EVALUATING THE COST-EFFECTIVENESS OF SCHOOL-BASED PUBLIC HEALTH INTERVENTIONS. EXPLORING THE SCHOOLS' PERSPECTIVE

By

KATHRYN BREHENY

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

Health Economics Unit

Institute of Applied Health Research

College of Medical and Dental Sciences

University of Birmingham

October 2019

ABSTRACT

Public health interventions are widely implemented in schools, although health economics methods have predominantly developed in healthcare. This thesis explores a broader perspective in economic evaluation to better support schools' evidence needs and decision making.

A systematic review was conducted on challenges and solutions for economic evaluation in schools. Semi-structured interviews were conducted with school staff participating in the Birmingham Daily Mile trial, to establish the decision-making process and considerations when prioritising public health interventions. A wellbeing measure (the Middle Years Development Instrument [MDI]) was validated for the first time in UK children, as a precursor to inclusion in an exploratory economic evaluation of the Daily Mile.

Interviews found that school leadership staff wanted evidence of cost-effectiveness. Health outcomes did not align with schools' priorities and staff recognised the opportunity cost of delivering interventions. The MDI performed well in children aged 7-10, with several domains demonstrating responsiveness to changes in quality of life and self-rated health

In conclusion, this thesis demonstrates disparities between economic methods used in a healthcare context and schools. The first economic evaluation informed by school decision makers is reported, with further work required to enable comparability with other economic evaluations and estimate children's opportunity costs.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisors Emma Frew, Joanna Coast and Iestyn Williams for their advice and support. I could not have asked for better supervision.

I would like to acknowledge colleagues that contributed to the Birmingham Daily Mile trial, particularly for providing assistance with the analyses and facilitating access to the qualitative study participants. I am very grateful to the staff who gave up their time to be interviewed and the participating schools, staff and pupils.

Without the NIHR funding for my Research Associate position and Birmingham City Council's contribution to the Birmingham Daily Mile trial, this PhD would not have been possible.

I am grateful to the research community who have contributed to the submission of my thesis. Colleagues in HEU have provided invaluable support and advice. Comments and feedback received at conferences I have attended has also been extremely helpful in shaping my thesis.

Finally, I would like to thank my friends and family for their unwavering support and patience throughout my PhD. I promise this will be my last degree. Thank you to James Best for proofreading my thesis and being there throughout.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABC	Aberrant Behaviour Checklist
ACE-Obesity	Assessing cost-effectiveness in obesity
ACRA	Advisory Committee on Resource Allocation
AFLY5	Active for life year 5
APH	Academic public health
APPLE	Alberta Project Promoting active Living and healthy eating in schools
APPLE schools	A Pilot Programme for Lifestyle and Exercise
BAEW	Be Active Eat Well
BEI	British Education Index
BMI	Body Mass Index
BMIz	BMI z-score
CAMHS	Children and Adolescent Mental Health Services
CBA	Cost benefit analysis
CCA	Cost-consequence analysis
CCG	Clinical Commissioning Group
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CER	Cost effectiveness ratio
CETP	Cholesteryl ester transfer protein
CHOICES	Childhood Obesity Intervention Cost-Effectiveness Study
CHU9D	Child Health Utility 9D
CI	Confidence interval
CIEH	Chartered Institute for Environmental Health
CMA	Cost minimisation analysis
COPD	Chronic obstructive pulmonary disease
CUA	Cost utility analysis
DALYs	Disability Adjusted Life Years
DCE	Discrete choice experiment
DfES	Department for Education and Science
EAP	European Action Plan

EiC	Excellence in Cities
EPHOs	Essential Public Health Operations
ERIC	Educational resources information centre
GMC	General Medical Council
GP	General practitioners
HC	Healthcare commissioning
HCP	Healthy Child Programme
HE	Higher education
HEER	Health Economics Evidence Resource
HI	Health improvement
HIV	Human immunodeficiency virus
HK	Health intelligence/knowledge
HMIC	Healthcare Management Information Consortium
HP	Health protection
HPA	Health Protection Agency
HPV	Human papilloma virus
HRQL	Health related quality of life
HWB	Health and Wellbeing Boards
ICER	Incremental cost-effectiveness ratio
ID	Identifier
IMD	Index of multiple deprivation
IT	Information technology
LA	Local authority
LEA	Local education authority
MCDA	Multi-criteria decision analysis
MD	Mean difference
MDI	Middle Years Development Instrument
MET	Metabolic equivalent
MVPA	Moderate to vigorous physical activity
N/A	Not applicable
NCMP	The National Child Measurement Programme
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research

NIHS	National Health Interview Survey
NMC	Nursing and Midwifery Council
OFSTED	Office for Standards in Education, Children's Services and Skills
ONS	Office for National Statistics
PCT	Primary Care Trusts
PE	Physical Education
PH	Public Health
PHA	Public Health Agency
PHE	Public Health England
PHSE	Personal Social and Health Education
PPI	Patient and Public Involvement
PQAQ	Paediatric Quality Appraisal Questionnaire
PSA	Probabilistic sensitivity analysis
PSS	Personal and social services
PWI-SC	Personal-Wellbeing Index – School Children
QALY	Quality adjusted life year
QOL	Quality of life
RCT	Randomised controlled trial
ROI	Return on investment
RSC	Regional School Commissioners
SATs	Standard Attainment Tests
SBS-QoL	Short-Bowel Syndrome Quality of Life Questionnaire
SCT	Social Cognitive Theory
SD	Standard deviation
SEL	Social and emotional learning
SLSS	Heubner's Student Life Satisfaction Scale
SNaX	Students for Nutrition and eXercise
SPOT	Spend and Outcomes Tool
SRM	Standardised response mean
SROI	Social return on investment
SWLS-C	Satisfaction with Life Scale - Child
TA	Teaching assistant
UK	United Kingdom

UKPHR	UK Public Health Register
USA	United States of America
WELBY	Wellbeing adjusted life year
WEMWBS	Warwick-Edinburgh Mental Wellbeing Scale
WHO	World Health Organisation
WTA	Willingness to accept
WTP	Willingness to pay

INTRODUCTION

Economic evaluation is a method that allows decision makers to compare the costs and benefits of alternative interventions, enabling evidence-based decisions concerning investment, disinvestment or no investment. Methodology has predominantly developed in the context of the provision of primary and secondary care, whereby ill health is treated. In contrast, public health encompasses health promotion, prevention of ill health and protection of health. These contrasting remits mean that services are inevitability provided in different contexts and require different evaluation methods. In addition, their funding has been uncoupled, meaning responsibility for the prioritisation of these services in England has been transferred from the National Health Service (NHS) to local authorities. The methodological challenges have been recognised in recent years in both guidance and the literature, with some alternatives proposed (Edwards et al., 2013; NICE, 2014; Weatherly et al., 2009). Schools have been cited as an ideal setting to provide public health interventions targeted at children, yet this setting has many distinct features and potential issues not compatible with 'traditional' methods of economic evaluation. These include the responsibility for prioritisation, what should be maximised and how this can be captured in children, choice of outcomes and relevant costs.

Thesis aims

The aims of this thesis are to:

1. Explore the decision-making structure in schools regarding health and wellbeing initiatives
2. Understand what outcomes are important in a school setting
3. Understand the costs of providing public health interventions in schools

This thesis examines the methodology used in economic evaluations in schools and prioritisation of public health interventions in this setting. Building on the recommendation that wellbeing is an important outcome, a measure of wellbeing is evaluated in a UK sample for the first time and used in an exploratory economic evaluation from the schools' perspective.

Thesis structure

Chapter 1 explores the foundations of public health and delineates what it encompasses. The structure and delivery of public health in the UK is briefly discussed, before focussing specifically on the school setting. The wide remit and indistinct boundaries of public health mean that public health functions in schools are not necessarily readily recognised. Those easily defined through mandatory guidance and those less obviously identifiable are discussed, in relation to their provision and regulatory responsibilities. This allows the contextualisation of each of the thesis aims (Aims 1-3) within a school setting and considering the aims of public health systems. The chapter highlights both the important role of schools in delivering public health interventions and the opportunity to use economic evaluation to evaluate new public health initiatives in this setting.

Chapter 2 provides a discussion of the theoretical basis for economic evaluation and how it is operationalised, highlighting the methodological issues posed by public health, children's interventions and in education. Following an introduction to the rationale for economic evaluation, a description of the important methodological components is then provided. The challenges of evaluating public health interventions are presented and why methods used in adult populations should not necessarily be transferred to children. The chapter finishes by exploring how economic theory has been used within education policy and how attempts have been made to apply healthcare methods. Describing the foundations of economic evaluation and the methodological challenges posed by public health interventions and paediatric populations provides an indication of the additional difficulties that could occur when methods are applied in a school setting. This chapter demonstrates how the requirements and perspective of the decision maker can determine the analytic approach, although in contrast to healthcare, the decision maker (Aim 1) and the relevant outcomes (Aim 2) and costs (Aim 3) in schools are unexplored, which warrants further examination in this thesis.

Chapter 3 reports a systematic review of economic evaluations of school-based interventions addressing risk factors for obesity and obesity prevention. This review was conducted to identify approaches used in a school setting, and to assess whether they offer potential solutions to the challenges of evaluating public health interventions identified in Chapter 2. The review also intended to develop an understanding of current practice in regard to outcomes used (Aim 2) in economic evaluations and what costs are included in analyses (Aim 3). The inconsistent inclusion of costs and outcomes, and the ambiguity regarding the intended decision maker indicate that these are considerations that require exploration,

Chapter 4 introduces the Daily Mile and the Birmingham Daily Mile trial. The Daily Mile is a school-based initiative that is advocated as an initiative that can contribute to childhood obesity prevention. With no robust evidence supporting this claim, The Birmingham Daily Mile cluster randomised controlled trial (RCT) was conducted. The empirical studies reported in Chapters 5, 6 and 7 and which address the thesis aims, used this trial and the Daily Mile intervention as case studies. Chapter 4 describes the trial procedures and reports the clinical results.

Chapter 5 reports a qualitative study with school staff. The aims were to understand decision making in the school context (Aim 1), the costs associated with public health interventions in schools (Aim 3) and the outcomes that are valued (Aim 2). Semi-structured interviews were conducted with leadership staff, teachers and other staff with roles such as pastoral leads. The constant comparison method was used to analyse the data and the results report the main themes that arose from the data. The chapter assists in developing an understanding of what the schools' perspective might comprise.

Chapter 6 presents an evaluation of the psychometric properties of the Middle Years Development Instrument (MDI). This is a self-reported measure of wellbeing included in the Birmingham Daily Mile trial, which had previously not been used in the UK or a clinical trial. Wellbeing was raised as an outcome perceived as important by school staff; therefore the suitability of the MDI was evaluated. This is the first step in applying the findings that emerged from the second aim of the thesis (important outcomes in a school setting). Properties evaluated include construct validity, internal consistency and responsiveness. The findings of this chapter contribute to undertaking an economic evaluation from the schools' perspective, using a valid and responsive measure of wellbeing and reported in Chapter 7.

Chapter 7 reports an exploratory economic evaluation of the Daily Mile using data from the Birmingham Daily Mile trial. This analysis encompasses the findings obtained from exploring the three thesis aims, by presenting evidence potentially relevant to the identified school decision maker and including costs and outcomes considered important. Two perspectives are used, with the school perspective informed by the findings reported in previous chapters and compared to a public sector perspective. This is the first evaluation of the cost-effectiveness of the

Daily Mile. Both a cost-utility and cost-effectiveness analyses were undertaken, with the MDI data utilised to present the results in terms of wellbeing, in line with the views of school staff.

A discussion is presented in Chapter 8. The findings of the thesis are summarised into themes, bringing together the results of the quantitative and qualitative work. These themes are discussed in relation to current policy and literature. Chapters 3-7 each end with a discussion of the respective empirical study, therefore the strengths and weaknesses of individual studies and their findings in context are noted only briefly in this chapter. Directions for future research are proposed and the thesis concludes with a summary of the key contributions made.

CHAPTER 1 PUBLIC HEALTH

The aim of this chapter is to describe the role of public health. In order to understand what public health aims to achieve, this chapter firstly describes the determinants of health. The definition of public health and its core functions are then described. An introduction to public health in the United Kingdom (UK) and a more detailed focus on schools is also presented.

1.1 **Determinants of Health**

There are several factors that play a role in determining health. Early literature discussing the origins of disease categorised diseases as those determined at fertilisation, (genetic diseases and other diseases attributable to multiple genes) and those that only occur in an appropriate environment (pre-natal and post-natal environments) (McKeown, 1979). More recent classifications expand upon those not attributable to genetics such as lifestyle factors, the natural environment and social factors. Social factors include economic, political and cultural circumstances. These are otherwise known as the social determinants of health (Detels et al., 2009) and are the most dominant drivers of the unequal distribution of health that exists in the population (Graham, 2004).

The long-term decline in mortality rate since the 1800's has been largely attributed to a fall in deaths from infectious diseases, although a substantial decline in non-infectious conditions also contributed (McKeown, 1979) . Improved hygiene and a changing environment reduced the spread of infections, whilst vaccination and effective treatments improved survival. For non-infectious conditions, improved nutrition and a decline in infant mortality were the most important contributors to reducing mortality.

1.1.1 Genetics

There is an established evidence base demonstrating the contribution of genetics to the development of many common diseases, and an understanding that many are the result of a complex interaction between the environment and predisposing genetic characteristics (Sankar et al., 2004). Furthermore, the outcomes of most degenerative and infectious diseases are influenced by an individual's genetics (Gottesman and Collins, 1994). In contrast to multifactorial inheritance which describes the interaction between the environment and genetics, single gene disorders such as cystic fibrosis are the result of mutations in a single gene and are inherited in recognisable patterns. These occur in about 0.5% of live births (Hirschhorn and Cooper, 1961; Lobo, 2008). Chromosomal disorders, caused by chromosomal abnormalities can occur at fertilisation, with approximately 0.6% of live births having such an abnormality (Natarajan, 2002). An example is Trisomy 21 (Down's Syndrome), where individuals have an extra copy of chromosome 21. Chromosomal abnormalities that occur after fertilisation can be spontaneous, although some are radiation or chemically induced by cytotoxic agents (Fröhling and Döhner, 2008).

Cancer in particular has a strong heritable component, sometimes caused by defects in particular genes. For example, 85% of individuals with a mutation of the BRCA1 are likely to develop early onset breast cancer (Gottesman and Collins, 1994). This is an unmodifiable risk factor making prevention irrelevant in this case, although early detection is a strategy that can improve outcomes. Cancer is not the only heritable disease however, with individuals having the APOB or cholesteryl ester transfer protein (CETP) genotypes carrying a 70% risk of developing coronary heart disease (Smith et al., 2005). The susceptibility of individuals to exposure to environmental toxins can also have a genetic basis, which has implications for employment (Smith et al., 2005). Where genetic susceptibility is discovered through screening or family history, individuals can be targeted with medical, environmental and behavioural interventions to reduce their risk of developing these conditions (Khoury, 1997). Nonetheless, in many cases genetic influences on health are factors that are unable to be modified.

1.1.2 Lifestyle factors

Lifestyle factors such as an unbalanced diet, physical inactivity and substance abuse have a great influence on an individual's health (Hillger, 2008). Behaviours detrimental to health include the consumption of alcohol, drugs and smoking of tobacco (Scarborough et al., 2011). These are linked to both short and long-term health conditions, including psychological issues. Considering alcohol for example, it was estimated that 5.3% of deaths in the UK in 2005 were directly attributable to alcohol consumption and 17.1% of mortality associated with ischemic heart disease was alcohol related (Balakrishnan et al., 2009). Identifying those at risk and modifying these health harming behaviours, such as sedentary lifestyles is essential to improving population health. Physical inactivity increases the risk of several diseases and also has a negative impact on metabolism (DeBusk et al., 1994).

A well balanced diet with an appropriate energy intake is essential for preventing nutrition related diseases, obesity and its related complications. Obesity has been strongly linked with several cancers such as colorectal, pancreatic and liver (Song and Giovannucci, 2016). In a review of the health burden of obesity in the United States of America (USA) (Flegal et al., 2015) it was found that the population attributable fraction of overweightness was up to 8% for cancer, between 5-15% for all-cause mortality and 7-44% for cardiovascular disease. In a separate review the population attributable fraction of an absence of playing sport for type 2 diabetes was up to 13% in men and 29% in women (Al Tunaiji et al., 2014). Nutrition can be affected by individual likes and dislikes, culture, and social factors in addition to knowledge of what constitutes a good diet. Factors influencing nutrition are not restricted to lifestyle, with the environment and income also exerting influence (Hillger, 2008).

Smoking impacts almost all organs in the body and is a major cause of premature death. Among UK women, two-thirds of all deaths of smokers in their 50s, 60s, and 70s are attributable to smoking and smokers lose on average at least 10 years of lifespan (Pirie et al., 2013). The estimated economic cost of smoking in the UK is £3.27 billion, similar to the cost of alcohol consumption (£3.23 billion) (Scarborough et al., 2011). Smoking during pregnancy can affect brain development and contributes to adverse pregnancy outcomes such as premature birth and stillbirth (Marufu et al., 2015). Smoking is a significant risk factor for lung cancer (Pirie et al., 2013) and other malignancies associated with smoking include colorectal and liver cancer (US Department of Health and Human Services, 2014). In addition to lung cancer, other respiratory harms of smoking include chronic obstructive pulmonary disease (COPD) and increased infections of the respiratory tract (US Department of Health and Human Services, 2014). An individual participant meta-analysis demonstrated an impact of smoking on cardiovascular mortality, acute coronary events and stroke in older adults. Smoking advanced the risk of dying of cardiovascular disease by 5.5 years, with hazard ratios of 2.07, 1.98 and 1.58 for cardiovascular mortality, acute coronary events and stroke, respectively (Mons et al., 2015).

Smoking is a major public health challenge, and since 2004 legislation has been introduced banning indoor smoking in public places (Frazer et al., 2016). This aims to reduce exposure to second hand smoke and decrease cigarette consumption. In the UK such a ban was implemented between 2006 and 2007. Research demonstrating their effectiveness is now emerging, with such public health interventions have reportedly been successful. For example acute coronary events reduced by 11.2% in Rome, Italy (Cesaroni et al., 2008) and a meta-analysis found the risk of hospital admission for acute myocardial infarction decreased by 17% (Meyers et al., 2009). Other positive health impacts of widespread bans include improved perinatal health, and reduced respiratory illness and mortality (Frazer et al., 2016). Public health interventions to facilitate smoking cessation include pharmacotherapy and behavioural support and many have been shown to be cost effective (Maciosek et al., 2017). Preventing uptake is critical in reducing the burden of smoking related illness and school, community and family-based programmes have been developed to this purpose, again proving largely cost-effective (Wang and Michael, 2015).

