in their household, reporting on the main-earner's current or last main job. Individuals were assigned the category 'Not classifiable' if information was missing and as such an NS-SEC class could not be calculated. We re-coded the five-class NS-SEC version to form the three-class version which can be assumed to have a

hierarchy [17]. The classes from high to low SES represented managerial/professional, intermediate and routine/manual occupations.

Potential confounding variables of the relationship between SES, symptom severity and sickness absence included in the analysis, were age, sex, ethnicity, foreign travel in the ten days before disease onset, and urban/rural residency (based on Super Output Areas) [19].

Statistical analysis

Ordinal logistic regression was employed for the symptom severity outcome, and logistic regression for the binary absence outcome. Model parameters were estimated by maximum likelihood. For the ordinal logistic regression models, the proportional odds assumption was assessed using graphical methods [20, 21]. Generalised additive models (GAMs) were used to assess the linear relationship between the continuous age variable and the outcomes (Additional file 1). There was a linear relationship between age and the log-odds of sickness absence, therefore age was included as a continuous variable when modelling the absence outcome. The relationship between age and symptom severity was non-linear, therefore a categorical age group variable was included when modelling symptom severity.

A hierarchical approach was used for the multivariate regression modelling. Firstly, we fitted baseline models for each of our two outcomes (symptom severity and sickness absence) with age, sex and ethnicity as independent variables. Secondly, we added NS-SEC as an additional independent variable to the models and tested the improvement in model fit using generalised likelihood ratio statistics to compare nested models. Thirdly, we tested whether the inclusion of additional confounders (recent foreign travel and urban/rural residency) improved the model fit. Finally, to explore whether differences in disease severity explained any association between NS-SEC and sickness absence, we added symptom severity as a control variable to the model with sickness absence as an outcome.

Listwise deletion was used as the method of handling missing data. For the two outcomes, cases with missing data within any of the variables to be included in the models were excluded. Sensitivity analyses were performed using multiple imputation by chained equations to impute missing data values for all of the variables included in the models.

We undertook several robustness tests, repeating our analyses using alternative cut-offs for the symptom severity categories; including recurrent episodes of IID within the same individual; using cases of all ages; and stratifying results by child and adult age groups. We also examined the appropriateness of combining cases from the IID2 cohort and GP presentation studies. Analyses were conducted using R (version 3.3.1).

Results

The IID2 studies identified 1915 cases meeting our inclusion criteria of which 1270 were of school or working age and included in the sickness absence analysis (see Additional file 1 for flow diagram). Characteristics of the cases stratified by NS-SEC are shown in Table 1. Around half of cases were in managerial/professional occupations, and the vast majority were of White ethnicity (>90%). Cases in routine/manual compared to managerial/professional occupations were less likely to reside in rural areas, be female or have travelled abroad before their illness. Age and ethnicity were not associated with NS-SEC.

Symptom severity

The symptom severity score ranged from 2 to 40 and was positively skewed. The boundaries for the tertiles were: mild (score 2–9), moderate (score 10–15) and severe (score 16–40). In total, 1164 (61%) cases had complete data for the variables of interest.

The univariate associations between symptom severity and the exposures are shown in Table 2, and two nested multivariate models for symptom severity are displayed in Table 3. The addition of NS-SEC to the baseline model improved the model fit when comparing the likelihoods of the models (Likelihood ratio χ^2 31.7; $P < 0.001$). For those in routine/manual compared to managerial/professional occupations the odds of experiencing severe IID symptoms, versus mild or moderate symptoms combined, were two times greater (OR 2.2, 95%CI;1.66–2.87). The odds of experiencing severe symptoms were greater for those of Non-White compared to White ethnicity, and for those aged 15–24 years compared to 5–14 years, however these estimates were based on small numbers (43 cases were of Non-White ethnicity; 61 cases were aged 15–24 years). There was no improvement in the model fit when the variables urban/rural residency and recent foreign travel were added to the Baseline + NS-SEC model, and therefore these models are not presented.

Sickness absence

Of the 1270 cases of school or working age, 818 (64%) had complete data for the variables of interest (Additional file 1). Over half of the cases (62%) were absent from work or school following their illness. Amongst the absentees, the majority took 1–2 days sick leave (62%), and few took more than five days (8%).

The univariate associations between sickness absence and the exposures are shown in Table 2, and three nested multivariate models for sickness absence are displayed in Table 4. The addition of NS-SEC produced a better fitting model compared to the baseline model (Likelihood ratio χ^2 10.2; $P = 0.006$). Those in routine/manual compared to managerial/professional occupations had a higher odds of absence (OR 1.8, 95%CI;1.26–2.69).

Table 1 Unadjusted distribution of each variable by NS-SEC category, for the two analysis samples (IID2 study 2008–9)

	Cases ≥5 years of age (*n* = 1915)						Cases school/working age (*n* = 1270)					
	Percentage within each category of NS-SEC						Percentage within each category of NS-SEC					
	Managerial/ professional (*n* = 949)	Intermediate (*n* = 330)	Routine/manual (*n* = 337)	*p*-value[a]	All cases[b] (*n* = 1915)		Managerial/professional (*n* = 662)	Intermediate (*n* = 215)	Routine/ manual (*n* = 228)	*p*-value[a]	All cases[b] (*n* = 1270)	
Age group (years)												
5–14	12.2	9.4	9.5	0.342	10.6		17.5	14.4	14	0.337	16	
15–24	4.4	5.2	7.4		5.2		6.3	7.9	11		7.9	
25–44	24.1	22.7	23.7		22.9		34.6	34.9	35.1		34.5	
45–64	36	36.7	33.5		35.1		41.5	42.8	39.9		41.7	
65+	23.2	26.1	25.8		26.2							
Male	38	32.7	45.1	0.004	37.9		39.7	37.2	49.6	0.014	41	
Ethnicity Non-White	3.4	4.8	3.6	0.468	4		3.9	7	4.8	0.185	5.1	
Rural residence	30.6	30	19	<0.001	28.8		30.3	28.4	20.6	0.019	27.6	
Travelled before illness	14.3	10.3	7.7	0.004	11.9		15.6	13.1	7.9	0.013	12.9	
Symptom severity												
Mild	38.3	34.4	20.3	<0.001	24.5		40	36	20.7	<0.001	26.8	
Moderate	34	33	35.5		24.9		36.7	32	38		27.1	
Severe	27.7	32.6	44.2		23.4		23.3	32	41.3		22.1	
Absent work/school	55.1	54.2	63	0.028	53.4		61.4	62.7	71.6	0.023	61	

IID infectious intestinal disease, *NS-SEC* National Statistics Socioeconomic Classification, *SD* standard deviation

Figures expressed as percentages except where stated otherwise

[a]Statistical significance of relationship between NS-SEC and each variable, tested using χ^2 test

[b]Total number of cases includes those with missing NS-SEC

Missing data (%) for cases ≥5 years and cases school/working age, respectively: NS-SEC 1.6 and 1.3; Urban/rural 0.1 and 0.2; Foreign travel 0.6 and 0.3; Symptom severity 2.7 and 2.4; Absence 3.6 and 2.7

Table 2 Univariate associations for IID symptom severity and sickness absence outcomes (IID2 study 2008–9)

	Severe symptoms versus mild or moderate symptoms combined OR (95%CI)	Sickness absence versus no sickness absence OR (95%CI)[a]
	Cases with complete data ≥5 years of age (*n* = 1164)	Cases with complete data school/working age (*n* = 818)
Age group (years)		
5–14	reference	
15–24	2.88 (1.59–5.31)	
25–44	0.99 (0.68–1.45)	
45–64	0.70 (0.49–1.01)	
65+	0.60 (0.41–0.89)	
Age (years)		0.98 (0.97–0.99)
Sex		
Female	reference	reference
Male	0.92 (0.74–1.14)	0.94 (0.70–1.25)
Ethnicity		
White	reference	reference
Non-White	2.27 (1.28–4.10)	3.13 (1.38–8.41)
NS-SEC		
Managerial/ professional	reference	reference
Intermediate	1.21 (0.92–1.61)	1.13 (0.78–1.66)
Routine/manual	2.18 (1.67–2.86)	1.77 (1.22–2.58)
Residence		
Urban	reference	reference
Rural	0.82 (0.65–1.03)	0.98 (0.71–1.34)
Travelled before illness		
No	reference	reference
Yes	1.21 (0.88–1.66)	0.66 (0.44–0.99)
Symptom severity		
Mild		reference
Moderate		3.88 (2.75–5.51)
Severe		5.99 (4.07–8.95)

CI confidence interval, *IID* infectious intestinal disease, *NS-SEC* National Statistics Socioeconomic Classification, *OR* odds ratio
[a]Since the absence outcome was common, the odds ratios should not be interpreted as risk ratios

When symptom severity was added to this model the odds of absence for those in routine/manual compared to managerial/professional occupations was attenuated and rendered non-significant (OR 1.4, 95%CI;0.92–2.07). There was a dose-response relationship between symptom severity and the odds of absence. Those with severe compared to mild symptoms had five times the odds of absence (OR 5.3, 95%CI;3.54–7.93). Again, there was no improvement in the model fit when the variables urban/rural residency and recent foreign travel were added to the Baseline + NS-SEC model, and therefore these models are not presented.