1.1.3 Environmental factors

The physical environment can affect health both directly and indirectly (Tones and Green, 2010). Seen as external factors out of an individual's control, direct effects include air pollution or a lack of clean drinking water. For example, Cutler and Miller (2005) estimated the impact of introducing clean water technologies in the USA. Water filtration was responsible for decreases in typhoid fever mortality (46%), total mortality (16%), infant mortality (43%) and child mortality (46%). Indirect effects have a less causal relationship, for example, a lack of green space for exercise could result in high cardiovascular disease rates in the local population (Tones and Green, 2010).

1.1.4 *Social determinants of health*

Even in developed countries, the poorest in society have shorter life expectancies and experience more illnesses than the better off (Dahl, 1994; Wilkinson and Marmot, 2003). This indicates that health is sensitive to the social and economic environments, which are largely modifiable. Establishing which factors contribute to poor health and understanding how these influence biological processes is important for guiding effective interventions designed to improve population health and reduce inequalities. There is a body of literature exploring the association between behaviour and biological processes such as stress and cortisol level (Burke et al., 2005). The scientific literature surrounding these will not be explored in depth here; rather the social determinants of health informed by this evidence base will be described.

The social gradient of health describes how health is reduced amongst the poorest in society compared to the richest. Disadvantages such as poor education and living conditions can contribute to disease and premature death and tend to accumulate in the same populations. In the USA, high minority and low education neighbourhoods have access to fewest physical activity facilities such as swimming pools or parks. These areas also have higher obesity and achieve less weekly physical activity (Gordon-Larsen et al., 2006). Living in stressful circumstances can contribute to several negative psychological states, such as low self-esteem, anxiety and social isolation. Illnesses such as cardiovascular disease and infections are also more prevalent in this population and can accumulate in the most vulnerable. A child's early years, even before birth, have an enduring impact on their emotional (Markussen et al., 2003) and physical health, and cognitive function in later life (Whitehead and Dahlgren, 1991). Nutritional deficiencies and substance abuse in pregnancy can impair foetal development and have long-term impacts on biological processes (Black et al., 2008; Wakschlag et al., 2002). A stimulating environment and secure attachment relationships aid healthy development, and a good education in childhood is associated with exposure to healthy lifestyle characteristics such as exercise and not smoking (Bradley and Corwyn, 2002). A longitudinal study of 3,032 individuals in the USA examined childhood disadvantage and health in adulthood (Ferraro et al., 2016). Being raised in a disadvantaged home was negatively associated with health, educational attainment, social integration and sense of control. Lifestyle risks such as smoking and drinking were more prevalent in this group.

Poverty, defined as a lack of economic resources, is both an outcome and facilitator of poor health (Townsend, 1979). Examples include needing to accept inferior living conditions such as damp and cold accommodation or the opportunity to earn a good wage being restricted by physical and emotional health issues (Marmot and Bell, 2012). Lacking the material resources that facilitate good health, such as good nutrition, can contribute to critical public health issues such as obesity and its associated complications (Marmot et al., 2008; Whitehead and Dahlgren, 1991). Diet is one of the largest determinants of diseases that cause early mortality, such as cardiovascular disease and cancer (Juonala et al., 2011). In addition to having the necessary economic resources, knowledge of what constitutes a healthy diet is critical. This is a key public health issue where legislation and communication can facilitate health gains (Parmenter et al., 2000).

Damaging health behaviours such as drinking and smoking can be a response to economic and social disadvantage, yet can worsen inequalities and result in falling further down the social gradient (Marmot, 2005; Whitehead and Dahlgren, 1991). Substance abuse is associated with indicators of social deprivation such as unemployment for example (Henkel, 2011). The socially excluded, those with few qualifications or those lacking social support are populations that have less access to material resources facilitating health. The feelings of being valued and cared for also have a protective effect on health, whereas an absence of these relationships is associated with more psychological disorders such as depression (Marmot, 2005).

1.2 **Defining Public Health**

The most widely accepted definition of public health, and that adopted by the UK Faculty of Public Health, is as follows: *"Public Health is the science and art of preventing disease, prolonging life and promoting health through the organised efforts of society"* (Acheson, 1988). In a report to the UK Government, Wanless (2004) extends this definition to identify facilitators and the importance of knowledge, adding "…*organised efforts and informed choices of society, organisations, public and private, communities and individuals*." Public health and medicine are complementary disciplines, although public health focuses on the health of the population in contrast to medicine's focus on individuals' health (Mann, 1997). The core public health functions of health promotion, prevention (prolonging of life) and protection included in Acheson's definition will now be described in addition to the World Health Organisation's (WHO) conceptualisation of public health (WHO, 2012).

1.2.1 Health Promotion

The Ottawa Charter for Health Promotion (WHO, 1986) conceptualised health promotion as '*the process of enabling people to increase control over and improve their health'*. Health was seen as a positive state which is conditional upon physical capabilities, social and personal resources (Breslow, 1999). This definition of health goes beyond the absence of disease and instead sees it as a resource for living, with health being seen as a spectrum, ranging from negative to positive. The role of health promotion therefore is to improve position on this continuum, in contrast to prevention which is avoidance of movement to the negative end (Breslow, 1999). The Ottawa Charter (WHO, 1986) proposed three strategies for health promotion. This first is advocacy, to create the condition for facilitating good health. Enabling a supportive environment and providing the knowledge and skills to make healthy decisions is the second and finally, mediation between groups to ensure the pursuit of good health. The aim of health promotion interventions is to enable people to take action to change determinants of health which are within their control (Wolfram and Fuchs, 2008). As the responsibility for change is perceived as being with the individual, behavioural theories have been used to inform interventions.

The psychologist Bandura proposed that as individuals have a certain degree of control over their health, by managing their lifestyle they can maximise their opportunity to live a long and healthy life (Bandura, 2005). The Social Cognitive Theory (SCT) (Bandura, 1998) applied to health promotion provides guidance on how to modify the socio-cognitive factors that control the self-regulation of behaviours and promote or impair health, for example health goals (Bandura, 2004). Similar to the World Health Organisation (WHO) Essential Public Health Operation (EPHO) (described in detail in Section 1.2.4) of health promotion's focus on education, the first factor is knowledge. An adequate understanding of health risks and benefits are a precondition for behaviour change, without which individuals have little reason to take any action (Bandura, 2004). Perceived self-efficacy is also a precursor to successful initiation and maintenance of change . Self-efficacy is an individual's beliefs in their capabilities to organise and complete actions required to achieve a level of attainment (Bandura, 1998) and without this individuals have little incentive to act. Another factor critical to behaviour change is the outcomes that individuals expect to achieve. These could be physical results, social reactions, an individual's personal evaluation. Seeing desired outcomes motivates or discourages positive or negative behaviour. Linked to knowledge, being able to see how behaviour can contribute to long-term and short-term goals enhances motivation. Goals provide self-incentives and guide development of enduring positive habits (Bandura, 1998).

The determinants of behaviour shared by the SCT and goal theory can be summarised as knowledge of health risks and benefits, belief in control you can have over your health, expected outcomes and health goals, the plans that individuals make to achieve their goals and the social facilitators and barriers that exist (Bandura, 2005). The Self-Management Model was developed to promote healthy lifestyles and is based upon the SCT principles of motivation and self-regulation (DeBusk et al., 1994). The core aspects of the interventions are providing information guides, self-monitoring health habits, setting small goals and providing feedback. The programme can be tailored to individual's self-efficacy beliefs and motivation, which can aid the narrowing of health inequalities (Bandura, 2005).

1.2.2 Prevention

Prevention is concerned with taking action on specified health issues in particular populations, for example lung cancer prevention. This is in contrast to promotion which instead targets the whole population to enable them to strengthen their health (Wolfram and Fuchs, 2008). There are three types of prevention; primary, secondary and tertiary. Primary prevention is the prevention of the development of clinical risk factors (Frame et al., 1997) and aims to reduce risk behaviours or risk factors and reduce the chances of acquiring an infectious disease. Examples include influenza vaccination in the UK for older people and healthcare workers or stop smoking services to avert lung cancer or cardiovascular disease. Whereas primary prevention aims to avert disease, secondary prevention aims to identify and treat asymptomatic individuals with existing risk factors or pathologies through screening (Wolfram and Fuchs, 2008). The intention is to achieve better health outcomes than those that would be achieved if disease was identified at a later stage. Tertiary prevention is closely related to medical care, with the intention of preventing the progression of disease, alleviating symptoms and averting disability (Tulchinsky and Varavikova, 2014). Services could include self-management programmes for chronic diseases such as diabetes or chronic kidney disease.

1.2.3 Health Protection

Health protection is the domain of public health that protects the population against communicable diseases and non-communicable environmental hazards (Nicoll and Murray, 2002). These comprise the first two components of the health protection function, with emergency planning and response being the third. Justification for a strong health protection service is provided by the emergence of new infections, exposure to infections abroad, novel chemical threats to health and new vaccinations (Nicoll et al., 2001). The UK Department of Health's *Getting Ahead of the Curve* policy published in 2002 predominantly set out an infectious disease strategy whilst also including health protection as a whole (Department of Health, 2002). It was developed in response to the identification of environmental and infectious threats such as foot and mouth disease and anthrax attacks in the USA. The report advocated the development of the Health Protection Agency (HPA).

The HPA provided expertise and either provided or commissioned services to protect the public in terms of prevention, disease control and surveillance (Roberts and Haworth, 2002). It combined public health laboratory services, microbiology, radiological protection and chemical incident response teams, but also required collaboration with the Environment Agency, Food Standards Agency, the NHS and local authorities (Roberts and Haworth, 2002). The HPA became part of Public Health England (PHE) in 2013 with the aim of developing coordinated services, providing better information and expanding surveillance beyond infectious diseases. PHE produces annual reports such as Sexual and Reproductive Health Profiles to assist local authorities to develop needs assessments and improve the health of their population (PHE, 2016b). They can benchmark performance against similar regions to identify areas for improvement. In Birmingham for example diagnosed human immunodeficiency virus (HIV) prevalence is 2.65 per 1,000 (aged 15-59) compared to 5.83 in Manchester and 0.68 in York (PHE, 2016b).

1.2.4 WHO Essential Public Health Operations

In 2012, the WHO adopted a set of ten Essential Public Health Operations EPHOs (WHO, 2012) (Table 1) underpinned by Acheson's definition. The 2012 European Action Plan (EAP) for strengthening public health services and capacities (WHO, 2012), describes the EPHOs. It was produced to aid the development, monitoring, evaluation and implementation of actions addressing health challenges facing Europe at the time and those anticipated to arise in the future. The EPHOs were developed in consultation with WHO member states and are divided into core functions (items 1-5) and enabling operations (items 6-10) (WHO, 2012). Whilst there are ten items, they are not intended to reflect ten separate services and instead can be combined or adapted to local or changing needs. For example, health promotion and disease prevention are interdependent domains.

Table 1 WHO Essential Public Health Operations

Core Functions of Public Health	
1. Surveillance of population health and well- being	This requires the collection of data to establish the current health of the population and monitor trends over time. The development of health needs assessments based on this data aids the understanding of health status, disease burden and health inequalities. These can then guide the core public health services, which include health protection, health promotion and disease prevention.
2. Monitoring and response to health hazards and emergencies	There should be capacity to rapidly detect, respond to and communicate about health threats which could include communicable diseases and threats in the environment, workplace and the food chain.
3. Health protection, including environmental, occupational, food safety, and others	Health protection requires the technical expertise to identify, assess, manage and communicate about potential risks to the population's health. Risks can include those related to the environment, food safety, air quality and occupational hazards. The workforce should have regulatory mechanisms to protect the population and the ability to monitor compliance.

Core Functions of Public Health	
4. Health promotion, including action to address social determinants and health inequity	Health promotion recognises the impact of political, social, cultural and economic factors on life chances, in addition to genetic, behavioural and environmental determinants of health. Thus, health promotion activities should demonstrate an awareness of these issues and include actions to address them. The domain encompasses the promotion of changes in lifestyle and behaviours in addition to environmental and societal conditions to foster a culture of health and wellbeing. The use of education as a vehicle for promotion and changes to the health system to encourage it are also important.
5. Disease prevention, including early detection of illness	This domain aims to maximise the chances of a positive health outcome. It includes both primary and secondary prevention and both communicable and non-communicable diseases. Actions are delivered by the health sector to individuals in the population to avoid the development of disease e.g. immunisations (primary prevention) or to detect it at an early stage e.g. screening programmes (secondary prevention). Primary prevention in particular aims to reduce the impact of health inequalities and behavioural determinants of health.

Enabling Operations for Public Health	
6. Ensuring governance for health and well-being	It is essential that governance structures exist that allow the monitoring and evaluation of the implementation of services. The results can determine whether resources are allocated appropriately and inform future policy development.
7. Ensuring a sufficient and competent public health workforce	Developing a multidisciplinary workforce is critical to a public health system. Given the multitude of factors determining health, expertise beyond those medically trained in public health is required. The EAP recommends that the workforce consists of public health specialists (traditional public health occupations e.g. communicable disease prevention professionals or those involved in promotion, prevention or protection), health professionals (those working in the health service without an explicit public health function, e.g. general practitioners (GPs)) and non-health sector professionals (those who's decisions have an impact on health e.g. government policy makers).
8. Ensuring sustainable organisational structures and financing	To ensure that services are efficient, integrated and designed to achieve long-term health benefits, sustainable and adequately financed organisational structures need to be in place. Funding of services should be sufficient to maximise health gains in the present and future. The structures should be organised at a local, regional and nation level with collaborations with agencies outside of the public health sector also established.
9. Advocacy, communication and social mobilisation for health	Communicating public health information aims to inform, influence and motivate individuals to improve their health status whilst countering the promotion of behaviours leading to unhealthy lifestyles and use of hazardous products. Improving health literacy and facilitating the development of accessible information for multiple audiences are also key functions for improving population health.

Enabling Operations for Public Health	
10. Advancing public health research to inform policy and practice	Research priorities should be tailored to the population's public health challenges and then be used to support evidence based policy making. In addition to aiding the development of new services, research should assist in guiding methods for successful implementation, monitoring and evaluation. This also encompasses evaluating the long-term cost-effectiveness of services and consideration of the most appropriate methodology for conducting public health research.

1.3 Public Health in the UK

At a national level in England, PHE is an executive agency of the Department of Health and Social Care. PHE provides evidence based expertise and support to protect and improve health and wellbeing, reduce inequalities and support a strong economy (PHE, 2019a). A strong economy is facilitated by the population living long and productive lives, thus good health is an asset. Public Health Wales shares similar aims and responsibilities as PHE but is an organisation within NHS Wales (Public Health Wales, 2018). Each health board employs a Director of Public Health who is supported by the national infrastructure and manages local public health staff. In Scotland, the development of Public Health Scotland as health board within NHS Scotland is currently under consultation (Scottish Government, 2019). At present there is no national public health body, with local authorities and health boards responsible for public health through Integration Authorities. In Northern Ireland, the Public Health Agency (PHA) is an agency within Northern Ireland's Department of Health, and provides similar services to as PHE and Public Health Wales (Public Health Agency, 2019). Hereafter, public health is discussed in the context of England.

The UK Faculty of Public Health has identified six functions that local authorities are responsible for. These largely align with the core functions and enabling operations adopted by the WHO (Rechel and McKee, 2014) and comprise health protection, health improvement, health services, public health intelligence, academic public health and workforce development (Faculty of Public Health, 2014). Changes to the public health system in 2012 transferred responsibilities for public health from Primary Care Trusts (PCTs) to local authorities. The recognition that health and wellbeing are influenced by the environment, social, economic and cultural factors led to the belief that local governments are best placed to influence the broad factors that drive health and wellbeing in the population. Local responsibility allows the identification of, and response to, the immediate needs of the community but at a higher level, accountability is still placed with the government to improve and protect the health of the population. This change did not however abdicate responsibility for public health from the NHS and other partners such as the voluntary sector. Rather, these agencies must still work together to share responsibility for the health of the nation. Health and Wellbeing Boards (HWB) were also created to enable key members of the healthcare system to work together to improve health and wellbeing in their local populations.

Published in 2010, the Marmot Review (Marmot, 2010) detailed health inequalities facing England and policy objectives to reduce these inequalities. If everyone were to enjoy the same number of healthy years of life, productivity would increase and NHS costs for treating illnesses resulting from inequalities in society would decrease. Determinants of health are complex and intertwined, thus objectives to reduce inequalities are related and dependent on an individual's development across the lifespan. The findings from the review are of direct relevance to the remit of public health and contribute to performance indicators endorsed by the Department of Health.

The first policy objective was to give every child the best start in life (Marmot, 2010). Reducing inequalities in the foundations of their development will allow children to have the best opportunity to develop successfully across their life course. These areas include children's physical and emotional health, social skills and cognition and language abilities. Providing high quality maternity services and early years services has a long-term impact on children's health and wellbeing with early intervention seen as particularly important (Ramey and

Ramey, 1998). A second priority is to reduce inequalities in education, with education seen to be a key factor in physical and mental health, income, employment and quality of life (Lahelma et al., 2004). Related is creating fair and good employment. Unemployment contributes to poor health and being in good employment is protective of health (Sturm and Gresenz, 2002). Ensuring individuals can receive an income that would ensure a healthy life is another priority. Living standards contribute enormously to health and having the right nutrition, housing, medical care are critical (Barker and Osmond, 1986). Providing an environment that encourages and supports healthy lives allows individuals to thrive socially and improve their health and wellbeing. Examples include reducing social isolation and increasing green spaces (Lee and Maheswaran, 2011). The final priority recommendation was strengthening prevention and early identification of ill health. Key health behaviours such as smoking and obesity play an important role in chronic diseases in which inequalities in incidence are observed in society (Vallejo-Torres and Morris, 2010).

1.3.1 Public Health Outcomes Framework

The Public Health Outcomes Framework was developed by the Department of Health through a consultation with stakeholders and the public and describes what the public health system aims to achieve (Department of Health, 2013b). The two overarching priorities are increasing healthy life expectancy and reducing differences in healthy life expectancies between communities. It is acknowledged that these may take many years to achieve and consequently, a number of indicators have been developed to enable monitoring of progress to achieve these goals. The indicators cover four domains: improving wider determinants of health, health improvement, health protection, healthcare public health and preventing premature mortality. Within these domains more specific metrics are specified which provide objective measures of each local authority's performance.

1.3.1.1 Improving wider determinants of health

The wider determinants of health that local authorities are expected to consider largely reflect those prioritised by the Marmot Report (Marmot, 2010). These are social determinants of health that are distinct from biological and genetic factors and include living and working conditions, culture and the environment (Bambra et al., 2009). To achieve indicators such as reducing violent crime, homelessness and school readiness, requires collaboration between many partners within local authorities, such as schools and the police. The establishment of HWBs aims to facilitate these partnerships.

1.3.1.2 Health improvement

There is no agreed upon, bounded definition of health improvement (Abbott et al., 2002). Interviews with individuals working in Primary Care Organisations (now Clinical Commissioning Groups (CCGs)) found ambiguity surrounding what particular health needs and requirements should be improved and there was also uncertainty regarding the commissioning responsibilities of the NHS and Public Health for health improvement. More recently, the Public Health Outcomes Framework has clarified the scope of such responsibilities. According to the Public Health Outcomes Framework (Department of Health, 2013b), health improvement aims to facilitate healthy choices and the maintenance of healthy lifestyles. Indicators include breastfeeding, excess weight in children and adults, alcohol related admissions to hospital and cancer screening. Health improvement programmes commissioned by local authorities address these indicators although the NHS delivers some services such as screening.

1.3.1.3 Health protection

The role of health protection is to ensure the population is protected from health threats in addition to reducing inequalities (Department of Health, 2013b). Indicators in this domain include chlamydia diagnoses, mortality attributable to air pollution and vaccination coverage. Like health improvement indictors, collaboration between several agencies such as the NHS and local transport planning is essential to achieving these indicators.

1.3.1.4 Healthcare public health and preventing premature mortality

This domain is the broadest and is delivered by all areas of public health. It intends to reduce the number of people living with preventable ill health and premature deaths from conditions such as liver and cardiovascular disease and cancer. Preventable ill health includes tooth decay, avoidable sight loss and suicide. Many of the indicators are also included in the NHS Outcomes Framework domain; Preventing People from Dying Prematurely, thus responsibility is shared and collaboration between agencies is critical.

1.3.2 The Funding of Public Health

Since 2013 the Department of Health has provided local authorities with a ring-fenced grant to improve public health. The distribution of funding allocated to local authorities is calculated based on a formula which aims to reflect the local population’s needs and has three components addressing mandatory, non-mandatory and substance misuse services (Department of Health, 2013a). Developed by the Advisory Committee on Resource Allocation (ACRA), weighted populations are calculated based on resident populations adjusted for relative need, geographical variation in the cost of providing services, age and gender. For drugs services, populations are also adjusted for outcomes. Local authorities are obliged to commission some non-discretionary services such as sexual health provision, although the remainder of the budget can be allocated to fund services to improve health and wellbeing in the community.

There are a number of public health services that are provided by NHS England, funded by a ring-fenced budget paid to NHS England by the Secretary of State. This agreement reflects the joint purpose of these agencies to protect and improve the population's health. In 2018-2019 these services included immunisation programmes, screening programmes, Child Health Information Systems, public health services for adults and children who are detained and sexual assault services (Department of Health, 2015).

1.3.3 Providers of Public Health Services

In 2014 it was estimated that in the UK the public health workforce comprises between 36,000 and 41,000 workers. This figure refers to staff fulfilling a 'core' public health role, defined as "*All staff engaged in public health activities who identify public health as being the primary part of their role*" (Centre for Workforce Intelligence, 2014). The eleven roles included in this remit are detailed in Table 2 and exclude individuals who may promote health as only part of their job, for example teachers or GPs. This wider workforce potentially encompasses up to 15,000,000 who influence health and wellbeing (PHE, 2016a). Since the Health and Social Care Act (2012) transitioned public health responsibilities to local authorities, the primary employers of specialist public health staff are local authorities or PHE. Some staff are still placed within the NHS and those practicing academic public health are based in universities. Table 2 provides a brief description of each role, where they are predominantly employed and the public health functions that their role addresses.

1.3.3.1 Specialists and consultants

Public health specialists and consultants have undertaken extensive speciality training in public health and occupy senior roles in local authorities, PHE and the NHS. They have a high level of expertise, with the intention of influencing the health of whole communities by using science and information to determine priorities and engage with partners to implement policies. They can be involved in delivering all public health functions although may choose to specialise in one or more. Specialists must be registered either with the General Medical Council (GMC) or the UK Public Health Register (UKPHR). Directors of public health are public health consultants employed in a statutory position in most local authorities. They take responsibility for determining the overall objectives for public health locally and sit on Health and Wellbeing Boards, providing public health input.

1.3.3.2 Public health managers

Public health managers are primarily employed in local authorities and are tasked with delivering projects and programmes in their locality, working with external partners and local authority staff on initiatives such as sexual health. They are usually although not necessarily qualified in public health at Masters level.

1.3.3.3 Public health scientists and public health academics

Public health academics are employed in universities and fulfil lecturing and research roles in public health. Public health scientists also conduct research and are mostly employed by the NHS and PHE, typically fulfilling a health protection role such as communicable disease control. They support the identification and delivery of public health objectives.

1.3.3.4 Intelligence and knowledge professionals

Intelligence and knowledge professionals work at a local and regional level and are located within PHE and local authorities. They are responsible for obtaining, managing, analysing and disseminating primary and secondary data relating to health, social and demographic characteristics of their populations. They also have a role in developing data for inclusion in performance indicators, therefore their expertise is used across all public health functions.

1.3.3.5 Public health and school nurses

Public health nurses have completed additional training to specialise in an area of public health such as infection control. They predominantly have health improvement and health protection roles. They are employed by local authorities, PHE and the NHS. In contrast school nurses are solely employed by local authorities and work in schools with a health improvement and health protection function. They coordinate and implement strategies such as the Healthy Child Programme (HCP). Both school nurses and public health nurses are professionally registered with the Nursing and Midwifery Council (NMC).

1.3.3.6 Health visitors

Health visitors work with primary healthcare services to assess the health needs of the community, although since 2015 they have been commissioned by local authorities. They are trained nurses or midwives, registered with the NMC and in possession of higher qualifications in child health, health promotion and education.

1.3.3.7 Environmental health professionals

Environmental health professionals have undertaken accredited training in environmental health and are regulated by the Chartered Institute for Environmental Health (CIEH). They are employed mostly in local authorities and their role is to improve, monitor and enforce health standards at a local level. This can be in several different domains, such as food safety or housing for example.

Table 2 Public Health Workforce

	Primary employer				**Setting**	**Main public health function**					**Role**
	LA	PHE	NHS	HE		HI	HP	HC	HK	APH	
Public health consultants and specialists	✓	✓	✓		Policy	✓	✓	✓	✓	✓	Influence health of entire communities
Directors of public health	✓				Policy	✓	✓	✓	✓	✓	Determine objectives of PH locally
Public health academics		✓		✓	Research					✓	Focus on PH research
Public health managers	✓				Community	✓	✓	✓	✓	✓	Deliver projects or programmes
Public health scientists		✓	✓		Research, Delivery		✓				Scientific role supporting PH objectives
Intelligence and knowledge professionals	✓	✓			Research, Delivery				✓		Data analysis, informatics and presentation of PH information
Public health nurses	✓	✓	✓		Delivery	✓	✓				Specialists in PH nursing
Health visitors	✓				Community	✓	✓				Assess the health needs of individuals, families and the community

	Primary employer				Setting	Main public health function					Role
	LA	PHE	NHS	HE		HI	HP	HC	HK	APH	
School nurses	✓				Community	✓	✓				Deliver PH functions in schools
Public health practitioners	✓	✓			Community	✓	✓	✓	✓	✓	Work across the PH system to deliver PH programmes
Environmental health professionals	✓				Community	✓	✓				Work to improve, monitor and enforce public and environmental health standards

Local authority (LA), Public Health England (PHE), National Health Service (NHS), higher education (HE), health improvement (HI), health protection (HP), healthcare commissioning (HC), health intelligence/knowledge (HK), academic public health (APH), public health (PH)

1.4 School Based Public Health Provision

1.4.1 The School Role in Delivering Public Health Functions

As children spend much of their waking time in schools, schools are an ideal location for developing their understanding of healthy lifestyles and delivering public health interventions. Until September 2020 there is no national curriculum requirement regarding health and wellbeing (Department for Education, 2019a), although the provision of education surrounding public health issues is addressed in the framework on which schools are inspected (OFSTED, 2019). In addition, school nurses can aid the delivery of certain public health initiatives and teaching.

1.4.2 School nurses

School nurses are qualified nurses with an additional specialist public health qualification. The aim of the service is to improve the health and wellbeing of young people and the role of schools includes coordination and delivery of public health interventions to school aged children. This includes contributing to; reducing obesity, promoting good emotional and sexual health and preventing substance misuse. The school nurse's remit is however much broader, also covering safeguarding, transitioning through life and working with communities. School nursing services are commissioned by local authorities and are funded by the public health grant (PHE, 2014). The provision of school nurses is not mandated, however non-statutory guidance on their role is provided by the Department of Health and PHE. School nurses are key contributors to the provision of the HCP, Personal Social and Health Education (PHSE) in schools, in addition to collecting data for the National Child Measurement Programme (NCMP) (NHS Digital, 2018).

1.4.3 Healthy Child Programme

The Healthy Child Programme (HCP) for children aged 5-19 is the government's early intervention and prevention public health programme (Department of Health, 2009). It provides a good practice framework for services to enhance young people's life chances. At a school level, school nurses lead, coordinate and deliver the majority of aspects of the programme. These include immunisations, supporting mental wellbeing, safeguarding, substance misuse, maintaining healthy weight and targeted support for those identified as at risk. As the HCP encompasses many aspects of a child's health, several providers are involved in the implementation of the services such as CCGs. Services provided in schools are however funded by local authorities.

1.4.4 The National Child Measurement Programme

The NCMP was established in 2005 and measures the height and weight of children in Reception and year 6 at state maintained schools and academies in England (NHS Digital, 2018). The primary aim of the NCMP was to provide health surveillance data on children's weight status however it has also been used to provide feedback on a child's weight status to their parents. It is intended that feedback will result in positive health related behaviour change. Local authorities are mandated to collect the data and school nurses are mostly responsible for this. Data is provided to NHS Digital and is available at a local level.

1.4.5 OFSTED

The Office for Standards in Education, Children's Services and Skills (OFSTED) regulates and inspects schools (OFSTED, 2019). Schools are assessed on several domains and are rated on a scale of 1 (outstanding) to 4 (inadequate). In the domain of personal development, behaviour and welfare, schools must demonstrate that students are able to keep themselves healthy in order to achieve a rating of outstanding. This includes making informed choices about eating, fitness and mental wellbeing (OFSTED, 2019). A school may be rated inadequate if students are unable to demonstrate such understanding. When judging a school's effectiveness in leadership and management, OFSTED also take into consideration the use of the Physical Education (PE) and Sport Premium (described in section 1.4.7) and its impact on outcomes for students.

1.4.6 Personal Social and Health Education (PSHE)

PHSE is currently a non-statutory subject, although the majority of schools elect to include it in the curriculum. In September 2020 PHSE will be succeeded by a mandatory Relationships and Health Education curriculum, with schools encouraged to adopt it from September 2019 (Department for Education, 2019a). It emphasises the importance of good physical health and mental wellbeing, schools must have good reason to depart from the guidance. PHSE contributes to schools' statutory responsibilities to promote personal wellbeing and health and the government recognises its importance in tackling public health issues (OFSTED, 2012). In a review of PHSE provision, OFSTED recommended that schools should ensure that teachers are trained in PHSE; and that schools track student's engagement in extra-curricular activities that develop their personal skills.

The prioritisation of wellbeing in the new PHSE curriculum suggests it is an important consideration for schools. Indeed, the guidance notes the interdependence of physical health and mental wellbeing, with physical exercise and time outdoors being cited as important contributors to maintaining wellbeing (Department for Education, 2019a). A PHE review has examined the link between wellbeing, health and academic attainment (PHE, 2014). Literature identified showed that better emotional wellbeing in primary school was a significant predictor of academic progression. Significant correlations were observed between pupil engagement, academic attainment and pupil wellbeing. (Gutman and Vorhaus, 2012). Physical activity has also been associated with academic attainment (Booth et al., 2014) and wellbeing (Gavin et al., 2012; Ussher et al., 2007). A systematic review of the impact of school based physical activity interventions on wellbeing did not find conclusive evidence of effect, although no studies measured children's wellbeing from their perspective (Rafferty et al., 2016).

There is considerable discussion regarding the definition of wellbeing (Dodge et al., 2012; McLellan and Steward, 2015) and whether it can indeed be summarised as a concise and bounded concept which can be quantified. Wellbeing literature spans psychology, sociology, medicine and economics (McLellan and Steward, 2015), culminating in many conceptualisations and outcome measures. Furthermore, it has been highlighted that many definitions are in fact descriptions, usually of what reflects or contributes to good or poor wellbeing (Dodge et al., 2012).

The absence of a definition, or perhaps the implicit assumption of its existence might explain the lack of clarity in many UK policy documents and reviews. Ofsted do not define it (OFSTED, 2019) and The Local Government Association Peer Learning Programme report: *Improving children and young people's mental health and emotional wellbeing* (Local Government Association, 2018) refers to it in the context of 'health and wellbeing', 'mental health and wellbeing' and 'emotional wellbeing'. References to the NHS and Child and Adolescent Mental Health Services (CAMHS) indicate that wellbeing was conceptualised medically, on the basis that facilitating good mental health/wellbeing will result in better life outcomes and reduced NHS service use. In addition, a recent Department for Education review of mental health and wellbeing provision in schools does not distinguish between the terms and does not define what is meant by wellbeing (Department for Education, 2018a). The review asserts that schools should nurture and support mental health and wellbeing, but the documentary analysis concludes that schools do not have formal polices in place to achieve these goals. More encouragingly, a recent PHE report on universal approaches to improving children and young people's mental health and wellbeing (PHE, 2019b) provides an infographic demonstrating the overlapping and distinct factors contributing to these concepts, although despite focusing on subjective wellbeing, the report does not define it.

The sociological perspective largely uses objective indicators to describe wellbeing, however relying on external indicators (e.g. financial resources) distances the concept from a subjective one (Fegter et al., 2010). Individuals with access to the same material resources or in the same life circumstances are unlikely to have similar wellbeing for example (Fegter et al., 2010). The psychological perspective utilises the term subjective wellbeing, incorporating personal feelings and emotions (McLellan and Steward, 2015). One definition of subjective wellbeing suggests it consists of life satisfaction and positive and negative moods and emotions (Diener and Suh, 1997). Additional factors also cited in the literature include; the ability to fulfil goals, happiness, self-esteem and purpose in life (Pollard and Lee, 2003). Asking about an individual's satisfaction with their life involves a personal judgement about what a good life is, which could be an efficient method of assessing their wellbeing (Layard, 2016). This narrow question synthesises wellbeing to one dimension however, ignoring the other components. One conceptualisation attempts to solve the problem of describing, not defining wellbeing. It refers to a balance point between an individual's resource pool and the challenges they face (Dodge et al., 2012). When their psychological, social and physical resources are sufficient to meet a particular challenge, their wellbeing is stable.

QOL and wellbeing are both subjective concepts. Like the controversy surrounding the definition of wellbeing, there is also discussion whether QOL and wellbeing are synonymous. It is argued that QOL is one component of wellbeing for example (Fegter et al., 2010). The WHO defined QOL as a person's perception of his/her position in life within the context of the culture and value systems in which he/she lives and in relation to his/her goals, expectations, standards, and concerns (WHO, 1994). It has been noted that this definition is strikingly similar to the concept of subjective wellbeing (Camfield and Skevington, 2008). The need for an outcome measure that can be used for treatments where the objective was not lifesaving has led to a focus on measuring health-related quality of life (HRQL). Due to the goals of the health service, only aspects of quality of life that could be impacted by health or health care were considered relevant (Peasgood et al., 2014). A way of distinguishing health-related aspects identifies HRQL focuses on deficits in functioning (e.g. pain, depression), in contrast to wellbeing that focuses on assets in functioning, such as positive emotions and resources (Centers for Disease Control and Prevention, 2018).

In this thesis, wellbeing is conceptualised as a multifaceted concept not limited to health. Drawing on Dodge et al. (2012), an individual's wellbeing can be impacted by many factors and the interrelationship between them results in an individual's level of wellbeing. These factors include self-esteem, family circumstances, social relationships and health status for example.

1.4.7 PE and Sport Premium

A legacy of the London 2012 Olympic Games was a commitment from the Government to create a sporting habit from a young age and to encourage the development of healthy and active lifestyles (Department for Education, 2015) . Sports Premium funding is allocated to primary schools to make additional and sustainable improvements to the quality of sports they offer. Examples include; employment of specialist sports coaches and increasing the availability of extra-curricular activities. Funding is provided by the Department of Education and Department of Health and Social Care to local authorities, who distribute allocations to schools. Since 2018, revenue from the Soft Beverage Industry Levy (HM Revenue and Customs, 2016) also contributes, with Sports Premium funding reportedly doubling (HM Treasury, 2018). In 2018/2019, each local authority maintained primary school received £16,000, with academies receiving £16,000 and an additional £10 per pupil (Education and Skills Funding Agency, 2018). Schools must publish how their budget is spent online and OFSTED also include an evaluation of the Sports Premium spending in their assessment. In order to demonstrate the Sport Premium's effectiveness, tools are available to measure and report the impact the funding has in terms of participation, attainment and sustainability.

1.4.8 School food standards

School Food Standards are mandatory legislation aiming to ensure that children in England receive nutritious and high quality meals at school, whilst additionally promoting good nutritional health (Department for Education, 2019c) .The current standards were introduced in 2015 and dictate the frequency with which different food groups can be provided, such as limiting deep fried food to two lunches per week and prohibiting confectionary across the whole school day. As the guidance only applies to food provided by the school, packed lunches prepared at home are not covered. Another policy introduced in 2014 entitled all children in reception, year 1 and year 2 to a free school lunch (Department for Education, 2019c). Funding is provided by the Department for Education and each meal attracts funding of £2.30 per eligible pupil. This policy was informed by pilot studies demonstrating that children performed better academically and ate more healthily, with the poorest children benefitting most. To-date, there has been no quantitative evaluation of school food standards (Rose et al., 2019) although a National Institute for Health Research (NIHR)-funded study is on-going at the University of Birmingham (University of Birmingham, 2019).