Table 3 Multivariate models for severe IID symptoms, versus mild or moderate symptoms combined for cases ≥5 years of age (IID2 study 2008–9)

	Baseline model OR (95%CI)	Baseline + NS-SEC OR (95%CI)
Age group (years)		
5–14	reference	reference
15–24	3.01 (1.66–5.57)	2.70 (1.48–5.02)
25–44	1.02 (0.69–1.50)	0.96 (0.65–1.42)
45–64	0.73 (0.51–1.06)	0.69 (0.47–1.00)
65+	0.64 (0.43–0.95)	0.60 (0.41–0.90)
Sex		
Female	reference	reference
Male	0.95 (0.77–1.19)	0.90 (0.72–1.13)
Ethnicity		
White	reference	reference
Non-White	2.11 (1.18–3.83)	2.03 (1.14–3.70)
NS-SEC		
Managerial/professional		reference
Intermediate		1.21 (0.91–1.61)
Routine/manual		2.18 (1.66–2.87)
Log-likelihood	−1255.1	−1239.3
Deviance	2510.2	2478.6
AIC	2526.2	2498.6
BIC	2566.7	2549.2
Number	1164	1164

AIC Akaike information criterion, *BIC* Bayesian information criterion, *CI* confidence interval, *IID* infectious intestinal disease, *NS-SEC* National Statistics Socioeconomic Classification, *OR* odds ratio

Sensitivity analyses

We undertook several robustness tests. Similar results to those reported were observed when analyses were conducted with recurrent episodes of IID included with clustering at the individual level accounted for using mixed-effects models, and when the boundaries of the symptom severity categories were changed so that there was an equal 12 point severity score difference within each category (data not shown). Results from multiply imputed datasets, and analyses involving cases of all ages and stratified results by child and adult age groups, also confirmed those from the main analyses (Additional file 1), however ethnicity was not associated with symptom severity when analyses were performed using the imputed datasets. Additionally, comparable associations were found when investigating predictive factors for the duration of absence among absentees (Additional file 1). Lastly, the

Table 4 Multivariate models for sickness absence due to IID for cases of school/working age (IID2 study 2008–9)

	Baseline model	Baseline + NS-SEC	Baseline + NS-SEC + Severity
	OR (95%CI)[a]	OR (95%CI)[a]	OR (95%CI)[a]
Age (years)	0.98 (0.98–0.99)	0.98 (0.98–0.99)	0.99 (0.98–1.00)
Sex			
Female	reference	reference	reference
Male	0.95 (0.71–1.28)	0.91 (0.68–1.22)	0.92 (0.67–1.26)
Ethnicity			
White	reference	reference	reference
Non-White	2.66 (1.16–7.22)	2.58 (1.12–7.00)	1.91 (0.80–5.31)
NS-SEC			
Managerial/ professional		reference	reference
Intermediate		1.13 (0.77–1.66)	1.05 (0.70–1.59)
Routine/manual		1.83 (1.26–2.69)	1.38 (0.92–2.07)
Symptom severity			
Mild			reference
Moderate			3.60 (2.54–5.14)
Severe			5.27 (3.54–7.93)
Log-likelihood	−531.0	−525.9	−482.4
Deviance	1062.0	1051.9	964.9
AIC	1070.0	1063.9	980.9
BIC	1088.9	1092.1	1018.5
Number	818	818	818

AIC Akaike information criterion, *BIC* Bayesian information criterion, *CI* confidence interval, *IID* infectious intestinal disease, *NS-SEC* National Statistics Socioeconomic Classification, *OR* odds ratio
[a]Since the absence outcome was common, the odds ratios should not be interpreted as risk ratios

appropriateness of combining cases from the IID2 component studies was supported by analyses indicating the relationships between NS-SEC and the outcomes were not significantly different between the cohort and GP presentation studies (Additional file 1).

Discussion

We analysed data from the largest population-based survey of IID conducted in the UK, and found that IID cases of lower SES compared to high were more likely to experience severe symptoms, and were more likely to be absent from work or school. The association between SES and sickness absence was largely explained by greater symptom severity amongst the more disadvantaged groups.

Our findings are comparable to those of other studies that have analysed measures of IID severity and SES, however these studies are sparse in number, and have tended to focus on children under five years of age. Our findings suggest that the association between SES and IID severity is true for the whole (all age) population, not just for young children. We identified one British study which analysed data from the population-based Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC), to assess predictive factors for the duration of diarrhoeal episodes in children less than six months of age [22]. The authors found that infants living in rented versus mortgaged/owned accommodation (a suggested indicator of SES) had greater odds of experiencing diarrhoea for six or more days. However, this association became non-significant after adjustment for duration of breast feeding, with longer spells of breast feeding providing protection against prolonged diarrhoea.

Whilst very few cases were admitted to hospital in our sample (<1%), our findings are somewhat similar to those of studies conducted in hospital settings. At this severe end of the disease spectrum, one UK-based study found low SES was associated with longer time to discharge for children hospitalised with gastroenteritis in univariate analysis [23]. Similarly, among American children less than five years of age hospitalised with gastroenteritis, those enrolled in Medicaid (a proxy measure for low SES) experienced longer average length of stay, compared to children not enrolled, when no other factors were taken into consideration [24]. In contrast, multivariate analysis revealed that education level and income were not related to length of stay for Canadian children less than five years of age hospitalised with rotavirus gastroenteritis, whereas regularly seeing a physician for a medical condition was associated with longer hospital stays [25].

These findings might suggest that the association between SES and IID severity could be mediated by socially patterned factors that impair immune response, such as lack of breast feeding in infancy and multimorbidity [26, 27], both of which are more prevalent among lower socioeconomic groups [28, 29]. Additional biologically plausible mechanisms which might help to explain a greater burden of severe IID in lower socioeconomic groups, but are as yet to be substantiated in this context, include increased levels of chronic stress, smoking, and nutritional deficiencies, all of which display social gradients and are associated with immune system compromise [30–34]. The potential mediating role of immune suppressing variables on the relationship between SES and symptom severity warrants further investigation.

We found IID cases of lower SES compared to high had greater odds of sickness absence due to IID, and this was largely explained by greater symptom severity amongst cases of lower SES. In a cohort of UK civil servants, age adjusted rates of sickness absence due to gastroenteritis, were over six and four times higher for men and women respectively, in lower employment grades compared to high [12]. Conversely, self-reported sickness absence for

gastroenteritis in a cohort of Dutch employees was unrelated to education level in univariate analysis [13]. These conflicting findings may, in part, be due to the different populations studied, since our age, sex and ethnicity adjusted results for absence were akin to those observed in the UK-based study of civil servants [12]. However, neither study investigated the role of symptom severity, which was identified as an important mediator of the relationship between SES and sickness absence in our analysis.

There are several limitations to this analysis. The validity of our results depended upon the unbiased and accurate self-reporting of symptoms and sickness absence among cases. If those of lower SES perceived their symptoms differently to those of higher SES, which has been observed in studies investigating perceptions of pain across socioeconomic groups [35, 36], our results could be a mere artefact of the severity measurement. Nonetheless, the variables used to derive the symptom severity score in our study were related to the presence and duration of symptoms, which are rather more objective measures of severity compared to, for example, a subjective rating of symptom severity from mild to severe.

There was a large amount of missing data, particularly within the NS-SEC and symptom severity variables (Table 1). Listwise deletion as a method of handling missing data can produce unbiased estimates when data are missing completely at random [37]. However, the odds of whether data were missing or not within the NS-SEC and symptom severity variables, were associated with other variables within the dataset, supporting the idea that missing data were missing at random, rather than missing completely at random. Sensitivity analyses were therefore performed using multiple imputation by chained equations to impute missing data values (Additional file 1). Results from multiply imputed datasets confirmed those from the main analyses, suggesting that any bias resulting from the use of listwise deletion, was minimal. Ethnicity however was not associated with symptom severity when analyses were performed using the imputed datasets.

Cases identified in the IID2 cohort and GP presentation studies were combined for this analysis. Individuals in managerial/professional occupations, those aged 55+ years and those of White ethnicity were over-represented in the cohort study compared to the UK population, and individuals in intermediate and routine/manual occupations and those aged 15–24 years in particular were under-represented [15]. Under-representation of lower socioeconomic groups is commonplace in population-based surveys [38], and could limit the external validity of our findings. Nevertheless, the internal validity of our findings should remain unaffected. It is possible that if non-participation or the design of the studies resulted in the under-representation of cases of lower SES who experienced milder symptoms, we may have overestimated the association between low SES and severe symptoms. However, within the cohort study this is unlikely as cases were captured prospectively. The GP presentation study may have been more prone to selection bias, since cases with more severe symptoms and those of lower SES may be more likely to present to their GP for an episode of IID [3], however as shown in Additional file 1, the relationship between NS-SEC and symptom severity was not significantly different between the cohort and GP presentation studies.

There is the potential for different pathogens to infect people of different SES, for example *Listeria* and norovirus have been associated with low SES in some studies [39, 40]. Unfortunately, we were unable to explore the role of pathogen type on the association between SES and symptom severity because for around 58% of the sampled cases no pathogen was identified [15]. The impact of pathogen type on the association between SES and symptom severity is unknown, however the severity of illness likely depends not only on the infecting pathogen but also on host factors and the dose to which the host is exposed [41]. The relationship between SES, pathogen type and IID symptom severity could be explored using a larger sample of cases, since for the majority a pathogen will not be identified.

Finally, the IID2 study also contained a retrospective telephone survey which gave higher IID incidence estimates compared to the IID2 cohort study [15], however we were unable to repeat our analyses with cases identified in the telephone survey because NS-SEC information was not collected. We were also unable to assess inequalities in sickness absence amongst those providing care for IID cases (caregiver informative was not collected) however this may be an interesting avenue for further research.

Conclusions

Our study sheds new light into an under-researched area and indicates that the consequences of having an IID may be unequally shared across socioeconomic groups. These consequences are potentially serious. Loss of working days due to sickness can have important economic consequences and these are likely to be more severe for more disadvantaged groups who might receive less adequate compensation from their employer. Loss of days from school can affect educational attainment [42], suggesting that the unequal effects of IID could exacerbate educational inequalities. Actions that reduce the risk of acquiring IID are unlikely to sufficiently address these inequalities; public health interventions also need to reduce their unequal consequences. Further research is required to understand the mechanisms explaining greater severity of illness in disadvantaged groups, and to identify ways to minimise the differential impact of IID on sickness absence.