1.4.9 Primary school accountability

Exploring the provision of public health in schools has highlighted a number of mandatary services and regulations related to health to which must be provided or adhered to. Within the inspection framework, OFSTED assess the quality of education delivered by schools (OFSTED, 2019). Whilst they do allow some flexibility, the basic curriculum in local authority-maintained schools must include religious education and age-appropriate relationship and sex education. From September 2020 this will include 'health education' (OFSTED, 2019). The quality of education is rated on the scale previously described (Section 1.4.5), ranging from outstanding to inadequate and contributes to the overall assessment of the school. An understanding of how to keep healthy and active also persists in the Personal Development domain. While not an explicit inspection domain, the use of the PE and Sports Premium (Section 1.4.7) is also considered when judging the effectiveness of a school. While these are important, schools have other significant responsibilities that contribute to the monitoring of their performance.

The prioritisation of public health therefore must be considered in the context of school's other accountabilities.

There are three outcomes or assessments for which UK primary schools are accountable (Department for Education, 2017a). These are attainment and progress measures (published data), floor standards (minimum expected attainment) and school inspections. If schools cause concern on any of these measures, they are eligible for formal action. This action is undertaken by the Regional School Commissioners (RSCs), who are required to intervene on behalf of the Secretary of State (Department for Education, 2016b). Intervention is deemed necessary under the following circumstances. The first is where a school has been judged by OFSTED as requiring significant improvement or requiring special measures. The second situation is when schools have failed to comply with warning notices about unacceptable performance, including results below floor standards. The final situation is when a school has been defined as coasting. The Secretary of State has a duty to issue an academy order to local authority maintained schools judged by OFSTED to be failing. The RSC is responsible for ensuring the school quickly becomes a sponsored academy. If an existing academy is judged to be failing, the RSC has the power to terminate the funding agreement and identify a new academy sponsor. Termination of the agreement is not a duty however, as in some circumstances it may hinder improvements.

Attainment and performance data against which schools are judged is derived from exams and teacher assessments at the end of Key Stage 2 (year 6, aged 10-11) (Department for Education, 2018c). These are used to inform performance tables which are released annually. More specifically, the measures include; the percentage of pupils achieving the expected or higher standard on a composite measure of reading, writing and maths, pupils' average scaled scores in reading and mathematics, and pupils' average progress in reading, writing and maths. The floor standard is the minimum pupil attainment or progress that the government expects a school to meet. Progress measures capture the progress pupils make between Key Stage 1 and Key Stage 2. To be above the floor, 65% of pupils must meet the expected standard in reading, writing and mathematics or the school should achieve sufficient average progress scores in all three subjects. Warning notices can be issued by local authorities and RSCs when there are concerns about a school's performance standards. Low standards of performance are those below what pupils would be expected to attain in the circumstances, standards below those previously attained and standards below those attained at comparable schools. Unlike action taken against inadequate schools, the local authority and RSC are more flexible and will consider contextual factors and their capacity to improve. Warning notices set out the actions that the governing body is expected to undertake to address the concerns, the period of time in which they should comply and the intended intervention if they fail to comply. It may be stipulated that a school would require additional support, require additional or alternative governance or that an academy order is necessary.

Coasting schools are schools where, over time, pupils do not fulfil their potential. The metric uses the same attainment and progress measures as the floor standard and is based on three years of data. Criteria to define a coasting school used in 2018 were; fewer than 85% of year 6 pupils achieved less than the expected standard in 2015, 2016 and 2017 in reading, writing and mathematics and average progress was less than -2.5 (reading and mathematics) and -3.5 (writing) (Department for Education, 2018c). Action taken against coasting schools is again quite flexible. RSCs will consider contextual factors such as characteristics of the school cohort, capacity to improve and wider achievements. It could be that no further action is taken following a notification of their coasting status. Potential interventions are similar to those offered to schools with unacceptable performance. If the school is already an academy, the same factors will be considered and the funding agreement could be terminated.

1.5 Conclusion

Public health is a multifaceted concept, which appears to be defined by its key functions. Acheson (1988) widely adopted definition provides a concise description of core functions that facilitate and protect good health, namely health promotion, prevention and prolonging of life. The WHO (Rechel and McKee, 2014) expands these functions, including monitoring and surveillance of health, accompanied by factors that enable these, such as a competent workforce. Public health provision which local authorities in England are funded to provide largely align with the WHO core functions and enabling operations. Funding is distributed to local authorities who allocate resources according to their population's needs. While services such as school nurses and the NCMP are funded through these means, other ways in which public health is provided in a school setting may not be funded through local authority public health budgets (e.g. childhood obesity interventions or wellbeing initiatives) and may be funded from school resources. There are alternative sources of funding to be drawn upon, such as the Sports Premium, but many mandated public health responsibilities of schools might be seen as cost-neutral, such as imparting knowledge about health. Passive teaching might not be an effective approach however, with other more novel initiatives perhaps being required to achieve an understanding of a healthy lifestyle for example. There is also incompatibility with the public health workforce, again with the exception of school nurses. This is despite schools being cited as an ideal setting to deliver public health interventions for children.

This chapter has described the way in which public health functions are delivered in schools and schools' other responsibilities. Health promotion, prevention, protection and academic attainment are intertwined, creating a complex but ideal setting to deliver public health interventions. Schools will consider new public health initiatives alongside other initiatives or responsibilities that serve to achieve the metrics upon which they are judged, which are predominantly academic attainment and children's progress (OFSTED, 2019). Allocating time as well as tangible costs to public health interventions in schools therefore requires the assessment of the resources required and potential benefits, which is an opportunity to use economic evaluation. Given the isolation of schools from the NHS and local authority public health funding in addition to competing priorities, methods advocated for economic evaluation in healthcare are perhaps challenged. This is initially discussed in Chapter 2, where methods for economic evaluation are described. This thesis also explores the appropriateness of 'traditional' methods using qualitative methods and is reported in Chapter 5.

While responsibility for public health is dispersed across sectors, this chapter has highlighted that schools have an important role for delivering interventions to school-aged children. The economics of public health interventions delivered in a school setting are the focus of this thesis and appears to be an under researched area. Chapter 2 introduces economic evaluation, discusses challenges in its application to public health and how economics has been used in schools previously.

CHAPTER 2 ECONOMIC EVALUATION[1]

2.1 Introduction

This chapter firstly aims to describe the theory behind economic evaluation and the possible approaches to conducting a health economic analysis. The choice of cost and outcome data is explored, with a specific focus on implementation costs. Finally, how economic evaluation methods have been applied in a school setting is discussed.

2.2 Theoretical underpinnings of economic evaluation

The rationale for conducting an economic evaluation of a healthcare intervention lies in the belief that health can be conceived as an economic good. As a scarce resource, how healthcare is distributed, who produces it and who pays for it are pertinent issues. Whether healthcare is the same as a consumer good and therefore whether it should be treated similarly in terms of provider and consumer behaviour has been discussed (Arrow, 1963). The primary distinguishing feature of health from other goods is uncertainty, and most other features stem from this. Illness is unpredictable, as are treatment responses and consequences (Arrow, 1963). Demand for treatment is variable and with the exception of preventative interventions is dependent on the occurrence of disease (Arrow, 1963). The providers of medical care have professional responsibilities and their decision making should be guided by clinical need not financial concerns (Arrow, 1963). A knowledge imbalance also exists, with the patient dependent on the medical professional's recommendations (Arrow, 1963). In most circumstances the

[1] Note on chapter: Small sections of this chapter have been submitted to the University of Birmingham as part of an essay towards an MSc in Health Economics and Health Policy (Breheny, 2013). Where appropriate, this essay has been referenced to indicate the use of previously submitted material.

patient is unable to test a product prior to its use and the quality of the product is difficult to determine even after its use. These differences would suggest that traditional methods and models of economic behaviour are unsuitable to be applied directly to the evaluation of health interventions.

2.2.1 Welfarism and extra-welfarism

Welfare economics is the traditional approach used in economics (Culyer, 1989). It is based on individualism, with social welfare being a function of individual welfare only. Those affected by social choices are the only views relevant, and individuals are the best judges of their own welfare. This excludes views of experts and other informed members of society. In addition, individual utilities are a function of the commodities consumed by individuals (Culyer, 1989). Another characteristic is that when judgements need to be made about the superiority of different situations, only utility aspects of decisions are considered. This excludes caring externalities for example. Within welfarism the Pareto principle is the value judgement used when deciding whether a choice is optimal. This is the proposition that social welfare increases only if the welfare of a member of society increases and no one is worse off. Pareto optimal decisions can be ranked in order of the potential utility obtained. In this sense, utility can be seen as the level of satisfaction obtained from the consumption of goods and services and an individual's aim is to maximise their utility (Culyer, 1989).

Extra-welfarism is a more recent approach and is seen as a more relaxed approach than welfarism (Culyer, 1989). The predominant departure from welfarism is that outcomes other than utility such as equity, quality of utility or capabilities can be considered (Brouwer et al., 2008). This is the 'extra' in extra welfarism. Another difference is that the individual affected need not be the primary source of valuation. Who makes value judgements is up to the decision maker. Extra-welfarism also permits the weighting of benefits according to characteristics of those receiving it. This is useful when equity might be a consideration. Finally, extra-welfarism allows interpersonal comparisons. The ability to quantify outcomes enables the monitoring and evaluation of populations, whilst also directly responding to decision makers' needs.

2.2.2 Equity and Efficiency

The efficient use of resources is a phrase fundamental to economic evaluation. Whilst the broad concept of efficiency is to maximise the gains from resources available (Knapp, 1984) there are several different types, or definitions of efficiency (Morris et al., 2007). These are technical, productive, allocative and social efficiency (Morris et al., 2007; Palmer and Torgerson, 1999; Shiell et al., 2002). The different approaches to conducting economic evaluation that will be presented later mostly satisfy at least one of these definitions. The first type is technical efficiency and it is concerned with how best to achieve an objective (Shiell et al., 2002), considering only resource inputs and not costs. A technically efficient decision is observed when the best outcome is achieved from a given set of inputs or the fewest inputs possible are used to produce an amount of output (Morris et al., 2007). For example, delivering two healthy cooking sessions may have the same effect on dietary habits as three sessions. Two sessions are therefore more technically efficient. Productive or technical efficiency allow the direct comparison of alternative interventions targeting the same health outcome, that have different costs or effectiveness (Palmer and Torgerson, 1999). This type of efficiency aims to maximise the health outcome for a given cost, or to minimise cost for a given outcome. An example is comparing interventions for smoking cessation which achieve similar quit rates, where the lowest cost intervention would be the most technically efficient. The final type of efficiency goes beyond allocating resources to achieve a common improvement in health. Allocative efficiency enables resource allocation decisions that are both productivity efficient and distribute resources efficiently in society. This allows the comparison of interventions with different health outcomes and outcomes outside of health in order to maximise the health of society (Palmer and Torgerson, 1999). This is closely related to the Pareto principle and is sometimes also referred to as Pareto efficiency.

A concept that also arises when deciding how to distribute healthcare resources is equity. This is particularly relevant to the social determinants of health and the aims of public health in reducing health inequalities. The use of the term health inequalities in this thesis is synonymous with health inequities as this is the language mostly used within public health (PHE, 2018b). There is a distinction between the two terms however (Braveman and Gruskin, 2003). Equity in health has been defined as equal access to available care for equal need, equal utilization for equal need and equal care for all (Whitehead, 1991). Health inequities are disparities that are socially unfair and unjust (Braveman and Gruskin, 2003). The difference between health inequalities and health inequities is that inequities are unjust or unfair, for example between socially advantaged and disadvantaged social groups. Inequalities could be biologically plausible and legitimate such as males having a lower life expectancy than females. Another distinction within equity is horizontal and vertical equity. Horizontal equity refers to the equal treatment of equals and vertical equity is the unequal but equitable treatment of unequals, which can be conceptualised as positive discrimination (Mooney, 2000). Vertical equity is pertinent to public health, where narrowing inequalities is a core aim.

The Whitehead (1991) definition of equity previously described refers to access, utilisation and equal care for equal need. What is meant by need is a debated issue however (Culyer and Wagstaff, 1993). Need could suggest that those in worse health need more than those in better health, or those with a similar health state need the same (Le Grand, 1978). Another conceptualisation is that need is the resource required to achieve a health goal or the goal is important enough to justify the use of the resource. There must be a capacity to benefit and the resource is required to attain the health improvement (Culyer and Wagstaff, 1993). A simple definition is the amount of healthcare required to attain the same level of health, or equal health. A further definition of need is the amount required to achieve the maximum health improvement possible or to exhaust an individual's capacity to benefit from healthcare (Culyer and Wagstaff, 1993). In regard to equity of access, it could be defined as equality in the opportunity to use health services (Mooney, 1983) or equity in the time and costs incurred by the population in obtaining healthcare (Le Grand, 1978). Ensuring equity in health, achieving the targets of reducing health inequalities and the importance of efficiency in resource allocation mean that economic evaluations within public health are challenging. This will be explored further following a description of methods of economic evaluation.

The different approaches to economic evaluation will now be described. Each method has a theoretical foundation in one or more of the theoretical positions previously discussed.

2.3 Types of Economic Evaluation

As economic resources are finite, decisions have to be made regarding which healthcare programmes or interventions should be funded (Drummond et al., 2005). Economic evaluations are a way in which alternatives can be compared to inform such resource allocation decisions. The definition of an economic evaluation is as follows: *the comparative analysis of alternative courses of action in terms of both their costs and consequences* (Drummond et al., 2005). There are three methods of economic evaluation that satisfy this, namely cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA). A further method that only reports outcome and cost data is a cost consequence analysis (CCA) and a description of costs only is a cost analysis. Where there is evidence that consequences are equivalent, a cost-minimisation analysis (CMA) could also be considered.

2.3.1 Cost-minimisation analysis

CMA is conducted when all relevant outcomes are not significantly different and the least costly option is the most efficient use of resources (Drummond et al., 2005). In contrast to cost analysis, CMA requires that both the costs and consequences of the alternatives have been evaluated and there is evidence of similar consequences. Critics of CMA argue that using significance testing to determine equivalence without the analysis of uncertainty is inappropriate however (Briggs and O'Brien, 2001).

Table *3* presents the different methods and the measure of costs and consequences used.

Table 3 Types of Economic Evaluation

		Measurement /valuation of costs	**Measurement /valuation of consequences**
Cost minimisation analysis	CMA	Monetary units	None (consequences are equivalent)
Cost analysis		Monetary units	None
Cost consequence analysis	CCA	Monetary units	Any appropriate outcome
Cost-effectiveness analysis	CEA	Monetary units	Natural units (e.g blood pressure reduction)
Cost-utility analysis	CUA	Monetary units	Healthy years
Cost-benefit analysis	CBA	Monetary units	Monetary units

2.3.2 Cost consequence analysis

CCA is a descriptive method of presenting an economic evaluation, with the costs and consequences of an intervention reported without analysis (Gold and Siegel, 1996; Torrance, 1987). The costs and outcomes of alternative interventions can be compared, without the restrictions of a single outcome measure or focus on health benefit (Greco et al., 2016). It is a simple way of presenting the important information regarding the value of the treatment for use as a decision aid (Drummond et al., 2005), although capturing all the consequences is unlikely and there may be uncertainties in the data, particularly if costs and benefits occur in the future (Gage et al., 2006). Furthermore, the results of a CCA require subjective interpretation of all the evidence by the individuals reviewing the output as there is no decision rule on which to judge the alternative interventions (Drummond et al., 2005). CCA may be informative when interventions are funded outside of the NHS, implemented in non-health settings and where relevant outcomes may not be health. Indeed, the National Institute for Health and Care Excellence (NICE) endorse the use of CCA alongside CUA when evaluating public health interventions (NICE, 2014).

2.3.3 Cost-effectiveness analysis

The intention in CEA is to minimise the costs of achieving an improvement in a unit of health outcome, or to maximise a gain in a health outcome within a given budget (Breheny, 2013; Garber and Phelps, 1997). Health in this context is measured in a single natural unit, such as a unit of blood pressure or BMI. Using different outcome measures restricts the comparability of analyses and therefore CEA cannot be used directly for allocative efficiency decisions and can only inform decisions in healthcare (Claxton et al., 2006; Donaldson et al., 2002). When comparing similar health problems in similar populations, decisions regarding technical efficiency can be made using CEA however (e.g. comparing obesity interventions delivered to the same population and measuring health effects in BMI units) (Yeo et al., 2019), with natural units perhaps being the most appropriate outcome. This data may be more readily available from routine sources, compared to self-reported preference data which might be required to conduct CUA. Due to a lack of focus on utility, it is theoretically underpinned by extra-welfarism. The product of CEA is a ratio reflecting the cost of a unit of health benefit compared to the best alternative or current practice, namely the cost effectiveness ratio (CER) (Garber and Phelps, 1997; Morris et al., 2007;

Torrance, 1997). An incremental cost effectiveness ratio (ICER) is however cited as more appropriate as it reflects the additional benefit the alternative intervention provides (Drummond et al., 2005; Morris et al., 2007). Both CER and ICERs are expressed in natural units such as days of pain avoided (Mauskopf et al., 1998) and the comparator's measure of effectiveness must be expressed in the same units (Breheny, 2013; Garber et al., 1996). In contrast to CBA, CEA does not attach a monetary value to the outcomes of an intervention. In order for the resulting ratio to be interpretable or useful, there must be a decision rule in place. This rule will mostly reflect what is thought of as the optimal use of resources as the health effect will be maximised and the benefits obtained at the lowest cost (Garber et al., 1996). This is a predominantly subjective monetary threshold recommended to be the best use of resources.

2.3.4 Cost-utility analysis

CUA is a form of CEA, however the outcomes are not expressed in the natural units prescribed by CEA and are instead a function of the quantity and quality of life (QOL) to reflect healthy years of life (Drummond et al., 2005). This allows the comparison of interventions with different clinical outcomes and the comparison of interventions for different indications. Costing methods in CUA are comparable to CEA however. The aim of CUA is to maximise quality adjusted life years (QALYs) gained within a limited budget (Johannesson and O'Conor, 1997), with an efficient intervention generating a positive number of QALYs for the smallest cost per QALY gained (Breheny, 2013; Harris, 1987). Like CEA, CUA is therefore underpinned by extra-welfarism. In CUA, the outcome is most commonly QALYs, but could be disability adjusted life years (DALYs) for example (Drummond et al., 2005). A QALY is the time in a state of health adjusted for the QOL in that state, represented by a health state value (Drummond et al., 2005). These values can be elicited from the general population, allowing their preferences to be incorporated into the analyses. For the appraisal of new technologies, CUA is the methodology currently preferred by NICE (NICE, 2014). However, as CUA is focused on health benefits and does not capture the wider welfare impact of interventions, NICE advocates the use of CCA and CBA in addition to CUA when evaluating public health interventions.

2.3.5 Cost-benefit analysis

In CBA both the inputs (costs associated with treatment) and outputs (benefits of treatment) are valued in monetary units. Neither are limited to those incurred in healthcare, allowing multi-sectoral analyses capturing the costs and benefits falling within broader sectors, such as the criminal justice system for example (McIntosh and Luengo-Fernandez, 2006). This is pertinent for public health interventions, which typically are intersectoral (Weatherly et al., 2009). There is an assumption that the decision maker can allocate resources across sectors (Claxton et al., 2007), and at the present time in the UK local authorities do have responsibility for several budgets. If the benefits outweigh the costs it could be assumed that there is a welfare benefit to society (Breheny, 2013; Robinson, 1993a) and a worthwhile use of resources (Currie et al., 1999; McIntosh and Luengo-Fernandez, 2006). This is an intuitive outcome, which could be interpreted by decision makers without expertise in health economics. As the result is expressed as a monetary value, the resource use can be compared to the potential benefits that could be obtained elsewhere in society (Robinson, 1993a), enabling decisions of allocative efficiency to be made (Yeo et al., 2019). This also eliminates the need for a willingness to pay threshold to interpret the results, as required in CUA and CEA (Svensson and Hultkrantz, 2017). Obtaining the benefits of treatment as monetary values has been approached using contingent valuation studies or discrete choice experiments (DCEs) (Ryan et al., 2003). Contingent valuation encompasses both willingness to pay (WTP) and willingness to accept (WTA) approaches. In essence, DCEs and contingent valuation studies attempt to elicit what an individual would pay to experience a health benefit or avoid an illness (Breheny, 2013; Robinson, 1993a). Alternatively the human capital approach can be used, where monetary weights are placed on healthy time and the benefits are assessed as the value of future earnings (Drummond et al., 2005). The theoretical underpinning of CBA is welfare economics. The compensation aspect of welfarism (Cohen, 1995) is addressed in that individuals should consider how much they would accept to be compensated for a loss of health (McIntosh and Ryan, 2002), allowing the intervention to be compared to other financial expenditures means that overall societal benefit can be evaluated (Breheny, 2013; Robinson, 1993a). While these approaches have facilitated the operationalisation of placing a monetary value on health, critics argue that this is conceptually problematic (Svensson and Hultkrantz, 2017). There are however advantages, such as allowing the inclusion of intangible costs such as process utility (Drummond et al., 2005).

2.3.6 Return on Investment

Return on investment (ROI) traditionally is a method used to assess the economic benefits of an investment relative to its costs from the perspective of an investor (Brousselle et al., 2016). In contrast to CBA where the cost-benefit ratio is the intervention costs minus the benefits, ROI is the financial benefits minus the intervention costs and divided by the intervention costs. The resulting value can be presented as a percentage reflecting the return on the investment. This can assist in justifying public health interventions being investments instead of expenses and encouraging discussions of affordability instead of worth

(Brousselle et al., 2016). ROI can be used to demonstrate the economic value of investing in public health interventions, however its use is hindered by the same difficulty experienced when using CBA, namely putting a price on health (Brousselle et al., 2016). Monetary values can be assigned to medical costs averted and productivity increased, although costing intangibles and externalities is difficult. It is also not possible to incorporate equity considerations into the analyses and the results of ROI are vulnerable to ranking or comparison which may be inappropriate in public health (Brousselle et al., 2016). Prioritising contrasting services such as smoking prevention over childhood obesity services due to a higher ROI involves a judgement of their relative value for example (Brousselle et al., 2016). NICE (2011) refer to ROI as a general term encompassing techniques used to compare the costs and benefits generated by an investment, to include CBA and multi-criteria decision analysis (MCDA).

Social return on investment (SROI) is a framework for measuring the value of a programme beyond financial accounting (UK Cabinet Office, 2012). The addition or reduction of social and environmental value of interventions is captured within this framework, allowing a broader assessment of value than is allowed by approaches such as CBA (UK Cabinet Office, 2012). Like CBA, both the costs and benefits in SROI are valued in monetary terms, although the outcome is social value instead of financial value (Leck et al., 2016). The main difference between the SROI and CBA approaches is the involvement of stakeholders. Stakeholders are individuals or organisations that will experience positive or negative change as a result of the initiative being evaluated (Leck et al., 2016). Mixed methods are employed to determine the outcomes and their indicators, although quantifying social value is limited by data availability and subject to similar difficulties as valuing health (Arvidson et al., 2010). Due to the subjective and iterative development of a SROI evaluation, the analyses must be accompanied by a narrative that reports the methods used and judgements made (Leck et al., 2016). It still remains that as stakeholders can determine outcome selection, the results cannot always be compared to other activities and between sectors (Arvidson et al., 2010). The identification of appropriate stakeholders when effects might occur in the future and unexpected outcomes might occur also challenges their inclusion (Leck et al., 2016). Additional methodological steps in SROI beyond CBA include consulting stakeholders on the proposed outcomes of the SROI analysis and the requirement for external assessment of the analysis for assurance. The Her Majesty's Treasury's Green Book notes that SROI is a practical tool for use in small and large organisations valuing outcomes for different stakeholders. For CBA, the Green Book instead values costs and benefits for the whole of society (HM Treasury, 2003; UK Cabinet Office, 2012).

A review explored the financial benefits of public health interventions in high income countries (Masters et al., 2017). Fifty-two published studies examining ROI or cost-benefit ratio of public health interventions were identified. Most interventions were cost saving, with positive ROI or cost-benefit ratios. National as opposed to local interventions yielded higher returns on investment, attributed to their one-off nature such as a vaccination or legislation. Health promotion and healthcare public health delivered at a local level had lower returns, generally requiring sustained implementation, of greater complexity and were resource intensive. Another review examined the use of SROI in public health (Banke-Thomas et al., 2015), finding 40 SROI analysis. The majority were conducted in health promotion (N=12) or mental health (N=11). The SROI ratios were all positive, suggesting they provide value for money.

Since funding of public health in England was transferred to local government, (NICE, 2014) has applied a flexible approach to the type of economic evaluation used to evaluate public health interventions. This is due to their wider responsibility for welfare as well as health and the smaller population served, with more identifiable needs. Whilst CUA and CEA remain the preferred methods where health is the primary benefit and to enable the comparison between programmes, CCA and CBA will also be considered. CCA allows commissioners to evaluate additional outcomes important to them including equity. CBA allows the comparison with non-health interventions competing for investment as the benefit is in monetary terms.

2.4 <u>Vehicles for conducting economic evaluations</u>

When making a resource allocation decision, decision makers need to understand whether an intervention is cost-effective compared to current care (Sculpher et al., 2006). Incorporating an economic evaluation into a clinical trial can provide decision makers with estimates of the cost-effectiveness of interventions compared to standard practice (Petrou and Gray, 2011). Advantages include the potential to explore associations between treatment outcomes; patients' health related quality of life (HRQL), complications and associated resource use. A pragmatic trial for example can reflect routine care, with the intervention received and follow-up similar to that outside of a trial situation (Petrou and Gray, 2011), thus providing real world evidence of treatment cost and effectiveness. The duration of clinical trial limits the extent to which the long-term cost effectiveness of an intervention can be evaluated. This is a situation where decision analytic modelling is beneficial. Decision analytic modelling synthesises existing evidence to estimate the cost effectiveness of an intervention (Petrou and Gray, 2011). Trial based and model based economic evaluation methodology will now be described briefly, and will be drawn upon at the end of the chapter, where the challenges of economic evaluations of public health interventions are summarised.

2.4.1 Trial based economic evaluations

As mentioned, clinical trials can be used to provide estimates of an intervention and its comparator's costs and effects, thus informing a full economic evaluation. The procedure will not be explored in depth, but briefly, patient-level resource use data (e.g. GP attendances) and effectiveness outcomes are collected alongside the trial (Glick et al., 2014). These are valued, with costs assigned to the resources used. This data can then be used to examine the incremental difference in costs and outcomes of the interventions being compared. While this approach allows an assessment of the direct impact of the intervention on costs and outcomes, there are several challenges that limit the potential for generalisable conclusions to be drawn from the results. These include the choice of comparator treatments, collection of suitable outcome data and the extent to which the trial reflects usual care (Sculpher et al., 2006).

Whether an RCT is always the most appropriate study design for the assessment of the causal impact of a public health intervention on health is a debated issue (Rychetnik et al., 2002). Systematic reviews of RCTs are typically considered the gold standard evidence of effectiveness (Evans, 2003). Although there is a belief that the RCT is only suitable for the evaluation of simple, unadaptable and standardised interventions targeted at individuals (Rychetnik et al., 2002), which conflicts with most public health interventions. Cluster RCT designs are able to address the issue of non-individual delivery and can be a good compromise when interventions are delivered in clearly-defined settings such as schools. Cluster designs may still be impractical in other community settings, where contamination is still highly likely. The ability to conclude that there is unequivocal evidence of effect, or no effect is also challenging. No effect could be due to quality of implementation, suitability of the outcomes or adaptation of the intervention for example. Process evaluations can aid the interpretation of results, although these are resource intensive and the ability to delineate the potential cause and effect mechanisms is limited.

Process evaluations can also inform assessments of the transferability of the RCT findings. To be useful in decision making, it must be possible to judge the transferability of the findings to settings outside of the one where it was evaluated (Rychetnik et al., 2002). In the case of school-based public health interventions, this could be another geographical area, different type of school governance or different age group for example.

RCTs do not usually have a follow-up period long enough to detect the desired outcome (Frieden, 2017) instead relying on surrogate outcomes. In a school setting, children will leave, usually progressing to a different establishment or finishing education entirely. This hinders longitudinal data collection and if it is attempted, perhaps ultimately results in a selective sample. Using linked routine data is a solution, once the appropriate permissions and administrative processes are completed.

2.4.2 Decision analytic modelling

As mentioned, health economic modelling can be used when data from a single clinical trial is insufficient to inform decision making. Situations when modelling can be used include extending the results of a trial when the time horizon was too short to reflect long-term cost and outcome implications, linking intermediate endpoints to final outcomes, applying results from one context to another, conducting treatment comparisons where direct comparisons have not been made in trials and answering questions that cannot be informed by primary data (Barton et al., 2004; Brennan and Akehurst, 2000).

Patient healthcare pathways, reflecting the progression of disease are mapped and in the case of more complex interventions, the implications falling outside of the health sector also considered (Brennan and Akehurst, 2000). The impact of treatments on costs and consequences are estimated using evidence from existing literature which is used to populate the model. Where there are several sources of important evidence, the data can be synthesised using techniques such as meta-analysis. There are several types of model structure that can be used in economic evaluation. Decision trees are the simplest, whereby patients follow a clear pathway with probabilities determining which ‘branch’ of the tree is followed. Costs and outcomes for each treatment can be estimated from the relevant pathways. Decision trees are not suited to all clinical problems however and are best used when there is a one-off treatment decision (Barton et al., 2004). Markov models are a second type, where individuals transition between different states of health in cycles. Markov models can incorporate recurrence and are able to represent pathways experienced in chronic conditions, in contrast to the one-off treatment choice represented in decision trees (Barton et al., 2004). Individual simulation models are the final type. These allow for the interaction between individuals, which is valuable when modelling infectious diseases for example. These models are computationally complex however (Brennan and Akehurst, 2000).

The challenges of conducting economic evaluations of public health interventions will be discussed at the end of this chapter. The use of modelling is particularly relevant to public health interventions as the health impacts can manifest long after a trial ends and there can be broad implications across sectors and individuals which may be impossible to capture (Squires et al., 2016). While modelling can offer some solutions to the challenges discussed, it is also unable to address others. For example, appropriate estimates of costs and effects are still required to populate a model.

2.4.3 Economic evaluations alongside natural experiments

Another potential solution is conducting economic evaluations alongside natural experiments, for which guidance has recently been published (Deidda et al., 2019). Natural experiments are studies conducted when manipulating exposure to an intervention for the purposes of research is not possible (Craig et al., 2008). This can be in situations where its unethical or impossible to manipulate exposure or policies are introduced quickly. There are different acceptable approaches, including pre and post implementation measures or comparisons of those exposed and unexposed. The challenges and pitfalls of conducting economic evaluations alongside RCTs are well recognised, as are the issues with assessing the cost-effectiveness of public health interventions. Many of these also apply to natural experiments, although they provide the opportunity to evaluate implementation in the real world. Collecting data to undertake a comprehensive CUA, incorporating all inter-sectoral costs and outcomes, and capturing spillover effects is probably unfeasible (Deidda et al., 2019). A degree of pragmatism is therefore required. Observational data are largely used to estimate effectiveness, meaning utility data is probably scarce. Healthcare resource use could be captured from routine sources, although the inter-sectoral costs associated with public health interventions are unlikely to be so readily available. Schools could provide routine data however, such as academic performance in Standard Attainment Tests (SATs) and NCMP data.

2.5 Measuring Effectiveness

Conducting an economic evaluation requires evidence demonstrating the interventions' effectiveness. Evidence can be obtained from clinical trials, systematic reviews or the synthesis of several sources of data (Drummond and Jefferson, 1996). RCTs provide the most reliable evidence of clinical effectiveness, however they may not report the outcomes of interest or compare the relevant interventions (Drummond and Jefferson, 1996). The comparator treatment may not be current treatment practice, or a placebo control may be used where there is a treatment alternative available. Furthermore, it is unlikely that the trial duration would be long enough to capture all the health and cost impacts of an intervention. Due to these limitations it is common that data from several sources is synthesised and an economic model developed to extrapolate the available data to align with the time horizon of the analysis (Drummond et al., 2005). The time horizon of an economic evaluation should be long enough to capture all important differences in costs and effects of an intervention (NICE, 2014). In most cases this is likely to be a lifetime however a shorter time horizon might be appropriate in some scenarios (Drummond and Jefferson, 1996). Such situations would usually be when there is no difference in long-term QOL and mortality (NICE, 2014).

2.5.1 Clinical endpoints

Health improvement as a result of an intervention is an outcome used in CEA and is usually measured in natural units relevant to the condition or intervention (Torrance, 1986). These could be cases detected, cases averted, lives saved, or reductions in mortality and morbidity for example (Torrance, 1986). A physical activity intervention may present an ICER as cost per Body Mass Index (BMI) unit change or change in daily activity level (Cradock et al., 2017) and a polypill intervention for cardiovascular disease prevention may use cardiovascular event averted or life years gained (Becerra et al., 2015). These are single dimensions of health however, not reflecting the multi-dimensional nature of disease (Gafni and Birch, 1995) and restricting evaluations to interventions with directly comparable outcomes (Birch and Gafni, 1992). This has implications for informing decision makers aiming to achieve allocative efficiency, with CEA only informing technical efficiency (Donaldson, 1990). Survival as an outcome may also be inappropriate as some interventions treat only morbidities and not extend life or improvements in survival may be at the expense of a deterioration in quality of life (Torgerson and Raftery, 1999b). As already mentioned, mortality would likely have to be extrapolated using economic modelling also.

2.5.2 Health-related quality of life

HRQL has been described as an individual's emotional and physical functioning which contribute to their overall quality of life (Torrance, 1987). Health state utilities are values assigned to a state of health to reflect the health-related quality of life in that state, with 1.0 representing full health and 0.0 representing death (Torrance, 1986). These values are obtained by developing descriptions of health states and asking patients or the public to judge them, or asking patients experiencing health states to value their own health. For the purpose of economic evaluation, previously used values in existing literature could also be applied. The most commonly used outcome in CUA is the QALY and is the measure preferred by NICE (NICE, 2014). This measures outcomes in terms of both morbidity and mortality (Torrance, 1986), achieved by adjusting the length of life by the health state utility reflecting time in that state.

The EQ-5D (EuroQol Group, 1990) is a generic measure of HRQL that can be used to generate utilities. It has five dimensions that assess the ability to perform activities affected by physical functioning, pain and discomfort and an individual's emotional functioning in terms of anxiety and depression. The measure is not disease specific, which enables the comparison of interventions with differing health effects. A valuation exercise using the general population was conducted to develop preference weights for each health state generated by the measure (Dolan et al., 1995), thus health state profiles can be weighted to represent how society values the state. This is a desirable characteristic when evaluating interventions in a publicly funded healthcare system. The SF-36 and short form SF-12 are alternatives to the EQ-5D, comprising 36 and 12 questions respectively assessing eight domains contributing to HRQL (Ware and Sherbourne, 1992). To enable a utility value to be calculated, SF-36 data is converted into a SF-6D preference based health profile (Brazier and Roberts, 2004).

A criticism of the definition of HRQL is that it excludes other factors affecting QOL, such as social functioning, economic, cultural and environmental considerations (Torrance, 1987). Alternative measures have been developed, evaluating HRQL more broadly, beyond health. The ICECAP-O (Grewal et al., 2006) for example is a measure of functioning for older people. Dimensions relate to their capabilities, such as attachment, enjoyment and control. This has the potential to be used in economic evaluations as values have been obtained in a valuation study. Furthermore, by not focusing on health specifically the measure could also be used to inform evaluations of social care interventions (Coast et al., 2008c).

2.5.3 Monetary terms

Initial methods to value health benefits or the cost of illness for the purpose of CBA used economic productivity as a proxy (Rice, 1967). This was measured by estimating the cost of premature death and a reduced capacity to function as a result of disability (Dunlop, 1975; Milton, 1981; Rice, 1967). This is the human capital approach, which is complicated by the estimation of work productivity for the unemployed such as the elderly, children or those unable to work. Gender and age imbalances in employment may bias estimates as well as the choice of earnings data (Donaldson, 1990). By focusing on economic worth to society, this approach does not address other impacts of morbidity that may be valued by individuals such as the burden of pain and suffering or loss of leisure opportunities (Rice et al., 1985). Individuals' aversion to risk of illness or death is another aspect that is not captured, in contrast to the QALY approach (Landefeld and Seskin, 1982).