Additional file

Additional file 1: Supplementary material and sensitivity analyses. (PDF 387 kb)

Abbreviations

ALSPAC: Avon Longitudinal Study of Pregnancy and Childhood; GAM: Generalised additive model; GP: General practice; IID: Infectious intestinal disease; NS-SEC: National Statistics Socioeconomic Classification; SES: Socioeconomic status; UK: United Kingdom

Acknowledgements

The authors thank Dr. John Harris and Dr. Peng Yin for their statistical advice.

Funding

This work was supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections [grant number NIHR HPRU 2012–10,038] at University of Liverpool in partnership with Public Health England (PHE), in collaboration with University of East Anglia, University of Oxford and the Institute of Food Research. Tanith Rose is based at the University of Liverpool. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or PHE.

Availability of data and materials

The datasets analysed during the current study are available in the UK Data Service repository, https://discover.ukdataservice.ac.uk/catalogue/?sn=7820&type=Data%20catalogue.

Authors' contributions

All authors contributed to the conception and design of the study. TR performed the analyses with guidance from DTR and BB. TR drafted the manuscript which was revised critically by DTR, BB, MV, JH, NL, SOB and MW. All authors approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The IID2 study was granted ethical approval by the North West Research Ethics Committee (07/MRE08/5). Participants gave written informed consent for their anonymised data to be used for future analyses.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1]NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, UK. [2]Department of Public Health and Policy, University of Liverpool, Liverpool, UK. [3]National Infection Service, Public Health England, London/Birmingham, UK. [4]Health Economics Research Centre, University of Oxford, Oxford, UK. [5]Department of Public Health and Policy Institute of Psychology, Health and Society, University of Liverpool, Whelan Building, Liverpool L69 3GB, UK.

Received: 14 March 2017 **Accepted:** 15 June 2017
Published online: 23 June 2017

References

1. Tam CC, Rodrigues LC, Viviani L, et al. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. Gut. 2012;61:69–77.
2. Food Standards Agency. The second study of infectious intestinal disease in the community (IID2 Study). 2016. https://www.food.gov.uk/science/research/foodborneillness/b14programme/b14projlist/b18021 Accessed 8 Nov 16.
3. Tam CC, Rodrigues LC, O'Brien SJ. The study of infectious intestinal disease in England: what risk factors for presentation to general practice tell us about potential for selection bias in case-control studies of reported cases of diarrhoea. Int J Epidemiol. 2003;32:99–105.
4. Teschke K, Bellack N, Shen H, et al. Water and sewage systems, socio-demographics, and duration of residence associated with endemic intestinal infectious diseases: a cohort study. BMC Public Health. 2010;10:767.
5. Pockett RD, Adlard N, Carroll S, Rajoriya F. Paediatric hospital admissions for rotavirus gastroenteritis and infectious gastroenteritis of all causes in England: an analysis of correlation with deprivation. Curr Med Res Opin. 2011;27:777–84.
6. Olowokure B, Hawker J, Weinberg J, Gill N, Sufi F. Deprivation and hospital admission for infectious intestinal diseases. Lancet. 1999;353(9155):807–8.
7. Dennehy PH, Cortese MM, Bégué RE, et al. A case-control study to determine risk factors for hospitalization for rotavirus gastroenteritis in U.S. children. Pediatr Infect Dis J. 2006;25(12):1123–31.
8. Biering-Sørensen S, Søndergaard G, Vitting Andersen K, et al. Time trends in socio-economic factors and risk of hospitalisation with infectious diseases in pre-school children 1985–2004: a Danish register-based study. Paediatr Perinat Epidemiol. 2012;26(3):226–35.
9. Wilking H, Höhle M, Velasco E, et al. Ecological analysis of social risk factors for rotavirus infections in berlin, Germany, 2007-2009. Int J Health Geogr. 2012;11:37.
10. Kristensen TR, Jensen SM, Kreiner S, Mikkelsen S. Socioeconomic status and duration and pattern of sickness absence. A 1-year follow-up study of 2331 hospital employees. BMC Public Health. 2010;10:643.
11. North F, Syme SL, Feeney A, et al. Explaining socioeconomic differences in sickness absence: the Whitehall II study. BMJ. 1993;306(6874):361–6.
12. Feeney A, North F, Head J, Canner R, Marmot M. Socioeconomic and sex differentials in reason for sickness absence from the Whitehall II study. Occup Environ Med. 1998;55:91–8.
13. Mohren DC, Swaen GM, Kant I, van Schayck CP, Galama JM. Fatigue and job stress as predictors for sickness absence during common infections. Int J Behav Med. 2005;12(1):11–20.
14. O'Brien SJ, Rait G, Hunter PR, et al. Methods for determining disease burden and calibrating national surveillance data in the United Kingdom: the second study of infectious intestinal disease in the community (IID2 study). BMC Med Res Methodol. 2010;10:39.
15. Tam C, Viviani L, Adak B, et al. The second study of infectious intestinal disease in the community (IID2 Study): Final report. Food Standards Agency, 2012. https://www.food.gov.uk/science/research/foodborneillness/b14programme/b14projlist/b18021 Accessed 8 Nov 2016.
16. Office for National Statistics. Pension trends. Basingstoke: Palgrave Macmillan; 2005.
17. Office for National Statistics. The National Statistics Socio-economic Classification (NS-SEC). https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010 Accessed 8 Nov 2016.
18. Rose D, Pevalin DJ, O'Reilly K, editors. The National Statistics Socio-Economic Classification: origins, development and use. Basingstoke: Palgrave Macmillan; 2005.
19. Office for National Statistics. Super Output Area (SOA). http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas%2D-soas-/index.html Accessed 8 Nov 2016.
20. Harrell FE, editor. Regression Modeling Strategies. New York: Springer-Verlag; 2001.
21. UCLA: Statistical Consulting Group. R Data Analysis Examples: Ordinal Logistic Regression. https://stats.idre.ucla.edu/r/dae/ordinal-logistic-regression/. Accessed 19 June 2017.
22. Baker D, Taylor H, Henderson J. ALSPAC study team. Inequality in infant morbidity: causes and consequences in England in the 1990s. J Epidemiol Community Health. 1998;52(7):451–8.
23. Conway SP, Phillips RR, Panday S. Admission to hospital with gastroenteritis. Arch Dis Child. 1990;65:579–84.
24. Ma L, El Khoury AC, Itzler RF. The burden of rotavirus hospitalizations among Medicaid and non-Medicaid children younger than 5 years old. Am J Public Health. 2009;99(Suppl 2):S398–404.
25. Ford-Jones EL, Wang E, Petric M, Corey P, Moineddin R, Fearon M. Hospitalization for community-acquired, rotavirus-associated diarrhea: a prospective, longitudinal, population-based study during the seasonal outbreak. Arch Pediatr Adolesc Med. 2000;154(6):578–85.

26. Jackson KM, Nazar AM. Breastfeeding, the immune response, and long-term health. J Am Osteopath Assoc. 2006;106(4):203–7.
27. Castle SC, Uyemura K, Rafi A, Akande O, Makinodan T. Comorbidity is a better predictor of impaired immunity than chronological age in older adults. J Am Geriatr Soc. 2005;53(9):1565–9.
28. Oakley LL, Renfrew MJ, Kurinczuk JJ, Quigley MA. Factors associated with breastfeeding in England: an analysis by primary care trust. BMJ Open. 2013;3(6):e002765.
29. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet. 2012;380(9836):37–43.
30. Cohen S, Doyle WJ, Baum A. Socioeconomic status is associated with stress hormones. Psychosom Med. 2006;68(3):414–20.
31. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull. 2004; 130(4):601–30.
32. Stämpfli MR, Anderson GP. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. Nat Rev Immunol. 2009; 9(5):377–84.
33. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr. 2008;87(5):1107–17.
34. Lund BM, O'Brien SJ. The occurrence and prevention of foodborne disease in vulnerable people. Foodborne Pathog Dis. 2011;8(9):961–73.
35. Dorner TE, Muckenhuber J, Stronegger WJ, Rásky E, Gustorff B, Freidl W. The impact of socio-economic status on pain and the perception of disability due to pain. Eur J Pain. 2011;15(1):103–9.
36. Mijković A, Stipčić A, Braš M, et al. Is experimentally induced pain associated with socioeconomic status? Do poor people hurt more? Med Sci Monit. 2014;20:1232–8.
37. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol. 2013;64(5):402–6.
38. Lorant V, Demarest S, Miermans PJ, Van Oyen H. Survey error in measuring socio-economic risk factors of health status: a comparison of a survey and a census. Int J Epidemiol. 2007;36(6):1292–9.
39. Phillips G, Tam CC, Rodrigues LC, Lopman B. Risk factors for symptomatic and asymptomatic norovirus infection in the community. Epidemiology & Infection. 2011;139(11):1676–86.
40. Gillespie IA, Mook P, Little CL, et al. Human listeriosis in England, 2001-2007: association with neighbourhood deprivation. Euro Surveillance. 2010;15(27):7–16.
41. O'Brien SJ, Halder SL. GI epidemiology: infection epidemiology and acute gastrointestinal infections. Aliment Pharmacol Ther. 2007;25(6):669–74.
42. Department for Education. The link between absence and attainment at KS2 and KS4: 2012/13 academic year. 2015. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/412638/The_link_between_absence_and_attainment_at_KS2_and_KS4.pdf Accessed 8 Nov 2016.