An alternative method of valuing the cost of a human life is the WTP approach. WTP values health effects in terms of what individuals are willing to pay or willing to accept in compensation for a change that will affect their health or probability of death (Mishan and Quah, 1971). The majority of contingent valuation studies in health use the WTP approach (Olsen and Smith, 2001). Theoretically, in contrast to the human capital approach, WTP captures intangibles contributing to individuals' wellbeing such as pain, suffering and leisure time (Landefeld and Seskin, 1982; Rice et al., 1985). In addition, it may implicitly include the rate of time preference as the weight given to benefits of living in the future (Landefeld and Seskin, 1982). Compared to QALYs which include only health gains, there is no restriction on what individuals express their preferences for, which could include the benefits of the process of receiving healthcare or the impact of receiving care on others for example (process utility) (Olsen and Smith, 2001). It is affected by ability to pay, with the well-off willing to pay more than those less well off (Landefeld and Seskin, 1982; Rice et al., 1985). Methodological challenges include a lack of comprehension of hypothetical risk situations (Landefeld and Seskin, 1982) or familiarity with valuing health as a commodity (Donaldson, 1990). Whether the question is posed as a WTP or WTA also produces differing estimates (Horowitz and McConnell, 2002), as do open ended or bidding approaches to eliciting values. (Frew et al., 2004).

2.6 Measuring costs

Opportunity costs are the benefits forgone or what could have been achieved by the next best use of resources. In the health system, this would be the health benefits that could have been achieved if the next best alternative treatment or intervention had been funded (Palmer and Raftery, 1999). When the most appropriate comparators are included in an economic evaluation, the opportunity costs can be observed in the estimated results. When estimating the costs of an intervention, typically the costs of the individual components are obtained to reflect the opportunity cost. Whilst this may be achievable for resources for which there is a market value such as a medical device for example, other costs are more challenging to estimate. Time spent receiving treatment or unpaid care is an opportunity cost, yet it has no market value (Palmer and Raftery, 1999). Valuing the opportunity cost of time is a subject of debate (Palmer and Raftery, 1999) however a technique used is the wage a working adult could expect to receive, adjusted for whether it is work or leisure time displaced (Robinson, 1993b).

The costs included in an economic evaluation reflect the value of the resources used (Morris et al., 2007). The process of costing firstly involves identifying the relevant resource implications of the interventions, whether an increase or decrease. Relevant resource use would be that which would change as a result of the intervention (Pyke, 1998), but would also depend on the perspective adopted and the time horizon of the economic evaluation. The quantities of these resources are subsequently ascertained and then unit costs attached (Morris et al., 2007). There are two main approaches to costing. A micro or bottom-up approach quantifies each individual element of an intervention separately and assigns a cost to each. This is most accurate, but also the most resource intensive approach and less generalisable to other contexts. Top down costing can be conducted when a micro method is unfeasible. The total cost of a service or resource is measured at an aggregate level and may be divided by the number of times it was used. This results in a less precise although potentially more representative estimate as the micro-costed value may only reflect the specific setting costed (Byford et al., 2003; Simoens, 2009).

The cost of interventions can be attributable to several providers including the NHS, public health and social care systems. Costs can also be incurred outside of the health sector by patients and carers however (Neumann et al., 2016). Costs included in an economic evaluation are those resulting from providing an intervention, any adverse events, services used by the patient and any related future costs (Mogyorosy and Smith, 2005). This would include materials, staff and expenses. Other costs may be incurred by other sectors or individuals, for example social care costs and informal care (Mogyorosy and Smith, 2005). Patients' out of pocket costs, lost productivity due to treatment, inability to work or premature death may also be included in an economic evaluation depending on the perspective or decision maker (Byford et al., 2003).

2.6.1 Long run and short run costs

Related to the time horizon, costs can be incurred both in the short or long-term. Short term or short run costs are usually variable costs initially. Fixed costs are those that remain the same in the short term, despite changes in volume (Byford et al., 2003). This could be the cost of employing a school nurse, which would be incurred independently of how much activity they undertake (Mogyorosy and Smith, 2005). In contrast, variable costs change with volume of activity. Direct costs such as medications used would be dependent on the number of individuals receiving in an intervention for example (Mogyorosy and Smith, 2005). What is actually defined as short run depends on the intervention and decision problem, although they are usually less than 3-5 years. This is because over the longer term changes in factors such as technology or the environment can introduce uncertainty (Simoens, 2009). The long-term is therefore the period where production is variable (Byford et al., 2003).

2.6.2 Perspective

The perspective is the viewpoint from which an economic evaluation is conducted. In terms of costs, a healthcare or payer perspective considers only the direct medical costs and benefits of an intervention currently and in the future, whereas a societal perspective considers all costs and benefits regardless of who incurs them (Gold and Siegel, 1996). A societal perspective may be preferable when important effects of an intervention are observed outside of the health system (Sanders et al., 2016), for example when patient and/or their carer's time or productivity are impacted or the intervention has significant implications for the education or legal systems. In terms of outcomes, the societal and NHS personal and social services perspective (PSS) perspectives are similar in that they both advocate including anyone who experiences health effects related to an intervention, which would include carers (Gold and Siegel, 1996; NICE, 2014). In the UK, NICE recommends that an NHS and PSS are adopted as its remit is to guide the efficient use of these resources (NICE, 2014). In some cases when benefits may be experienced in other sectors both the healthcare and societal perspectives may be reported or in other settings such as the USA, it is recommended that results of both are presented (Sanders et al., 2016). In England, local government is responsible for implementing public health interventions. The NICE Reference Case (NICE, 2014) therefore recommends a public sector perspective, taking the view of the government department which is administering the intervention. The guideline acknowledges that the department administering the intervention will not necessarily observe the benefits, although offers no solutions.

2.6.3 Discounting

Positive time preference describes the phenomena that individuals are not indifferent to costs and benefits over time, in fact people value things more in the present than in the future (Morris et al., 2007). Explanations include the idea that individuals will get more pleasure from an event now than in the future or that due to the possibility of death, we take into account the possibility that we may be unable to experience something in the future (Morris et al., 2007). Thus in the context of health economic evaluation, our perceived value of costs and health benefits decreases over time (Torgerson and Raftery, 1999a). Discounting is the analytical procedure in which the future costs and outcomes associated with an intervention are adjusted to account for time preferences and reflect their present value (Torgerson and Raftery, 1999a). The NICE reference case states that the results of economic evaluations should reflect the present value over the time horizon of the analysis, therefore both costs and benefits should be discounted at a rate set by the UK Treasury, currently 3.5% (NICE, 2014).

Whether costs and benefits should be discounted uniformly or with a different rate is a matter of debate (Claxton et al., 2011). Whether individuals have the same time preference for receiving health outcomes now as they do monetary benefits is an unresolved argument (Brouwer et al., 2005). The rationale for equal discounting is that the impact of time is independent of whether the commodity is a cost or an outcome and that there is a need for consistency. If you were to discount outcomes at a lower rate than costs, interventions appear to be more cost effective if they are delayed (Claxton et al., 2011). If higher discount rates are used for outcomes, thus attaching a lower weight to future health then cost-effectiveness ratios will worsen (Rappange et al., 2010). Complexities arise when comparing interventions that require immediate investment but do not yield health benefits until the future (Severens and Milne, 2004). An example is immunisation programmes (Walker et al., 2010). Enduring consequences of immunisation such as avoidance of death at an early age or permanent disability will have little weight. Also, immunisations that can have benefits in the distant future such as human papilloma virus vaccination (HPV) would be weighted less (Claxton et al., 2011; Severens and Milne, 2004). Although these issues are applicable to economic evaluation methods in general, they are particularly pertinent to preventative public health interventions (Rappange et al., 2010; Weatherly et al., 2009).

The NICE reference case for technology appraisals (NICE, 2014) recognises that interventions with potentially long-term health benefits (at least 30 years) may be disadvantaged by a discount rate of 3.5%. To mitigate this, it is suggested that a lower rate may be considered if there is evidence supporting long-term health benefits. Furthermore, for public health interventions, NICE proposes that a sensitivity analysis using a discount rate of 1.5% is presented alongside the reference case analysis (NICE, 2014).

2.6.4 Implementation Costs

Whilst methodological literature focuses on the inclusion of costs in relation to perspective and debates surrounding unrelated and future costs, little attention is paid to costs incurred across the full lifespan of interventions (Drummond et al., 2005; Neumann et al., 2016; Weatherly et al., 2009). Recommendations suggest that all resource use that is relevant to the analysis should be included (Neumann et al., 2016) and the length of time should be dependent on avoiding the misleading of decision makers (Drummond et al., 2005). NICE public health guidance recommends that all important costs are included and the time horizon chosen should reflect this (NICE, 2014). This issue is particularly relevant to behavioural interventions where ongoing delivery or monitoring may be required in addition to the initial set-up investment. In relation to such programmes, Ritzwoller et al. (2009) proposed that all intervention components required for intervention replication should be costed. The potential points at which costs could be incurred in developing, delivering and maintaining a behavioural intervention will now be explored using examples from the literature.

Difficulties evaluating behavioural interventions and hindering the accurate costing, demonstration of cost-effectiveness and eventual uptake of behavioural interventions have been highlighted by Ritzwoller et al. (2009). Firstly, the identification of all long-term costs and benefits is challenging. This aligns with difficulties observed evaluating public health interventions. Furthermore, few analyses reportedly explore the uncertainty surrounding implementation, such as delivery conditions and workforce requirements. Another issue is the focus on modelling long-term cost-effectiveness at the expense of establishing an accurate cost of implementing an intervention, which is more critical information for decision makers. A lack of standards for capturing costs of behavioural interventions is cited as a barrier and the reliance on retrospective data collection can make estimates unreliable. Finally, accounting records do not reflect the inevitable variation in how programmes are delivered.

2.6.5 Set-up costs

The WHO classifies start-up costs broadly as all costs that are incurred before an intervention is implemented or in a pre-implementation phase (WHO, 2007). Similarly, Johns et al. (2003) define the start-up costs as those incurred in the period between the decision to implement an intervention and its first application. In regard to the implementation of vaccines, start-up costs refer to the initial investments required to introduce a vaccine (Levin et al., 2014). This can include activities such as; training, guidelines and planning. These are distinct from recurrent or operational costs which are the costs of running the programme such as transport, monitoring and supervision. In behavioural intervention literature, development costs can be included as an initial cost (Ritzwoller et al., 2009). These include the resources to develop protocols, websites and materials to aid delivery. The cost of adoption is another recurring term used to describe costs incurred at an early stage (Lang and Connell, 2017). Adoption costs, or pre-implementation costs are those incurred when preparing an intervention (Lang and Connell, 2017).

2.6.6 Intervention costs

The intervention costs are those incurred when individuals begin to receive an intervention or it is put into practice. The WHO refer to these as post-implementation costs (WHO, 2007) and they could also be thought of as operational or recurrent costs (Levin et al., 2014). In a disease management programme, implementation or intervention costs were those that occurred after programme development and when it began to be delivered to patients (Tsiachristas et al., 2014). This included management costs, collection of quality indicators, the cost of materials and the costs of keeping the information technology operating. Ritzwoller et al. (2009) include; provision of materials, labour required to deliver the intervention, and the collection of data from individuals in order to tailor behavioural interventions to the patient. As an example, Saldana et al. (2014) included training foster parents.

2.6.7 Ongoing costs

Ongoing costs are those incurred for interventions involving repeated administration or oversight, not one-off delivery. This may be required to ensure a programme is sustained (Saldana et al., 2014) or an effect is maintained. Determining the effectiveness of such interventions involves a long-term assessment across the programme, therefore the concurrent cost implications should also be considered (Saldana et al., 2014). Possible costs could be repeated visits for monitoring, training or multiple sessions of an exercise intervention for example. For a diabetes self-management programme, these consisted of costs of the nurse support, patient retention and costs of managing the programme (Handley et al., 2008).

2.7 Allowing for uncertainty

Confidence in the results of an economic evaluation and consequently its potential use for decision making can depend on how uncertainty was considered. There are three broad areas of uncertainty: uncertainty about data inputs, uncertainty related to the extrapolation of results and uncertainty relating to analytical methods (Drummond and Jefferson, 1996). The potential uncertainties also vary depending on whether decision analysis or an individual patient data approach have been used (Jain et al., 2011).

In a decision model, the most appropriate parameters may be unclear (Jain et al., 2011), for example in a situation where clinical trials report conflicting effectiveness data (Briggs et al., 2012). Where patient level data is available, deciding which unit costs or utility weights to use will also introduce parameter uncertainty (Jain et al., 2011). Variability in treatment pathways or disease trajectories may mean that one model structure cannot reflect all patient pathways. Uncertainty could be a result of random variability between patients, as individuals may not necessarily respond to an intervention in the same way (Briggs et al., 1994). Incorporating all possible patient heterogeneity into a model structure would be challenging. Extrapolating results beyond the primary data source may also introduce uncertainty, for example modelling smoking relapse rate beyond the end of a clinical trial. How costs and effects are valued, the use of discounting and including indirect costs are methodological choices for the analyst (Briggs et al., 1994). Given the breadth of possible methodological approaches available and the possibility of equally appropriate methods, these choices can cause uncertainty.

Assessing the robustness of results of an economic evaluation requires that the potential impact of uncertainty is examined (Briggs et al., 1994) and NICE recommends that decision uncertainty is explored (NICE, 2014). In this case, the analysis aims to assess the confidence in the findings but it can also be used to ascertain the value of collecting further data to support the analyses (Briggs et al., 2012). There are several forms of sensitivity analysis. The simplest is one-way sensitivity analysis where parameters are varied individually across plausible ranges (Briggs et al., 1994). This enables the impact of changing single components to be assessed. Multiple parameters can also be varied simultaneously, perhaps to reflect different clinical scenarios. Threshold analysis is an approach that allows an exploration to find the critical value at which the conclusions of a study will change (Briggs et al., 1994). This may be useful in determining a cost of an intervention which would make it comparable to an alternative. An analysis of extremes can also be conducted using the best and worst possible parameters. If the results remain cost effective using the most pessimistic model inputs there could be confidence in the base case findings for example. Finally, probabilistic sensitivity analysis (PSA) can be employed to explore uncertainty surrounding different model inputs. This is the approach preferred by NICE to address parameter uncertainty (NICE, 2014). Probability distributions are assigned to uncertain model inputs and varied simultaneously. Monte Carlo analysis records the results of many simulations using randomly selected values from the distributions. This provides a large number of estimates of the potential cost effectiveness of an intervention

2.8 **Public Health Issues in Economic Evaluation**

The main components and considerations when conducting an economic evaluation of a health intervention have been described. These methods have largely been developed to aid decision making regarding clinical interventions for individual patients, for example a hip replacement (Weatherly et al., 2009). Difficulties arise when applying these methods to public health interventions for several reasons.

Many public health interventions could be described as complex (Edwards and McIntosh, 2019). According to the MRC, features of complex interventions include having several interacting components between interventions and controls in experimental studies, having a number of behavioural influences when delivering and receiving the intervention, the intervention targeting multiple organisations or groups, the number and variety of outcomes, and the flexibility and tailoring of the intervention over time (Craig et al., 2008). These characteristics present challenges for evaluating their cost-effectiveness. Conducting process evaluations alongside clinical trials can assist in understanding the mechanisms of impact, reasons for success and failure, and intervention fidelity and adaptation (Moore et al., 2015). It could identify one expensive but seemingly ineffective component for example, or adaptations that increased costs. Process evaluations can contribute to interpreting the results of an economic evaluation, although issues remain with their conduct. Without a clearly defined and bounded intervention, it is difficult to estimate its cost and the potential cost implications for different sectors. Choosing one appropriate outcome is difficult as the impacts could be diverse, observed in a variety of populations and not manifest immediately. Similarly, accurately capturing the resource implications over a long period of time might be difficult, yet this is important to understand to reflect its long-term cost-effectiveness. Not all public health interventions are complex, yet they still pose many challenges for economic evaluation. These challenges are described in Table 4.

Table 4 Challenges conducting economic evaluations of public health interventions

Challenge	
Measuring and determining effectiveness	
Determining effectiveness (Weatherly et al., 2009)	In contrast to clinical interventions, there are relatively few RCTs of public health interventions. This is understandable in some cases, for example testing the effectiveness of a preventative intervention could require a large sample and long-term follow-up. Determining unbiased estimates of effectiveness is therefore difficult and other sources may be required.
Determining cause and effect (Kelly et al., 2005)	Public health interventions can be multidimensional approaches potentially involving different mechanisms of behaviour change and involving different agencies. Determining which aspects of an intervention has contributed to the success or failure of an initiative can therefore be difficult.
When to measure effectiveness (Kelly et al., 2005)	Public health interventions usually aim to impact long-term health outcomes which may take years to become evident. The point at which success is determined is therefore difficult to judge.
How to measure effectiveness (Kelly et al., 2005; Weatherly et al., 2009)	Determining the most appropriate measure of effectiveness for public health is challenging. The QALY measure may not be sufficient to capture the wide range of health and non-health outcomes and those perhaps not directly targeted by an intervention. If the QALY was to be used, whether the results are comparable to a clinical intervention is also uncertain given the complex nature and sometimes indeterminate population.
Biological and social variation (Kelly et al., 2005)	In a clinical medicine, treatments are usually tested and subsequently administered to a refined population. In public health, interventions are implemented in much broader and undefined populations. As individuals do not respond consistently to interventions, determining a definitive level of effectiveness is potentially unachievable.

Challenge	
Individual or population measures (Kelly et al., 2005)	Whether change should be measured at the individual or population level is a matter for debate. Certain interventions may be more effective in certain groups so measuring success at a population level may understate the success of public health initiatives.
Incorporating equity (Weatherly et al., 2009)	An important issue is the compatibility between health promotion interventions to reduce health inequalities (WHO, 1986) and economic evaluation methods, which tend to focus on efficiency (Weatherly et al., 2009). The assumption that the value of a QALY is the same no matter who receives it is not aligned with the public health goal of reducing inequalities.

Costs	
Discounting future costs and benefits (Brouwer et al., 2005; Weatherly et al., 2009)	As public health interventions are largely preventative, the benefits are not usually seen immediately. As discounting is based on the premise that health outcomes and costs are valued more in the present than in the future, preventative services are disadvantaged. Choosing a discount rate is therefore contentious.
Involvement of multiple sectors (Weatherly et al., 2009)	The costs and consequences of public health interventions may fall on several different sectors therefore identifying these to include in an economic evaluation is complex. A significant issue is that an intervention's funder may not necessarily observe its benefits and cost savings could be made in other sectors as a result. For example, improving housing conditions could reduce healthcare utilisation.
Interventions	
Development of interventions (Kelly et al., 2005)	Compared to medication development, which is strictly regulated, public health intervention development is relatively flexible. This means that interventions may change during implementation, making the interpretation of results challenging.
Upstream and downstream interventions (Kelly et al., 2005)	Interventions operating upstream target the circumstances that result in negative health behaviours such as smoking. Downstream interventions aim to change those behaviours. These are interlinked, in that an upstream change might be a prerequisite for a downstream intervention to be effective. Adequately capturing the impact of context must therefore be considered.
The role of behaviour change (Kelly et al., 2005)	Public health interventions frequently require a change in behaviour to achieve the intended outcome. This behaviour change is an intermediate event, but a necessary requirement to observe the desired outcome. Modelling an individual's or population's behaviour is not an approach typically used in economic evaluations of clinical interventions.

2.9 Economic evaluation of interventions for children

The methods used to evaluate the cost-effectiveness of interventions have been described and the challenges posed by public health programmes identified. Additional challenges associated with interventions aimed at children will now be discussed.

2.9.1 Child development

Firstly, there are many distinct phases of child development in which they have different healthcare needs and are susceptible to different biological and environmental characteristics (Ungar and Gerber, 2009). For example, the social determinants of health already discussed can shape and influence the development of a child's physical and mental health and ongoing development. Delivering interventions at critical time points to achieve optimal outcomes is important and would influence the effectiveness and thus cost-effectiveness of an intervention. Preventative interventions targeted at disadvantaged children are most cost-effective in pre-school years for example (Doyle et al., 2009; Heckman, 2006). Children will also respond differently to treatment at different ages and will generally require different doses of pharmacological medication due to age, weight and size variation. The ability of a child to communicate their health status and symptoms also changes significantly throughout their development (Bevans and Forrest, 2009), with reliable self-reports emerging around age 12 (Ungar and Gerber, 2009). This will also be discussed in relation to outcomes, but determining what treatment is required and the success of an intervention will sometimes require a proxy report. Furthermore, adults will greatly influence the access a child has to healthcare (Ungar and Gerber, 2009).

2.9.2 *Measuring Costs*

As interventions targeted at children may be delivered in several different settings such as the community, schools and general practice; identifying the relevant costs may be challenging. The perspective of the economic evaluation will also be important in ascertaining what costs should be included in the analysis (Ungar and Gerber, 2009). Whether out of pocket costs paid by the child's caregivers are included in an economic evaluation conducted from societal perspective is a consideration for example (Lamsal and Zwicker, 2017). Whether productivity costs should be included in an economic evaluation is already a matter of debate, which is extended further when the productivity costs incurred when adults do not attend work to care for a child are considered (Lamsal and Zwicker, 2017). Although children are rarely in employment, their potential productivity could be affected by impaired development, either through; biological, environmental, social or behavioural factors. Cognitively and physically healthy children are more likely to attend school and become well-educated. Healthy and well-educated adults have better employment prospects (Bärnighausen et al., 2011). If these costs are not considered, the value of interventions could be underestimated. Obtaining data that enables an accurate prediction of future productivity costs would be challenging however.

NICE recommends that the time horizon chosen should incorporate all important costs and effects (NICE, 2014). Estimating long-term effects, impact on life expectancy and the possible associated costs will introduce significant uncertainty due to the difficulty of measurement. Over a child's lifetime it is possible that new technologies or treatments will be developed that may incur significant costs, yet also improve health outcomes. As with estimating potential cost implications of disruptions to a child's development, a lack of data would result in many assumptions and thus considerable uncertainty. As children experience less morbidity than adults on average, the majority of interventions directed at this population are preventative rather than treatment targeted (Ungar and Gerber, 2009). The difficulties in conducting economic evaluations within the realm of public health arise in these cases too. In the case of treatment interventions, evidence regarding treatment efficacy and costs in a paediatric population may be unavailable due to the paucity of clinical trials conducted in this population (Joseph et al., 2015).

2.9.3 Measuring Outcomes

In general, the measurement of health consequences in economic evaluations in children is similar to adults, whereby CEA uses measures of health status or mortality and CUA uses utility measures. However, choosing an appropriate outcome and obtaining data to support the analyses is more challenging (Ungar and Gerber, 2009). These challenges will be described briefly. Firstly, children are not always able to communicate signs and symptoms of illness (Rebok et al., 2001). Pain scales for example may not be appropriate for infants and adult proxy reports could be inaccurate (Huguet et al., 2010). Measuring these outcomes reliably is therefore difficult. The availability of suitable questionnaire instruments to capture health related quality of life, functioning and health state data is also limited (Thorrington and Eames, 2015). When an appropriate measure with psychometric evidence supporting its performance in the target population is available, administration may also be confounded by children's short attention span, awareness of being assessed and a desire to please the assessors or caregivers (Ungar and Gerber, 2009). Using a parent to complete an instrument on the child's behalf as a proxy is possible, although a lack of agreement between responses has been demonstrated in some domains (Eiser and Morse, 2001).

Agreement between children and proxies tends to be poor for social and emotional domains, but good for physical functioning (Eiser and Morse, 2001). In order to calculate QALYs for use in CUA, preference based measures are required. As already discussed, standard gamble and time trade off methods are complex and thus would be unsuitable for use in children due to their lesser ability to understand risk and time. Measuring WTP for the purpose of CBA is also not attempted for similar reasons and children's financial inexperience (Ungar and Gerber, 2009).

There are several questionnaires developed to assess children's health related quality of life. There is a youth version of the EQ-5D, the EQ-5D-Y (Wille et al., 2010), although there are no child tariffs (Feng et al., 2019; Kreimeier et al., 2018). Instead the adult value set is applied to derive utilities. Alternatives include the CHU9D (Stevens, 2009), HUI-2 and HUI-3 (Horsman et al., 2003), all of which have been used in paediatric CUA (Thorrington and Eames, 2015).

While HRQL and clinical outcomes are considered important when making resource allocation decisions in healthcare, these may not be relevant for decisions in an education setting. As mentioned in Chapter 1 (Section 1.4.9), primary schools are judged on academic outcomes. As budgets are also limited in the education sector and there are different ways in which education can be delivered to maximise attainment, the principles of economic evaluation have been applied in this setting. This will now be discussed.

2.10 **The Economics of Education**

The origins of the economics of education lie in the development of the human capital approach. Simply, differences in income are partly due to variations in labour quality, which is the human capital that has been acquired by individuals.

Human capital includes qualities that cannot be separated from a person, unlike capital goods for example. It can include knowledge, skills, health and values (Becker, 1975), with some specifying that it must have productive value to the economy. Education and training are therefore important investments to increase an individuals' human capital. A detailed examination of the theoretical basis of the human capital approach will not be explored. This is beyond the scope of this thesis and is relevant only insofar that it stimulated the application of the principles of economics to education.

The economics of education has been defined as the study of how society chooses to use and distribute scarce resources to produce training, the development of knowledge, skills, mind and character over time (Cohn and Geske, 1990). It can be used to contribute to decisions on how education can be best produced, the distribution of education, how much should be spent on education and what should be funded (Cohn and Geske, 1990; Dearden et al., 2009). Whilst economists are unlikely to have the expertise to advise the development of teaching techniques, they can evaluate inputs and outputs informed by teachers to provide recommendations to improve education (Cohn and Geske, 1990).

2.10.1 Use of economics in education policy

2.10.1.1 Increasing resources

One strategy hypothesised to assist in raising attainment is the provision of additional resources. The Excellence in Cities (EiC) programme was launched in 1999. It was an outcome of the 1997 White Paper: Excellence in Schools (Department for Education and Employment, 1997), which recognised that economic resources were important for supporting best practice and success in education. The government pledged a greater share of national income would be spent on education, affirming its priorities of education, employment and investing in the future. The EiC aimed to increase standards in urban schools and included seven strands that included additional support for the most and least able, removing barriers to learning and improving information technology (IT) resources. Partnerships were formed between schools and local education authorities (LEAs) who developed partnership plans detailing how resources would be used. The Department for Education and Science (DfES) then released funds on agreement of the plan. On average, the EiC cost £120 per pupil per year and by 2005/2006 total expenditure on the programme was £386 million (Kendall, 2005). An evaluation of the programme (Kendall, 2005) showed an improvement in mathematics attainment at the end of Key Stage 3 in the most disadvantaged schools, but no impact on English or science. A simple CBA was conducted, equating a one level increase in attainment to two years of additional education. This had a wage return of eight percent. The programme was deemed to be potentially cost-effective if the benefits did endure into participant later lives and resulted in higher earnings.

2.10.1.2 Incentives

The Education Reform Act, 1988 (Education Reform Act, 1988) introduced market principles into the school system with the aim of raising educational standards. Instead of allocating schools according to locality, children attended schools chosen by their parents, as long as school capacity allowed. The intended outcome of parental choice was the increased accountability of schools and the introduction of competition. Funding was linked with pupil numbers, incentivising schools to maximise their popularity. LEAs were obliged to allocate 75% of school budgets based upon student enrolment, with the other 25% allocated based on objectively assessed need. Unpopular schools would be financially disadvantaged and would have to adapt to satisfy parent demand.

Popular schools would gain pupils and receive increased funding. Evidence for the impact of this reform has been mixed and critics argue that it has increased inequalities in education. The most popular, best resourced schools are over-subscribed resulting in the use of geographical proximity selection criteria. Inflated house prices in desirable catchment areas subsequently restrict access of the disadvantaged to these schools.

2.10.1.3 Autonomy

Increasing schools' autonomy in terms of funding and governance has been used as approach to improving school standards (Academies Act, 2010). The introduction of autonomy to the school system was intended to encourage more innovative policies to help raise standards and increase competition between schools. The initial act in 2002 (Education Act, 2002) was in response to concerns that some schools in particular local authorities were delivering an inadequate education. As a result of the reforms there are now broadly two types of publicly funded primary schools in England. Community schools (local authority maintained) were the previous status-quo, receiving funding from local authorities and following the national curriculum. In 2002 the UK government introduced the 'academy system', where underperforming existing schools were converted to academies, receiving funding directly from the government and becoming autonomous from the local authority (Education Act, 2002; Eyles and Machin, 2018). Amongst many differences from community schools, they can set their own term times, follow their own curriculum and implement their own employment and performance management policies.

Prior to the Academies Act 2010, all academies were managed by independent sponsors, with a self-appointed board of governors. Some academies are part of multi-academy trust, with a single sponsor responsible for the management of several schools. Since 2010, schools achieving good or outstanding OFSTED rating are able to become academies (converter academies) without the need for an independent sponsor (Eyles and Machin, 2018). Instead they are run by an academy trust. As a result, academies now dominate the English secondary school system and represent approximately 25% of primary schools (Andrews, 2018). Academies can also be set up without the need to convert an existing school and are referred to as free schools. Demonstrating the impact of introducing school autonomy on education standards is complicated by the change in eligibility, with high performing schools able to become academies. Research has however shown that pre-2010 academies raised secondary school outcomes (Andrews, 2018; Eyles and Machin, 2018), although in primary education being part of a group of high performing local authority schools or in a high performing academy chain, was more important than being an local authority or academy school in general (Andrews, 2018).

2.10.2 Economic Evaluations in Education

As education is usually government funded, in the USA in particular there have been attempts to use economic evaluation methods to compare the value of policies that have similar educational goals (Levin, 1988; Long et al., 2015). Despite its potential to demonstrate the most efficient use of public resources, uptake has been limited and there has been little interest from policymakers in thorough evaluations, reportedly due to a lack of familiarity with the methodology (Levin and Belfield, 2015). CBA and CEA have both been employed in this setting, although CUA is a method not yet used (Levin and Belfield, 2015). Depending on the intervention’s aim, effectiveness outcomes have included reading ability (Hollands et al., 2016) and high school completion (Bowden and Belfield, 2015). In this setting there may be multiple benefits, such as improved attendance for students and teachers. Regarding costs, the largest input of educational interventions is usually labour (Levin, 1988), although this will depend on role and qualifications. Further resources required include facilities, equipment and materials.

An example of how analyses informed by economics have been used to assess an education intervention is now provided. To establish a transparent and standardised method to comprehensively cost an education intervention, a cost analysis was conducted to understand the resource use associated with a social and emotional learning (SEL) intervention in elementary schools (Long et al., 2015). SEL aims to develop social and emotional competencies, facilitating responsible decisions and the management of behaviour. Such skills have been associated with higher academic achievement, positive attitudes and reduced aggression (Durlak et al., 2011). It was recognised that whilst material costs were readily available in accounting records, teacher time utilised implementing interventions was not. As these were proportionally the greatest cost component, they required greater scrutiny and accurate estimation. Costs were estimated retrospectively using data from a three-year RCT. Using an ingredients approach (Levin, 1988), resources and their quantities were identified using interviews, reports and the internet. In addition to estimating material and facility costs, a detailed investigation of the activities and procedures was conducted to understand the opportunity costs attributable to teachers for programme implementation. Specific activities included initial training, ongoing training, internal meetings to evaluate progress and delivery of the programme over three years. The results of this cost analysis showed that year 1 of the intervention was the most expensive ($683,106) and year 3 the least ($566,426) (Long et al., 2015). The costs of the intervention over this period were still substantial, with ongoing support and programme delivery resulting in investment not being limited to the initial set-up of the programme. Although effectiveness was not addressed in this investigation nor costs of a comparator, the detailed costs will reportedly enable a more accurate estimation of cost-effectiveness of school-based interventions.

It is acknowledged that economic evaluation methods used in healthcare are more advanced and that published analyses are of higher quality (Levin and Belfield, 2015). Unaddressed methodological areas in education include addressing uncertainty in effectiveness outcomes and the use of statistical methods such as bootstrapping. As this literature is in its infancy, lessons can be learned from methods used in healthcare. When evaluating health interventions delivered in schools, experiences from the developing body of education research could also be drawn upon to inform approaches, particularly to costing. However, the differences in aims must be noted alongside their associated opportunity costs. Educational interventions would likely target educational attainment, although perhaps not directly for example through behavioural improvement. Consequently, their implementation would displace an activity with a similar learning objective. Placing a health intervention in school would probably displace an educational activity, thus carrying an opportunity cost to a child's learning. When comparing the methods these distinctions should be kept in mind.

2.11 **Conclusion**

This chapter has described the methodological approaches used in economic evaluation to inform the allocation of limited resources. While there are prescriptive frameworks that can be used to guide analysis, such as using a CUA if the intervention is intended to be used in the NHS, approaches to economic evaluation outside of this context are less defined. Moreover, unlike the NHS the decision maker in a school setting is undefined. The difficulties that arise when attempting to apply the advocated methods of economic evaluation to public health interventions have been highlighted, in addition to challenges that arise when assessing interventions for children. These are well recognised and documented in the literature. Economic evaluations in a school setting provide an additional level of complexity. While attempts have been made to evaluate education initiatives to inform government policy, predominantly in the USA, methods have largely been transferred from a health setting, rather than being refined to suit the specific circumstances in education. Through the mapping of public health provision in schools and review of methods used in economic evaluation, notable differences between education and health are as follows:

1. In schools, maximizing health is not necessarily the primary aim
2. Despite delivering and potentially funding interventions, schools will not benefit from health effects observed in the future
3. In schools, the funding of interventions is less clear
4. In schools, the decision maker and decision rules are not well defined
5. In schools, the costs of delivering public health interventions are not well understood, including the opportunity cost of time

Having identified these differences and the challenges of conducting economic evaluations of public health interventions, Chapter 3 reports a systematic review of methods used in economic evaluations of school based public health interventions. Conducting a thorough review of the literature provides an understanding of approaches currently used in this setting and facilitates the refinement of this thesis' research aims, which are provided at the end of Chapter 3.

CHAPTER 3 SYSTEMATIC REVIEW OF ECONOMIC EVALUATIONS OF PRIMARY SCHOOL BASED PREVENTATIVE INTERVENTIONS TARGETING RISK FACTORS FOR OBESITY

3.1 Introduction

As demonstrated in Chapter 1, public health interventions can be delivered in many settings, by different providers and using a variety of mechanisms to produce an effect. Consequently, evaluating their relative cost-effectiveness poses complex challenges (Edwards et al., 2013; Weatherly et al., 2009), as discussed in Chapter 2. These interventions can result in changes in costs across diverse sectors, including education and criminal justice as well as health and care systems. Equally, benefits resulting from these interventions often go beyond those captured in more traditional health technology assessments, (improvements in general wellbeing, for example, rather than health). Evaluating the impact of interventions on inequalities provides a further challenge, and although recent recommendations have suggested conducting equity impact analyses (conducting analyses by subgroup) or using equity weights (Cookson et al., 2017), appropriate data may not be available, and methods are not yet standardised.

Schools have been cited as an important setting for delivering public health programmes to children (Story et al., 2009). As children spend a significant amount of their time in school, in some senses it provides an ideal environment for the delivery of public health interventions but such interventions may also impact on objectives relating to academic attainment. Positive impacts may result from improving concentration for example, but these interventions also have the potential to divert resources from these key academic goals (Littlecott et al., 2018). Evaluations of school-based interventions thus need to capture all of these benefits and costs if they are to have traction in decision making across different sectors. An evaluation that assesses only use of healthcare resources but neglects expenditures or broader opportunity costs incurred by schools, or that assesses only health-related quality of life in terms of outcomes, is unlikely to provide the information necessary for taking decisions in school-based public health. This is recognised by organisations such as NICE in the UK, which allows for different methods that are acceptable for alternative decision problems. The US Panel on Cost-Effectiveness in Health and Medicine (Sanders et al., 2016) recommends presenting an 'impact inventory', encompassing the costs and effects both inside and outside the health sector.

Despite the difficulties, the cost-effectiveness of many school-based health promotion interventions has been assessed and several reviews published. A 2015 review examined the cost-effectiveness of physical activity health promotion (Korber, 2015) in children and adolescents. There were no restrictions on the type of analysis or outcomes, and the review included multi-component interventions, e.g. those including nutrition and complex interventions. Eleven primary school studies were identified and most were of at least 'good' quality, assessed using the Paediatric Quality Appraisal Questionnaire (PQAQ) (Ungar and Santos, 2003). The majority took a societal perspective, although what sectors these encompassed was not reported. A recent synthesis of full economic evaluations of childhood obesity interventions identified 21 school-based preventative programmes (Zanganeh et al., 2019). These comprised 75% of evaluations included and 20 interventions were assessed as cost-effective. Given the rapid publication of literature reporting economic evaluations in this setting, increasingly reviews are limiting their focus to schools. One recent review examined economic evaluations of lifestyle interventions delivered in schools (Oosterhoff et al., 2018). Twenty-three studies were identified, and findings included comparators being poorly defined and most studies reporting cost-effectiveness not cost-utility analyses. Another examined model based economic evaluations of school-based physical activity interventions (Batorova and Sørensen, 2019). Eight models were reviewed, with conclusions including estimating long-term effectiveness is challenging and the transparency of modelling methods was poor. Given the abundance of challenges of evaluating public health interventions highlighted in Chapter 2 (Edwards et al., 2013; Weatherly et al., 2009) and the distinctive features of schools as a setting, there is a need to identify potential solutions in the literature.