Appendix 6.5: Relationship between socioeconomic status and gastrointestinal infections in developed countries: A systematic review and meta-analysis

PLOS ONE

RESEARCH ARTICLE

Relationship between socioeconomic status and gastrointestinal infections in developed countries: A systematic review and meta-analysis

Natalie L. Adams[1,2,3☯]*, Tanith C. Rose[1,2☯], Jeremy Hawker[1,3], Mara Violato[1,4], Sarah J. O'Brien[1,2], Benjamin Barr[1,2], Victoria J. K. Howard[2], Margaret Whitehead[1,2‡], Ross Harris[3‡], David C. Taylor-Robinson[1,2‡]

1 NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, United Kingdom, **2** Department of Public Health and Policy, University of Liverpool, Liverpool, United Kingdom, **3** National Infection Service, Public Health England, London, United Kingdom, **4** Health Economics Research Centre, University of Oxford, Oxford, United Kingdom

☯ These authors contributed equally to this work.
‡ MW, Rh, and DCTR are joint senior authors on this work.
* Natalie.Adams@phe.gov.uk

OPEN ACCESS

Citation: Adams NL, Rose TC, Hawker J, Violato M, O'Brien SJ, Barr B, et al. (2018) Relationship between socioeconomic status and gastrointestinal infections in developed countries: A systematic review and meta-analysis. PLoS ONE 13(1): e0191633. https://doi.org/10.1371/journal.pone.0191633

Editor: Mussaret Bano Zaidi, Hospital General O'Horan, MEXICO

Received: February 13, 2017

Accepted: December 26, 2017

Published: January 23, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections [grant number NIHR HPRU 2012-10038] at University of Liverpool in partnership with Public Health England (PHE), in collaboration with University of East Anglia, University of Oxford

Abstract

Background

The association between socioeconomic status (SES) and health is well-documented; however limited evidence on the relationship between SES and gastrointestinal (GI) infections exists, with published studies producing conflicting results. This systematic review aimed to assess the association between SES and GI infection risk, and explore possible sources of heterogeneity in effect estimates reported in the literature.

Methods

MEDLINE, Scopus, Web of Science and grey literature were searched from 1980 to October 2015 for studies reporting an association between GI infections and SES in a representative population sample from a member-country of the Organisation for Economic Co-operation and Development. Harvest plots and meta-regression were used to investigate potential sources of heterogeneity such as age; level of SES variable; GI infection measurement; and predominant mode of transmission. The protocol was registered on PROSPERO: CRD42015027231.

Results

In total, 6021 studies were identified; 102 met the inclusion criteria. Age was identified as the only statistically significant potential effect modifier of the association between SES and GI infection risk. For children, GI infection risk was higher for those of lower SES versus high (RR 1.51, 95% CI;1.26–1.83), but there was no association for adults (RR 0.79, 95% CI;0.58–1.06). In univariate analysis, the increased risk comparing low and high SES groups

and the Institute of Food Research. Natalie Adams is a PhD student based at the University of Liverpool and Public Health England. Tanith Rose is a PhD student based at the University of Liverpool. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

Competing interests: The authors have declared that no competing interests exist.

was significantly higher for pathogens spread by person-to-person transmission, but lower for environmental pathogens, as compared to foodborne pathogens.

Conclusions

Disadvantaged children, but not adults, have greater risk of GI infection compared to their more advantaged counterparts. There was high heterogeneity and many studies were of low quality. More high quality studies are needed to investigate the association between SES and GI infection risk, and future research should stratify analyses by age and pathogen type. Gaining further insight into this relationship will help inform policies to reduce inequalities in GI illness in children.

Introduction

Gastrointestinal (GI) infections are common. Estimates suggest around 25% of people in the UK suffer an episode of infectious intestinal disease (IID) per year.[1,2] Several risk factors for GI infection have been investigated in the literature, including environmental risk factors such as population density and rurality, and individual-level risk factors such as sex and ethnicity.[3–6] Age has been identified as an important risk factor for GI infection, with the youngest age groups most at risk.[3,7,8] Yet for some potential risk factors such as socioeconomic status (SES), the association has been less clear. Inconsistent results have been observed among studies, with some reporting higher rates of GI infections among lower socioeconomic groups[8–10] and others observing the opposite.[11,12] Lower risk of *Campylobacter*, *Cryptosporidium* and norovirus has been identified in more disadvantaged areas.[5, 13–16] In contrast, incidence was found to be higher in more disadvantaged areas for listeria and rotavirus.[17,18] Disadvantaged children were found to have higher risk of non-typhoidal *Salmonella*, rotavirus and norovirus.[19–21] A systematic review exploring the impact of SES on laboratory confirmed foodborne illness in developed countries identified 16 studies across four pathogens with mixed results, differing by pathogen.[22] These results demonstrate the ongoing disagreements within this area of research.

A systematic review was warranted to summarise and understand the contradictory findings observed in the literature and explore the relationship for GI pathogens which are predominantly transmitted via person-to-person, waterborne and environmental routes as well as the foodborne route which has been studied previously.[22] Our review aimed to explore the relationship between SES and a full range of GI infections to assess the magnitude and direction of the association, and suggest possible explanations for any observed differences.

Methods

Search strategy and selection criteria

We conducted a systematic review and meta-analysis. The exposure of interest was SES, measured at the individual or aggregate level by income, education, occupation, employment or area-level deprivation. The primary outcome of interest was incidence/prevalence of any symptomatic GI infection, including syndromic definitions of GI infections without a laboratory diagnosis. These were included as various socioeconomic or healthcare seeking behavioural factors could influence whether an individual is diagnosed with a GI infection.

The methods for this study have been described in detail in the study protocol (https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-016-0187-7). [23] Full details of the inclusion/exclusion criteria are reported in Table 1.

We included studies that analysed the same individuals if they analysed different exposures or outcomes. Where more than one study analysed the same individuals using the same outcomes and exposures, only one study was included based on the study with the greatest focus and amount of information on the relationship between SES and GI infections.

Three search strategies were used to identify relevant literature. Electronic searching of three databases was performed: MEDLINE (Ovid); Scopus and Web of Science Core Collection. Search terms were piloted prior to selection and comprised specific GI infections and symptom-based terms, socioeconomic and inequality terms and developed countries of interest (Table A in S1 File). The results were restricted to publications that used data collected after 1980, to ensure the results were as relevant as possible to the present day.

Secondly, we searched the reference lists of studies selected for inclusion in the review to identify relevant articles that were not captured via electronic searching. Finally, grey literature was searched by entering the terms "gastrointestinal infection", "gastroenteritis", "diarrhoea", "diarrhea", "socioeconomic", "social class", "income", and "deprivation" into the Google internet search engine and Google Scholar search application. The first 100 results from each search were screened for inclusion.

Titles and abstracts of the publications were screened independently by two authors (NA and TR) to ensure consistency in the application of inclusion/exclusion criteria. Discrepancies were discussed and resolved through a consensus process. The full text for studies deemed relevant after title and abstract screening, were sought and screened in the same way.

Data analysis

Data were extracted into a standardised Excel spreadsheet by one reviewer and were checked by the second reviewer. Data extracted included: aim/hypothesis; study design; level of analysis; country; sample size; age; measurement of GI infection; measurement of SES; covariates

Table 1. Inclusion and exclusion criteria.

Inclusion criteria
1. Studies quantitatively measuring the prevalence or incidence of any symptomatic gastrointestinal infection in a representative population sample
2. Studies quantitatively measuring socioeconomic status at an individual or aggregate level by occupation, income, education, employment or area deprivation
3. Studies reporting a quantitative association between the first two inclusion criteria i.e. reporting an association between gastrointestinal infection and socioeconomic status
4. Studies written or translated into English language
5. Studies reporting on human subjects
6. Subjects selected from the populations of countries that are members of the Organisation for Economic Co-operation and Development (OECD), reporting data after 1980 or the date that they became a member of the OECD
7. Studies reporting on data collected after 1980
8. Observational studies
Exclusion criteria
1. Unrepresentative population sample
2. Outbreak reports
3. Studies analysing travel related cases only
4. Review studies
5. Case reports

https://doi.org/10.1371/journal.pone.0191633.t001

and results. For studies where quantitative data were reported in text form only, authors were contacted to obtain the relevant data.

Risk of bias and quality assessment of the studies were conducted by the review team independently and then reconciled. The Liverpool University Quality Assessment Tool (LQAT) was used for this review, which allowed for the methodological quality of the studies to be assessed using a tool specific to each study design.[24] LQAT incorporates a star rating system to assess and quantify absence of bias, misclassification and confounding. It has been used in a number of other reviews[25,26] and has been independently evaluated against other quality assessment tools.[27] Discrepancies between reviewers in the quality assessment of the studies were discussed, re-examined and resolved.

Both harvest plots and meta-analysis were used to synthesise the data. Harvest plots were created to display and summarise the results of the studies and the subgrouping graphically.[28] Each reported association between SES and GI infection was represented by a single bar. The height of the bars were used to indicate the quality score of the studies from which the associations arose, so that the strength of the evidence could be determined, and greater weight given to conclusions drawn from the most methodologically robust and reliable studies. An inclusive strategy was used for the harvest plots, allowing all studies to be captured graphically, irrespective of whether quantitative estimates were provided. The findings from the harvest plots were used to inform the methods used in the meta-analysis and lead to potential explanations for the contrasting findings observed in the literature.

Subgroup analyses were performed on study design factors and potential moderating factors of the relationship identified a priori,[23] including: pathogen type (based on predominant mode of transmission–foodborne (*Campylobacter*, *Salmonella*, *Yersinia enterocolitica*); person-to-person (viral GI infections, *Shigella*); waterborne (*Giardia*, *Cryptosporidium*); environmental (Shiga toxin-producing *Escherichia coli* [STEC]); age; country (based on climate and level of development); methods used to sample GI infection cases; methods used to measure SES; and level of analysis (aggregate or individual). Predominant mode of transmission for each pathogen was assigned following consultation with experts in GI pathogens and based on available literature, however it is noted that for some pathogens, most notably STEC, there are multiple transmission routes. The Human Development Index[29] was used to classify the countries by relative level of development, and climate zones were assigned based on the Köppen system.[30] Separate tables and harvest plots were created for each subgroup, detailing the number of studies finding a positive, negative and no association, across the categories of the subgroup.

Meta-analyses were conducted in R (version 3.3.1) using an inverse variance random-effects model on combined results. Where necessary, standard methods were used to calculate the risk ratios and confidence intervals.[31] Where studies analysed the same cases, or provided numerous estimates for the relationship between SES and GI infection, only one estimate was retained in the meta-analysis to avoid the double counting of cases. For example, where studies provided estimates for numerous SES measures, the most commonly used SES measure across all of the studies (education level) was chosen if available. Estimates that were adjusted for potential confounding variables, such as age and sex, were chosen over univariate estimates. Eleven studies provided more than one estimate but the cases used for each estimate were considered independent of each other, so all estimates were included in the meta-analysis. Studies that belonged to this category included those with estimates for children and adults, and estimates for different pathogens since it was assumed that it would be unlikely for a case to be infected with more than one pathogen. However, a potential issue when including multiple estimates from single studies in random-effects meta-analyses, is the within-study variability of the different estimates would be treated as between-study variability, therefore studies

with multiple estimates would have received a disproportionately high weight in the pooled estimate. Therefore fixed-effect meta-analyses were used to combine estimates from the same study, allowing these pooled estimates to be combined with the remaining studies using random-effects meta-analysis.[31]

Statistical heterogeneity was assessed by applying the I^2 statistic with values of 30 to 60%, 50 to 90% and 75 to 100% used to denote moderate, substantial and considerable levels of heterogeneity, respectively.[31] Random-effects meta-regression[32,33] and subgroup meta-analyses were conducted to investigate potential moderating factors of the relationship between SES and GI infections, guided by the harvest plot findings. Sensitivity analysis on the basis of study quality was conducted to explore the robustness of the meta-analysis. Small study effects, which can be viewed as an indication of publication bias, were assessed using a funnel plot.

Results

Following duplicate removal, the database search identified 6021 citations, and 344 were full-text screened. Of these, 102 were regarded as eligible for inclusion in the review and 77 were eligible for inclusion in the meta-analysis (Fig 1). Table 2 shows the summary characteristics of the included studies. The majority of studies were conducted in Europe, had ecological study designs, used laboratory records to identify GI infection cases, and did not stratify by age. Education level was identified as the most commonly used measure of SES across the studies. Full details of the studies can be found in Table B in S1 File.

The majority of the studies were graded as low quality (n = 56). Of these there were four cross-sectional, 35 ecological, eight cohort and nine case-control studies. Twenty-seven studies were graded as being of medium quality, including seven cross-sectional, four ecological, four cohort and 12 case-control studies. Finally, 19 studies were graded as high quality, seven cross-sectional, four ecological, four cohort and four case-control studies.

Fig 2 shows the harvest plot for GI infection by SES, stratified by age, method of identifying GI infection cases, and SES measure. Similarly, a harvest plot stratified by age, pathogen transmission route and SES measure is presented in Fig A in S1 File. Of the 102 studies included, there were 103 point estimates for the association between SES and GI infection risk for adults or children specifically, and these point estimates were represented graphically as bars in the harvest plot (Fig 2). In the harvest plots, each bar represents one study. The height of the bar represents the quality of the study. Studies are classed into those showing lower risk in disadvantaged individuals/areas, no association or higher risk in disadvantaged individuals/areas.

The harvest plot (Fig 2) illustrates that the relationship between SES and GI infection varied with age.

The results for children by method of GI data collection are presented in the upper half of Fig 2. There was a clear social patterning for children in the reviewed studies, showing higher risk of GI infection in disadvantaged children or no association between GI infection and SES; although most studies were of low quality. With the exception of a small number of laboratory record studies, none of the studies found a lower risk of GI infection in disadvantaged children. The Harvest Plot in Fig 2 also shows that there were gaps in the literature using GP presentation to explore the relationship between GI infection and SES.

The results for adults by method of GI data collection are presented in the lower half of Fig 2. The pattern for adults different from that for children, with most studies weighted towards lower risk of GI infection in disadvantaged adults or no association. There were far fewer studies exploring the association between GI infection and SES in adults and notable gaps in studies exploring the association using hospitalisation or GP presentation data and only low quality studies using hospitalisation data.

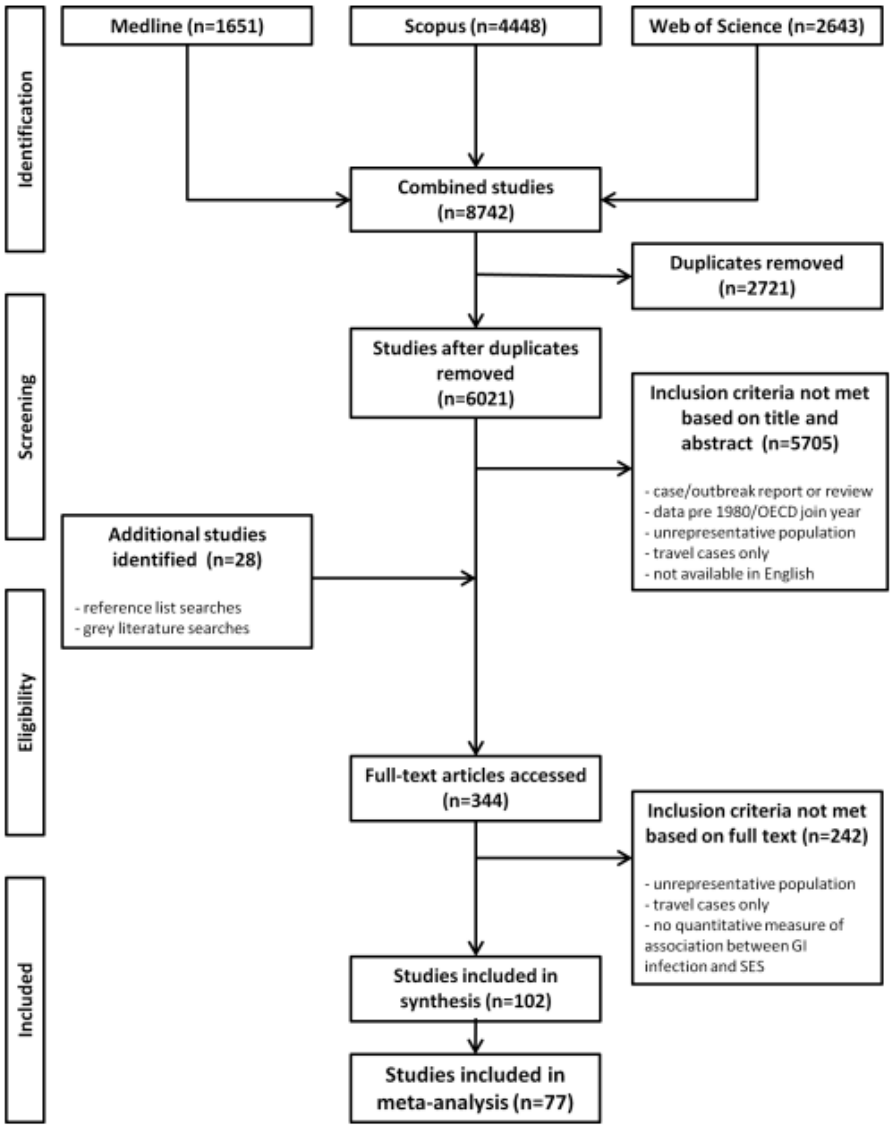


Fig 1. Flow diagram of studies included in the systematic review and meta-analysis.

https://doi.org/10.1371/journal.pone.0191633.g001

Table 2. Characteristics of included studies.

Study characteristics	Studies* (number)
Total	102
Year of publication	
Before 2000	17
2000–2005	15
2006–2010	38
After 2010	32
Level of analysis	
Individual	59
Area	43
Region	
Asia	3
Europe	49
North America	34
Oceania	16
Sample size	
<200	3
200–1000	25
1001–5000	15
5001–10000	9
10001–100000	5
>100000	45
Age category	
Children (<18 years old)	27
Adults	8
Mixed	61
Not stated	6
Gastrointestinal infection outcome	
Acute GI infection (syndromic)	41
Campylobacteriosis	20
Cryptosporidiosis	4
Giardiasis	3
Hepatitis A	3
Listeriosis	1
Norovirus	1
Rotavirus	3
Salmonellosis	8
Shigellosis	3
Shiga toxin-producing *E. coli* infection	4
Yersinia enterocolitica	1
Multiple pathogens	10
Gastrointestinal infection measure	
Population-based survey	30
General practice (GP) presentation	5
Hospital admission	13
Laboratory records	52
Multiple measures	2
Socioeconomic status measure	

(*Continued*)

Table 2. (Continued)

Study characteristics	Studies* (number)
Deprivation	17
Education	22
Employment	7
Income	10
Occupation	8
Social class	10
Multiple measures	28
Study design	
Case-control	25
Cohort	16
Cross-sectional	18
Ecological	43
Quality	
High	19
Medium	27
Low	56

*See S1 File for full overview of included studies.

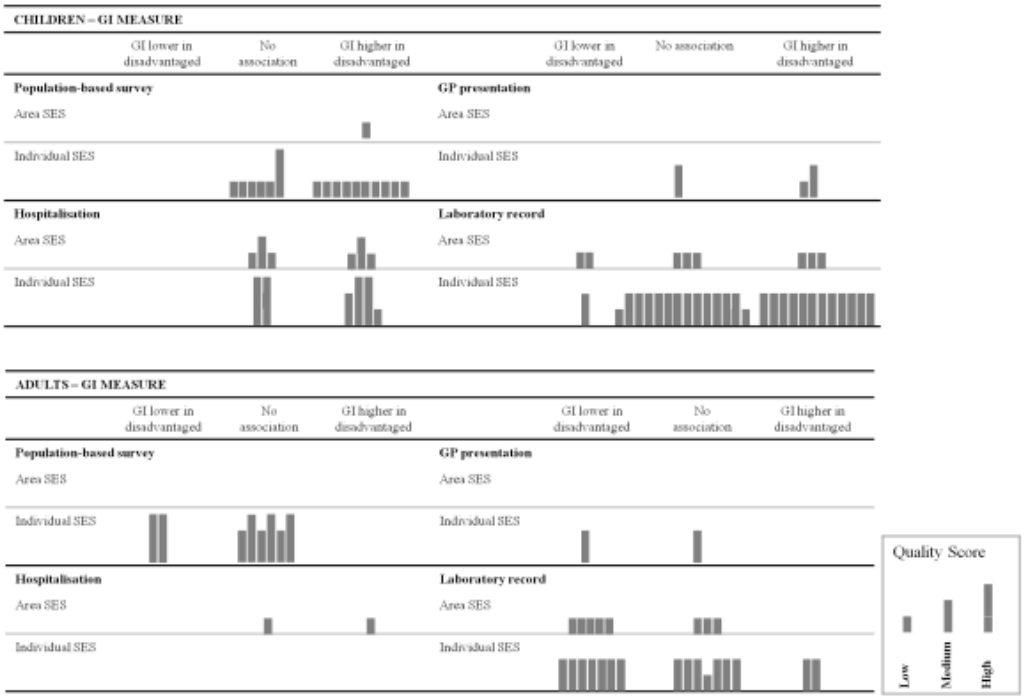
https://doi.org/10.1371/journal.pone.0191633.t002

Amongst the point estimates based on cases with a laboratory report, pathogens were grouped into predominant mode of transmission (displayed in Figure A in S1 File). There was no clear modifying role of pathogen type (based on the predominant route of transmission) on the relationship between SES and GI infection risk, although there were some differences by age. For children (upper half of Figure A in S1 File), as for the previous harvest plot, the results were socially patterned towards higher risk of foodborne (*Campylobacter*, *Salmonella*, *Yersinia enterocolitica*) and person-to-person (viral GI infections, *Shigella*) GI infection in disadvantaged children and no association for waterborne infections (*Giardia*, *Cryptosporidium*). Only three studies explored the relationship between predominantly environmental GI infections (STEC) and SES and none were of high quality. For adults (lower half of Figure A in S1 File), there were also notable gaps in studies exploring the relationship in environmental or waterborne GI infections. There was a clear pattern with studies reporting lower risk in more disadvantaged adults or no association for studies exploring the relationship between predominantly foodborne GI infections and SES, and these studies were generally of medium quality.

No clear difference was observed in the relationship between SES and GI infection, when comparing point estimates based on area and individual SES measures, or when comparing point estimates from different countries (based on level of development or climate) (data not shown).

Of the 102 studies included in this systematic review, 77 studies were included in the meta-analysis. These 77 studies contributed 83 effect estimates. Of the 25 studies that could not be included in a meta-analysis, 15 did not provide sufficient quantitative data, six did not use a dichotomous outcome and four analysed the same cases as other studies (Table B in S1 File). Since age was highlighted as a key potential effect modifier in the harvest plots, estimates from the same study stratified by age were retained individually in the meta-analysis to allow for the investigation of this variable.

The pooled risk ratio for GI infection comparing low verses high SES for all studies combined was 1.06 (95% CI 0.95–1.19), with considerable statistical heterogeneity (I^2 99.08%).



Each bar represents one study. The height of the bars represent the quality of the study. Studies are classed into those showing lower risk in disadvantaged individuals/areas, no association or higher risk in disadvantaged individuals/areas and are stratified by individual and area based SES measures.

Fig 2. Harvest plot for risk of GI infection by SES, stratified by age, GI infection measure and SES measure.

https://doi.org/10.1371/journal.pone.0191633.g002

Potential effect modifiers and sources of heterogeneity were further explored in a multivariate random-effects meta-regression in an attempt to quantitatively explain some of the heterogeneity. In univariate meta-regression, the risk of GI infection for low compared to high SES was on average significantly higher among studies that analysed hospitalised cases, and non-significantly higher among studies that analysed cases identified via population-based surveys and general practices, compared to studies that analysed laboratory recorded cases (Table 3). Amongst studies using laboratory records, the risk of GI infection for low compared to high SES was significantly lower among studies that analysed environmental pathogens, and significantly higher among studies that analysed person-to-person pathogens, compared to studies that analysed foodborne pathogens. The risk of GI infection between low and high SES groups was not statistically significantly different between studies conducted in countries with different climates and levels of development. Additionally, the risk of GI infection for low versus high SES was non-significantly lower among studies that used area-level compared to individual-level SES measures.

In multivariate meta-regression (excluding pathogen type since not all studies analysed specific pathogens), age was identified as the only statistically significant potential effect modifier;

Table 3. Univariate and multivariate random-effects meta-regression for GI infection risk between low and high SES groups.

		Univariate RR (95% CI)	Multivariate RR (95% CI)	Number observations
Method of sampling GI infection cases	Laboratory records	1 (ref)	1 (ref)	43
	Population-based survey	1.11 (0.85–1.44)	1.04 (0.75–1.43)	23
	GP presentation	1.18 (0.71–1.94)	1.02 (0.62–1.69)	5
	Hospital admissions	1.49 (1.08–2.07)*	1.24 (0.88–1.73)	12
SES measure	Individual level	1 (ref)	1 (ref)	50
	Area level	0.87 (0.69–1.09)	0.92 (0.70–1.22)	33
Age of participants	Adult	1 (ref)	1 (ref)	14
	Mixed ages	1.17 (0.88–1.54)	1.22 (0.90–1.66)	42
	Child	1.89 (1.40–2.55)***	1.87 (1.35–2.59)***	27
Country Human Development Index[a]	Upper tertile	1 (ref)	1 (ref)	39
	Middle tertile	0.98 (0.76–1.25)	1.09 (0.84–1.41)	30
	Lower tertile	1.04 (0.73–1.49)	0.88 (0.62–1.25)	14
Country climate	Temperate/Mediterranean	1 (ref)	1 (ref)	62
	Arid	1.05 (0.69–1.61)	1.01 (0.67–1.52)	7
	Snow	0.81 (0.60–1.10)	0.89 (0.67–1.19)	14
Pathogen type[b]	Foodborne	1 (ref)	-	28
	Waterborne	0.73 (0.46–1.14)	-	8
	Environmental	0.46 (0.23–0.91)*	-	3
	Person-to-person	1.65 (1.05–2.59)*	-	7

CI = confidence interval; GI = gastrointestinal; ref = reference category; RR = ratio of risk ratios; SES = socioeconomic status

a Higher values indicate higher level of human development.

b Not all studies analysed specific pathogens, therefore this variable was not entered into the multivariate model.

*p <0.05.

**p <0.01.

***p <0.001.

https://doi.org/10.1371/journal.pone.0191633.t003

the risk ratios for GI infection between low and high SES groups observed by studies that analysed children, were on average 1.87 (95% CI 1.35–2.59) times the risk ratios observed by studies that focused on adults (Table 3).

A forest plot for the studies stratified by age is shown in Fig 3. For children, the pooled risk ratio was 1.51 (95% CI 1.26–1.83) with I^2 97.87%. For adults, the pooled risk ratio was 0.79 (95% CI 0.58–1.06) with I^2 98.64%. In sensitivity analyses, the results were similar when restricting to studies of high and medium quality only.

Two main outliers were identified in the forest plot (Fig 3; Jackson et al.[34], Fullerton et al. [35]). Both of these studies were conducted in the United States using laboratory records.

A contour-enhanced funnel plot[36,37] was produced to assess publication bias (Figure B in S1 File). Points within the plot appeared largely symmetrical, indicating that publication bias was unlikely.

Discussion

In this systematic review and meta-analysis of observational studies from developed countries we found evidence of an association between lower SES and a higher risk of GI infections for children, but no association in adults. Overall, age explained a small proportion of the heterogeneity observed across the studies as a whole.

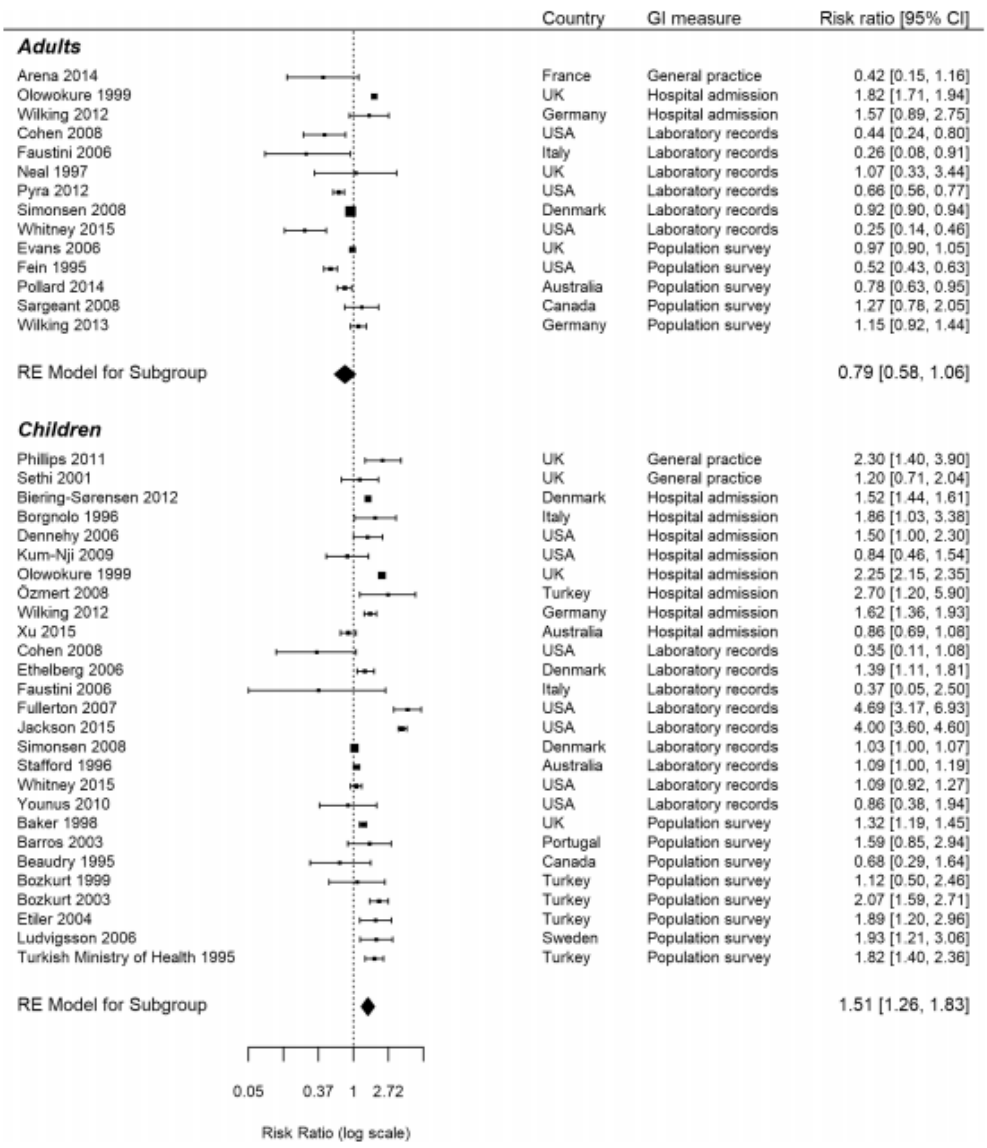


Fig 3. Forest plot for studies stratified by age.

https://doi.org/10.1371/journal.pone.0191633.g003

To the best of our knowledge, this study provides the first meta-analysis on this topic. We included a broad range of study designs and data sources, as well as definitions of GI infections. We used harvest plots[28] to summarise all studies, not exclusively studies with a quantitative estimate. This allowed the exploration of heterogeneity and provided important insights to inform our meta-analysis. Selection bias was mitigated by double screening throughout.

We explored the potential for publication bias, and this was not evident in our funnel plot. Subgroup analyses were defined a priori[23] which minimised the potential issues of performing multiple analyses of the data. Furthermore, these results reflect trends in inequalities of GI infections across numerous developed countries, adding confidence that the results may be generalisable.

Our review included syndromic definitions of GI infections in the absence of laboratory confirmation. We were hence able to identify literature on the burden of symptoms by SES and capture population groups who may not seek healthcare for their illness and consequently may not be included in studies which use laboratory data to identify cases. This was particularly important for this review as the decision to seek healthcare may be related to SES.

To explore sources of heterogeneity, stratified meta-analyses and meta-regression were performed on the basis of factors mentioned in the literature. Despite this, a large amount of heterogeneity remains unexplained. As seen in the forest plot, effect estimates were similar; however there were several outliers with wide confidence intervals combined with studies with narrow confidence intervals, which may provide some explanation for the extreme statistical heterogeneity observed. It is possible that factors that could not be adjusted for may explain the high residual heterogeneity. The studies covered a broad range of healthcare systems, with individual biases and caveats. SES was measured in numerous ways, and categorisation of low and high SES may have differed considerably between studies. The primary aims of the individual studies varied, as did the variables used to statistically adjust the associations between SES and GI infection risk. Further, the studies were conducted in socioeconomically diverse countries, including countries that have been in transition between developing and developed e.g. Turkey, which could potentially limit the interpretation of the results. However, our analysis of level of country development as a source of heterogeneity showed no statistically significant difference between studies conducted in countries with different levels of development. It should be noted, however, that the large amount of heterogeneity may have reduced the power to detect statistically significant modifiers in the meta-regression (Table 3), and therefore non-significance should not necessarily be interpreted as evidence that a potential modifier had no effect on the relationship between SES and risk of GI infection.[38]

Non-English language studies were excluded due to time limitations and costs of translating studies, and it is possible that bias may have been introduced by grouping the pathogen-specific studies by predominant mode of transmission, particularly for pathogens such as STEC which have multiple modes of transmission; however we consulted experts in GI pathogens for advice on the most appropriate groupings, and sensitivity analysis showed similar results when reclassifying STEC as a predominantly foodborne pathogen (data not shown). There were many ecological studies, studies conducted in Europe and studies assessed as generally low quality. Additionally, there were a few studies that focused on individual pathogens and stratified analyses by age. A number of studies had potential for bias due to study design; such as case-control studies, several of which selected controls based on the geographical residence of cases or through case-nomination, thereby potentially biasing the relationship between SES and GI infections towards the null. However, the results were similar when sensitivity analyses were conducted excluding studies which controlled for or matched by SES.

Despite the remaining heterogeneity, the evidence in this systematic review suggests that the relationship between SES and GI illness varies with age, with disadvantaged children at

greater risk of GI infection compared to more advantaged children. There are no other systematic reviews that have addressed this topic in developed countries. Newman et al.[22] undertook a systematic review of the association between SES and foodborne illness, a subset of GI infections; however they did not look at differences by age or different levels of healthcare reporting such as hospitalisation. Our results are in line with those of Newman et al.[22] for foodborne and laboratory confirmed pathogen-specific results, in that there were no consistent trends across all studies or pathogens for a single SES measure, perhaps indicating weakness in the measures of social class or differential effects of SES by pathogen type.[39]

We can speculate that children may be more likely to be taken to seek medical help regardless of SES, so the higher risk of GI infection seen in children might reflect real differences by SES, rather than bias due to differential healthcare seeking behaviour. Our differing findings for children compared with adults could also reflect differential exposures or immunity by SES in children. Within developing countries, *Campylobacter* is almost exclusively seen in disadvantaged children[40–43] while adults are rarely infected or identified,[44] potentially due to different healthcare seeking behaviour. This pattern is also seen for other bacteria and parasites.[45,46] We hypothesise that disadvantaged children are more exposed to these GI infections in childhood but that re-exposure leads to better immunity and subsequent asymptomatic infection later in life compared with their more advantaged counterparts.

Of note, the majority of studies that analysed hospital admission cases also analysed children only, making it difficult to separate out the potential modifying effects of these variables. In univariate meta-regression the risk of GI infection for low compared to high SES was significantly higher among studies that analysed hospitalised cases, compared to studies that analysed laboratory recorded cases, however this association was attenuated and rendered non-significant following adjustment for age in multivariate analysis (Table 3). Risk of person-to-person GI infection for low compared to high SES was significantly higher compared to the risk of foodborne GI infection in univariate meta-regression and risk of environmental GI infection was significantly lower. We recommend that future research investigating the association between SES and GI infection risk, should provide results stratified by adult and child age groups wherever possible, and additional insight may be gleaned by investigating the potential modifying role of pathogen type on the association.

To conclude, this systematic review finds that, in developed countries, disadvantaged children, but not adults, are at greater risk of GI infection compared to their more advantaged counterparts. Strategies to improve childhood socioeconomic conditions are likely to reduce the burden of GI illness. Additional high quality research is needed to investigate the association between SES and GI infection risk which stratifies results by age and pathogen type. Gaining greater insight into this relationship will help to inform policies to reduce the health inequalities identified.

Supporting information

S1 File. Supporting information File 1.
(PDF)

S2 File. PRISMA checklist.
(PDF)

Acknowledgments

This work was funded by the National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections.

Author Contributions

Conceptualization: Natalie L. Adams, Tanith C. Rose, Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Margaret Whitehead, David C. Taylor-Robinson.

Data curation: Natalie L. Adams, Tanith C. Rose, Victoria J. K. Howard.

Formal analysis: Natalie L. Adams, Tanith C. Rose, Ross Harris.

Funding acquisition: Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Margaret Whitehead, David C. Taylor-Robinson.

Investigation: Natalie L. Adams, Tanith C. Rose.

Methodology: Natalie L. Adams, Tanith C. Rose, Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Margaret Whitehead, Ross Harris, David C. Taylor-Robinson.

Project administration: Natalie L. Adams, Tanith C. Rose.

Supervision: Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Margaret Whitehead, Ross Harris, David C. Taylor-Robinson.

Validation: Natalie L. Adams, Tanith C. Rose, Ross Harris.

Visualization: Natalie L. Adams, Tanith C. Rose.

Writing – original draft: Natalie L. Adams, Tanith C. Rose.

Writing – review & editing: Natalie L. Adams, Tanith C. Rose, Jeremy Hawker, Mara Violato, Sarah J. O'Brien, Benjamin Barr, Victoria J. K. Howard, Margaret Whitehead, Ross Harris, David C. Taylor-Robinson.

References

1. Food Standards Agency. The second study of infectious intestinal disease in the community (IID2 Study). Available from: https://www.food.gov.uk/science/research/foodborneillness/b14programme/b14projlist/b18021
2. Food Standards Agency. Foodborne disease strategy 2010–15: an FSA programme for the reduction of foodborne disease in the UK. 2011. Available from: http://www.food.gov.uk/sites/default/files/multimedia/pdfs/fds2015.pdf
3. Tam CC, Viviani L, Rodrigues LC, O'Brien SJ. The second study of infectious intestinal disease (IID2): increased rates of recurrent diarrhoea in individuals aged 65 years and above. BMC Public Health. 2013; 13: 739. https://doi.org/10.1186/1471-2458-13-739 PMID: 24219653
4. Zappe Pasturel B, Cruz-Cano R, Rosenberg Goldstein RE, Palmer A, Blythe D, Ryan P, et al. Impact of rurality, broiler operations, and community socioeconomic factors on the risk of campylobacteriosis in Maryland. Am J Public Health. 2013; 103: 2267–75. https://doi.org/10.2105/AJPH.2013.301338 PMID: 24134343
5. Bessell PR, Matthews L, Smith-Palmer A, Rotariu O, Strachan NJ, Forbes KJ, et al. Geographic determinants of reported human Campylobacter infections in Scotland. BMC Public Health. 2010; 10: 423. https://doi.org/10.1186/1471-2458-10-423 PMID: 20633277
6. Davis MA, Moore DL, Baker KN, French NP, Patnode M, Hensley J, et al. Risk factors for campylobacteriosis in two Washington state counties with high numbers of dairy farms. J Clin Microbiol. 2013; 51: 3921–7. https://doi.org/10.1128/JCM.01433-13 PMID: 24025908
7. Majowicz SE, Horrocks J, Bocking K. Demographic determinants of acute gastrointestinal illness in Canada: a population study. BMC Public Health. 2007; 7: 162. https://doi.org/10.1186/1471-2458-7-162 PMID: 17640371
8. Olowokure B, Hawker J, Weinberg J, Gill N, Sufi F. Deprivation and hospital admission for infectious intestinal diseases. Lancet. 1999; 353: 807–8. https://doi.org/10.1016/S0140-6736(99)00611-X PMID: 10459964
9. Biering-Sorensen S, Sondergaard G, Vitting Andersen K, Andersen AM, Mortensen LH. Time trends in socio-economic factors and risk of hospitalisation with infectious diseases in pre-school children 1985–

2004: a Danish register-based study. Paediatr Perinat Epidemiol. 2012; 26: 226–35. https://doi.org/10.1111/j.1365-3016.2011.01255.x PMID: 22471682

10. Pockett RD, Adlard N, Carroll S, Rajoriya F. Paediatric hospital admissions for rotavirus gastroenteritis and infectious gastroenteritis of all causes in England: an analysis of correlation with deprivation. Curr Med Res Opin. 2011; 27: 777–84. https://doi.org/10.1185/03007995.2011.555757 PMID: 21294699
11. de Wit MAS, Koopmans MPG, Kortbeek LM, Wannet WJ, Vinjé J, van Leusden F, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol. 2001; 154: 666–74. PMID: 11581101
12. Pollard CM, Meng X, Williamson S, Dodds J, Binns CW. Eating out is associated with self-reported food poisoning: a Western Australia population perspective, 1998 to 2009. Public Health Nutr. 2014; 17: 2270–7. https://doi.org/10.1017/S1368980013002371 PMID: 24172074
13. Gillespie IA, O'Brien SJ, Penman C, Tompkins D, Cowden J, Humphrey TJ. Demographic determinants for Campylobacter infection in England and Wales: Implications for future epidemiological studies. Epidemiology and Infection. 2008; 136(12): 1717–25. https://doi.org/10.1017/S0950268808000319 PMID: 19000328
14. Nichols GL, Richardson JF, Sheppard SK, Lane C, Sarran C. Campylobacter epidemiology: a descriptive study reviewing 1 million cases in England and Wales between 1989 and 2011. BMJ Open. 2012; 2 (4): e001179. https://doi.org/10.1136/bmjopen-2012-001179 PMID: 22798256
15. Lake IR, Harrison FC, Chalmers RM, Bentham G, Nichols G, Hunter PR, et al. Case-control study of environmental and social factors influencing cryptosporidiosis. European Journal of Epidemiology. 2007; 22(11): 805–11. https://doi.org/10.1007/s10654-007-9179-1 PMID: 17891460
16. De Wit MA, Koopmans MP, van Duynhoven YT. Risk Factors for Norovirus, Sapporo-like Virus, and Group A Rotavirus Gastroenteritis. Emerging Infectious Diseases. 2003; 9(12): 1563–70. https://doi.org/10.3201/eid0912.020076 PMID: 14720397
17. Gillespie I, Mook P, Little C, Grant KA, McLauchlin J. Human listeriosis in England, 2001–2007: association with neighbourhood deprivation. Euro Surveillance. 2010; 15(27): 7–16. PMID: 20630146
18. Wilking H, Hohle M, Velasco E, Suckau M, Eckmanns T. Ecological analysis of social risk factors for Rotavirus infections in Berlin, Germany, 2007–2009. International Journal of Health Geographics. 2012; 11: 37. https://doi.org/10.1186/1476-072X-11-37 PMID: 22929067
19. Borgnolo G, Barbone F, Scornavacca G, Franco D, Vinci A, Iuculano F. A case-control study of Salmonella gastrointestinal infection in Italian children. Acta Paediatrica. 1996; 85(7): 804–8. PMID: 8819545
20. Sethi D, Cumberland P, Hudson MJ, Rodrigues LC, Wheeler JG, Roberts JA, et al. A study of infectious intestinal disease in England: risk factors associated with group A rotavirus in children. Epidemiology and Infection. 2001; 126(1): 63–70. PMID: 11293683
21. Phillips G, Tam CC, Rodrigues LC, Lopman B. Risk factors for symptomatic and asymptomatic norovirus infection in the community. Epidemiology and Infection. 2011; 139(11): 1676–86. https://doi.org/10.1017/S0950268810002839 PMID: 21205382
22. Newman KL, Leon JS, Rebolledo PA, Scallan E. The impact of socioeconomic status on foodborne illness in high-income countries: a systematic review. Epidemiol Infect. 2015; 143: 2473–85. https://doi.org/10.1017/S0950268814003847 PMID: 25600652
23. Rose TC, Adams N, Taylor-Robinson DC, Barr B, Hawker J, O'Brien SJ, et al. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review protocol. Systematic Reviews. 2016; 5: 13. https://doi.org/10.1186/s13643-016-0187-7 PMID: 26791956
24. Pope D. Introduction to systematic reviews [lecture]. Liverpool: University of Liverpool 2015.
25. Puzzolo E, Stanistreet S, Pope D, Bruce N, Rehfuess E. Factors influencing the large scale uptake by households of cleaner and more efficient household energy technologies: a systematic review. Evidence for Policy and Practice Information and Co-ordinating Centre. 2013. Available from: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3426
26. Rehfuess E, Puzzolo E, Stanistreet S, Pope D, Bruce NG. Enablers and barriers to large-scale uptake of improved solid fuel stoves: a systematic review. Environ Health Perspect. 2014; 122: 120–30. https://doi.org/10.1289/ehp.1306639 PMID: 24300100
27. Voss PH, Rehfuess EA. Quality appraisal in systematic reviews of public health interventions: an empirical study on the impact of choice of tool on meta-analysis. J Epidemiol Community Health. 2013; 67: 98–104. https://doi.org/10.1136/jech-2011-200940 PMID: 22851579
28. Ogilvie D. Fayter D. Petticrew, Sowden A, Thomas S, Whitehead M, et al. The harvest plot: A method for synthesising evidence about the differential effects of interventions. BMC Medical Research Methodology. 2008; 8: 8. https://doi.org/10.1186/1471-2288-8-8 PMID: 18298827
29. United Nations Development Programme. Human Development Index (HDI). Available from: http://hdr.undp.org/en/content/human-development-index-hdi

30. Met Office. Climate Zones. Available from: http://www.metoffice.gov.uk/climate-guide/climate/zones
31. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration 2011. Available from: www.handbook.cochrane.org
32. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random effects regression model for meta-analysis. Statistics in Medicine. 1995; 14: 395–411. PMID: 7746979
33. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Statistics in Medicine. 1999; 18: 2693–2708. PMID: 10521860
34. Jackson R, Smith D, Tabnak F, Vugia D. Disparities of shigellosis rates among California children by race/ethnicity and census tract poverty level, 2000–2010. Pediatr Infect Dis J. 2015; 34: 843–7. https://doi.org/10.1097/INF.0000000000000746 PMID: 26020406
35. Fullerton KE, Ingram LA, Jones TF, Anderson BJ, McCarthy PV, Hurd S, et al. Sporadic campylobacter infection in infants: a population-based surveillance case-control study. Pediatr Infect Dis J. 2007; 26: 19–24. https://doi.org/10.1097/01.inf.0000247137.43495.34 PMID: 17195700
36. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J Clin Epidemiol. 2008; 61: 991–6. https://doi.org/10.1016/j.jclinepi.2007.11.010 PMID: 18538991
37. The metafor Package: a meta-analysis package for R. Contour-enhanced funnel plot. Available from: http://www.metafor-project.org/doku.php/plots:contour_enhanced_funnel_plot
38. Hempel S, Miles JN, Booth MJ, Wang Z, Morton SC, Shekelle PG. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. Systematic Reviews. 2013; 2: 107. https://doi.org/10.1186/2046-4053-2-107 PMID: 24286208
39. Simonsen J, Frisch M, Ethelberg S. Socioeconomic risk factors for bacterial gastrointestinal infections. Epidemiology. 2008; 19: 282–90. https://doi.org/10.1097/EDE.0b013e3181633c19 PMID: 18300694
40. Da Silva Quetz J, Lima IFN, Havt A, de Carvalho EB, Lima NL, Soares AM, et al. Campylobacter jejuni and Campylobacter coli in children from communities in Northeastern Brazil: molecular detection and relation to nutritional status. Diagn Microbiol Infect Dis. 2010; 67: 220–7. https://doi.org/10.1016/j.diagmicrobio.2010.02.025 PMID: 20542202
41. Kakai R, Wamola IA, Bwayo JJ, Ndinya-Achola JO. Enteric pathogens in malnourished children with diarrhoea. East Afr Med J. 1995; 72: 288–9. PMID: 7555883
42. Lloyd-Evans N, Drasar BS, Tomkins AM. A comparison of the prevalence of campylobacter, Shigellae and Salmonellae in faeces of malnourished and well nourished children in The Gambia and Northern Nigeria. Trans R Soc Trop Med Hyg. 1983; 77: 245–7. PMID: 6868107
43. Fernández H, Vera F, Villanueva MP, García A. Occurrence of campylobacter species in healthy well-nourished and malnourished children. Braz J Microbiol. 2008; 39: 56–8. https://doi.org/10.1590/S1517-83822008000100013 PMID: 24031178
44. Coker AO, Isokpehi RD, Thomas BN, Amisu KO, Obi CL. Human Campylobacteriosis in Developing Countries. Emerg Infect Dis. 2002; 8: 237–43. https://doi.org/10.3201/eid0803.010233 PMID: 11927019
45. Nematian J, Nematian E, Gholamrezanezhad A, Asgari AA. Prevalence of intestinal parasitic infections and their relation with socio-economic factors and hygienic habits in Tehran primary school students. Acta Trop. 2004; 92: 179–86. https://doi.org/10.1016/j.actatropica.2004.06.010 PMID: 15533285
46. O'Ryan M, Prado V, Pickering LK. A millennium update on pediatric diarrheal illness in the developing world. Semin Pediatr Infect Dis. 2005; 16: 125–36. PMID: 15825143