The objective of this review was to summarise the literature reporting economic evaluations of primary school-based interventions for obesity prevention or to reduce risk factors for obesity. Specific behaviours targeted were physical inactivity and poor diet. The review focused on these interventions for several reasons. Firstly, they were relevant to the case study trial in which the empirical work reported in this thesis is based. The Birmingham Daily Mile trial is reported in detail in Chapter 4. Briefly, this is a cluster RCT examining the effectiveness and cost-effectiveness of the Daily Mile, a physical activity intervention designed to prevent childhood obesity in UK primary schools. Examining the existing reviews suggested that there would be sufficient publications to enable an understanding of the breadth of methods to be explored, whilst the topic was not so broad that the number of publications would be too large to be feasibly reviewed. This scope would also encompass both health promotion and preventative interventions, which may have divergent methodologies.

The review aimed to explore methods currently used in a school setting and establish how the challenges of evaluating public health interventions are addressed. This will help inform the economic evaluation of future school-based interventions. These challenges include equity considerations, implementation fidelity, the costs incurred in multiple sectors and the maintenance of effect.

The objectives of the review were as follows:

1. Summarise the methods used (e.g. costs, outcomes, models/trial based, discounting) in economic evaluations of primary school-based interventions for obesity prevention or to reduce risk factors for obesity.
2. Examine how the economic evaluations identified have addressed the published challenges associated with estimating the cost-effectiveness of public health interventions.

3.2 <u>**Methods**</u>

3.2.1 Study selection criteria

Full or partial economic evaluations comparing two or more interventions were included. This review did not aim to inform decision making, therefore both partial and full economic evaluations were reviewed. Only partial economic evaluations comparing costs (cost analyses) or clinical trials reporting costs were eligible. The interventions targeted children aged between 4 and 11 attending a primary school (or equivalent). The analyses compared interventions delivered in a primary school setting aimed at preventing either of the following risk factors for obesity: physical inactivity, unhealthy diet. Obesity prevention interventions were also included. To be included, interventions must have included at least one component that took place during school time (for example excluding after school clubs) and must have been actively delivered to the children (thus excluding not teacher training and strategies to enhance implementation).

3.2.2 Searches and screening

Search strategies were developed and were informed by published filters to identify economic evaluations (CRD, 2015), strings used in previous Cochrane systematic reviews of interventions in school settings and reviews of diet, physical activity or obesity interventions (Dobbins et al., 2013; Wolfenden et al., 2017). Whilst only primary school (or equivalent) interventions were of interest, the search strategy was broad to capture any school-based intervention at the initial stage of searching. Search strings were adapted for each of the electronic databases searched, which included; Medline, Embase, PsycINFO, Healthcare Management Information Consortium (HMIC), CINAHL, The NHS Economic Evaluation Database (last updated March 2015), EconLit, Educational resources information centre (ERIC), British Education Index (BEI) and SPORTDiscus. Search strategies are provided in Appendix A. Only English language articles were eligible and no date restrictions were applied. Searches were conducted on February 8th 2018. Grey literature in the form of public health reports, NHS and NICE publications were also searched to identify additional economic evaluations. Reference lists of included publications were checked to identify other potentially relevant articles.

Titles and abstracts were screened against the eligibility criteria by one reviewer (KB) and a second reviewer (EF) screened 10% of records, which were randomly selected. Full text articles were retrieved for all publications judged as potentially relevant and assessed against the eligibility criteria for a second time.

3.2.3 Data extraction

Data were extracted by one reviewer (KB) into an Excel database designed for this review. The form was piloted on a subset of identified papers and several modifications made. The form was subsequently piloted again successfully. The form contained bibliographic information, descriptions of the aims and interventions evaluated, characteristics of the economic evaluation (e.g. model or trial based, type of analysis), the costs (by sector), type of outcomes included in the analyses and author conclusions. Additional fields assessed how the studies addressed the challenges of evaluating public health interventions that have been reported in the literature. These included fidelity, intervention development and modification, modelling of behaviour change and equity.

3.2.4 Analyses

A narrative was used to present the results. The descriptive aims of the review lend itself to this approach and the broad methodological diversity anticipated by the inclusion criteria means that statistical synthesis would be inappropriate. This is the most commonly used approach in systematic reviews of economic evaluations (Mathes et al., 2014). No quality assessment of the economic evaluations was conducted because the aim of this review was to understand the methods used in practice, not to establish the robustness of results to inform decision making.

3.3 Results

The process of study selection is shown in Figure 1. Searches of electronic databases identified 7,574 records, and after deduplication 5,744 unique records remained. After reviewing the titles and abstracts for eligibility, 164 full-text publications were retrieved and assessed in more detail against the inclusion criteria. Thirty of these were deemed eligible, and a further eight publications were identified from other sources. A total of 38 publications was included in the review. Four publications (Babey et al., 2014; Barrett et al., 2015; Brown et al., 2007; Cradock et al., 2014; Cradock et al., 2017; Dollahite and Hosig, 1998; Gortmaker et al., 2015; Mernagh et al., 2010; Segal et al., 2005) evaluated more than one intervention and reported the results of the economic evaluations separately. These were assessed as independent economic evaluations; therefore 46 studies are discussed in the narrative synthesis.

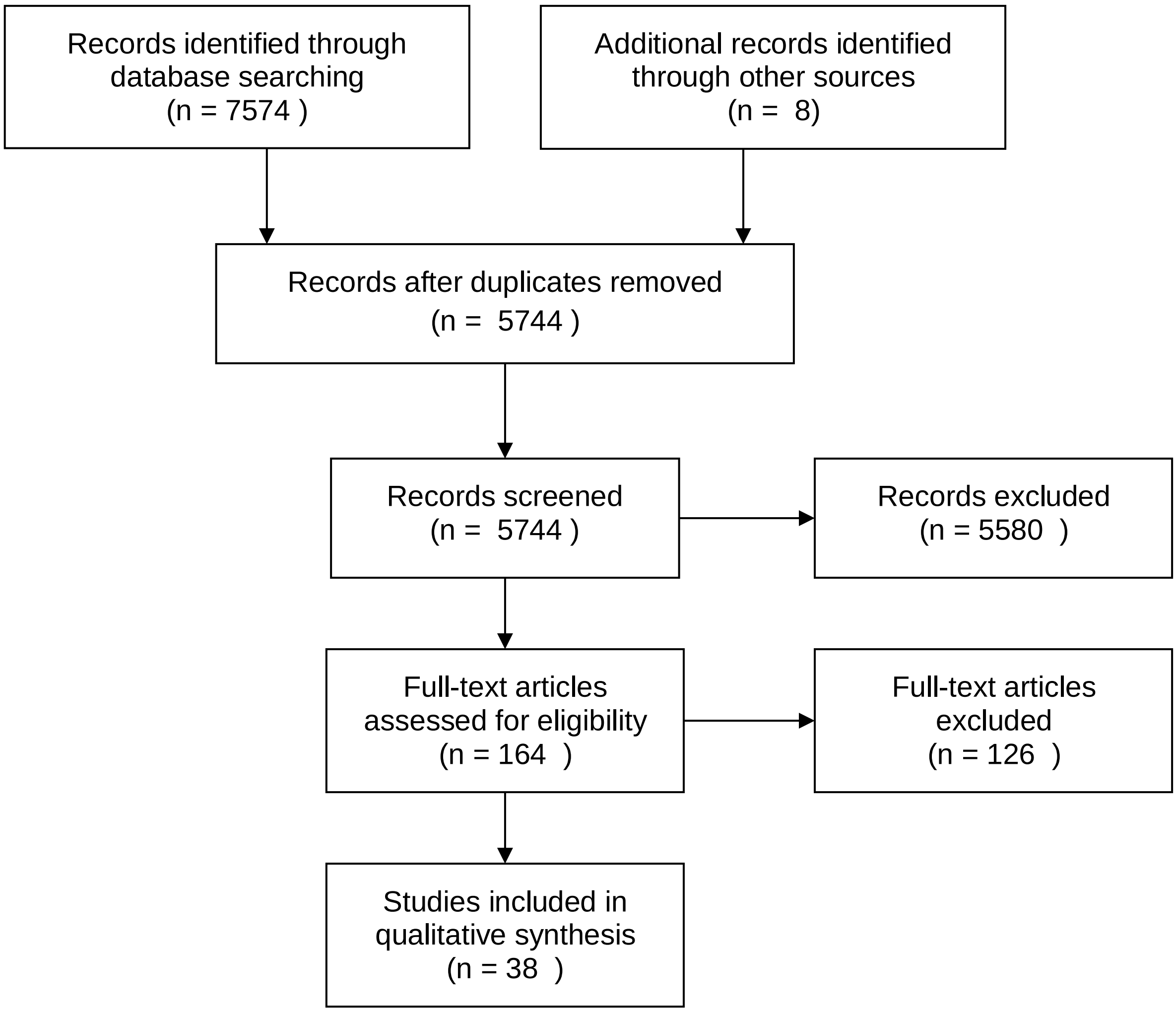

Figure 1 Flow diagram of study selection

3.3.1 Study Characteristics

Most analyses (18/46) were set in the USA (Babey et al., 2014; Barrett et al., 2015; Brown et al., 2007; Cradock et al., 2014; Dollahite and Hosig, 1998; Flores, 1995; Gortmaker et al., 2015; Graziose et al., 2017; Hendy et al., 2011; Ladapo et al., 2016; Manger et al., 2012; Reznik et al., 2015; Wang et al., 2017; Wang et al., 2003), followed by Australia (Eckermann et al., 2014; Moodie et al., 2011; Moodie et al., 2013; Segal et al., 2005; Waters et al., 2017) and the United Kingdom (Adab et al., 2018a; Lawlor et al., 2016; Vale et al., 2012; Wyatt et al., 2018). The majority were from the perspective of high-income countries (The World Bank, 2019), with the exception of one publication reporting results from the perspective of several countries (England, India, China, Mexico, Brazil, South Africa, Russia) (Cecchini et al., 2010) and two analyses conducted in China (Li et al., 2017; Meng et al., 2013). A minority (5/38) were published prior to 2008.

3.3.1.1 Interventions and comparators

All economic evaluations evaluated school-based interventions, targeted at children aged between 5 and 11. The interventions and comparators are summarised in Appendix B. One evaluated a community intervention but involved a school-based component (Moodie et al., 2013). Obesity prevention was the outcome predominantly targeted (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Cradock et al., 2017; Ekwaru et al., 2017; Gortmaker et al., 2015; Graziose et al., 2017; Hendy et al., 2011; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Segal et al., 2005; Tran et al., 2014; Wang et al., 2003; Waters et al., 2017; Wyatt et al., 2018). The remainder aimed to influence poor diet (Dollahite and Hosig, 1998; Eckermann et al., 2014; Li et al., 2017; te Velde et al., 2011; Vale et al., 2012), physical activity (Babey et al., 2014; Lawlor et al., 2016; Reznik et al., 2015; Wang et al., 2017; Wu et al., 2011) and cardiovascular risk factors (Flores, 1995; Segal et al., 2005). It is likely that the majority would therefore have some influence on excess weight. The majority involved multiple components (35/46), with single component interventions comprising either food policy changes (Vale et al., 2012), increasing physical activity alone (Babey et al., 2014; Barrett et al., 2015; Cradock et al., 2014; Reznik et al., 2015) or lesson-based education on diet (Graziose et al., 2017; Li et al., 2017) or physical activity (Moodie et al., 2011; Reznik et al., 2015). Several economic evaluations assessed adaptions or refinements of existing programmes. These included A Pilot Programme for Lifestyle and Exercise (APPLE schools) (McAuley et al., 2010; Mernagh et al., 2010), Alberta Project Promoting active Living and healthy Eating in schools (APPLE) (Ekwaru et al., 2017; Tran et al., 2014), Be Active Eat Well (BAEW) (Kesztyus et al., 2017; Kesztyus et al., 2013) and Active-PE, Active School Day, Active Recess (Barrett et al., 2015; Cradock et al., 2014; Cradock et al., 2017; Wang et al., 2017). Active-PE is a state-level policy devoting a minimum moderate to vigorous physical activity (MVPA) in PE lessons, which can be supplemented with a classroom curriculum and food policies. APPLE schools involve community coordinators in schools to encourage more physical activities and improvements in the availability of healthy food and water. The comparator intervention was ‘usual activities’ or ‘no intervention’ in most economic evaluations (37/46) (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Dollahite and Hosig, 1998; Eckermann et al., 2014; Ekwaru et al., 2017; Flores, 1995; Graziose et al., 2017; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Mora et al., 2015; Reznik et al., 2015; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012;

Wang et al., 2017; Wang et al., 2003; Waters et al., 2017; Wyatt et al., 2018). Alternative comparators included state policies on nutrition and tax (Gortmaker et al., 2015) and contrasting multicomponent interventions (Cradock et al., 2017).

The mechanisms of action were classified as either behaviour change (e.g. active travel, nutrition standards), education (e.g. grocery shop tours), or environment change (e.g. playground equipment). Most interventions were identified as using both education and behaviour change (24/46) (Adab et al., 2018a; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Eckermann et al., 2014; Ekwaru et al., 2017; Flores, 1995; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mernagh et al., 2010; Mora et al., 2015; Ohinmaa et al., 2011; Rush et al., 2014; Segal et al., 2005; Tran et al., 2014; Wang et al., 2003; Wyatt et al., 2018). Fifteen interventions used behaviour change alone (Babey et al., 2014; Barrett et al., 2015; Cradock et al., 2017; Gortmaker et al., 2015; Hendy et al., 2011; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Reznik et al., 2015; Vale et al., 2012; Wu et al., 2011). Only two interventions involved changing the environment, and these were combined with either behaviour change (Wang et al., 2017) or both behaviour change and education (Waters et al., 2017).

3.3.2 Economic evaluation characteristics

Study characteristics are summarised in Table 5. Partial economic evaluations accounted for nine studies reviewed (Cradock et al., 2014; Dollahite and Hosig, 1998; Flores, 1995; Hendy et al., 2011; Manger et al., 2012; Ohinmaa et al., 2011; Reznik et al., 2015; Waters et al., 2017). Seven were trials reporting costs (Cradock et al., 2014; Dollahite and Hosig, 1998; Flores, 1995; Hendy et al., 2011; Manger et al., 2012; Reznik et al., 2015; Waters et al., 2017), whilst two were cost-analyses (Ohinmaa et al., 2011; Wyatt et al., 2018). Trials reporting costs only were almost exclusively trials of multiple component interventions conducted in the USA (Cradock et al., 2014; Dollahite and Hosig, 1998; Flores, 1995; Hendy et al., 2011; Manger et al., 2012; Reznik et al., 2015). Although ROI is an approach being used increasingly in public health, only one ROI analysis was identified (Eckermann et al., 2014), assessing a national nutrition programme in Australia. No CBA were retrieved. The type of analyses used most frequently was a CEA (15/46) (Babey et al., 2014; Barrett et al., 2015; Cradock et al., 2017; Gortmaker et al., 2015; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; McAuley et al., 2010; Meng et al., 2013; Mora et al., 2015; Wang et al., 2017; Wu et al., 2011), followed by CUA (14/46) (Adab et al., 2018a; Brown et al., 2007; Cecchini et al., 2010; Ekwaru et al., 2017; Graziose et al., 2017; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Rush et al., 2014; te Velde et al., 2011; Wang et al., 2003). Four studies reported both a CEA and CUA (Li et al., 2017; Segal et al., 2005) and one a CCA and CUA (Vale et al., 2012).

Half of the studies (23/46) conducted economic modelling as a primary analysis (Babey et al., 2014; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2017; Ekwaru et al., 2017; Gortmaker et al., 2015; Graziose et al., 2017; Li et al., 2017; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Rush et al., 2014; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang et al., 2003; Wu et al., 2011), whilst four studies supplemented clinical trial data with a model (Li et al., 2017; Segal et al., 2005). Data from a cross-sectional study was used in one analysis (Ohinmaa et al., 2011). A model developed for the Assessing cost-effectiveness in obesity (ACE-Obesity) project in Australia was used as the basis for eight economic models (Barrett et al., 2015; Cradock et al., 2017; Gortmaker et al., 2015; Moodie et al., 2011; Moodie et al., 2013), each populating the model with data relevant to the intervention being evaluated. Simulation modelling was used to determine cost-effectiveness over one or two and 10-year time horizons, with the exception of two that modelled costs and outcomes over a lifetime horizon (Moodie et al., 2011; Moodie et al., 2013). A further eight studies used simulation modelling (Cecchini et al., 2010; Mernagh et al., 2010; Rush et al., 2014; Tran et al., 2014; Vale et al., 2012) and three reported decision trees (Brown et al., 2007; Graziose et al., 2017; Wang et al., 2003). Finally, Markov models were used in five studies (Ekwaru et al., 2017; Li et al., 2017; Segal et al., 2005). There was no evident trend of methods being used repeatedly for particular types of intervention. Of the trial-based analyses, 13/17 used findings from a cluster RCT (Adab et al., 2018a; Flores, 1995; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Li et al., 2017; McAuley et al., 2010; Meng et al., 2013; Mora et al., 2015; Reznik et al., 2015; Waters et al., 2017; Wyatt et al., 2018). As only school-based interventions were included, cluster trials would be the most appropriate design. Non-randomised designs were used in the remainder, and included comparisons against a historical control group, pre-post designs and observational studies (Hendy et al., 2011). The trial durations ranged from five weeks to six years, although most lasted one year. CUA was the primary analysis conducted in only one of the trial-based studies (Adab et al., 2018a).

Table 5 Study characteristics of included economic evaluations

	Number of studies (%)
Year published	
1995-2010	6 (15)
2011-2014	15 (39)
2015-2018	18 (46)
Type of economic evaluation	
Full	37
CEA	15 (41)
CUA	14 (38)
CEA/CUA	4 (11)
CCA	2 (5)
ROI	1 (3)
CCA/CUA	1 (3)
Partial	9
Trial reporting costs	7 (78)
Cost-analysis	2 (22)
Design	
Trial	19
Cluster RCT	13 (68)
Non-randomised cluster trial	3 (16)
Case control	1 (5)
Pre-post design 1 school	1 (5)
Cross-sectional review of costs	1 (5)
Model	27
Simulation	15 (56)
Markov	6 (22)
Cost-effectiveness ratios from estimates in the literature	2 (7)
Decision tree	3 (11)
Epidemiological	1 (4)
Based on ACE-Obesity	8 (3)

3.3.2.1 Perspective

As would be expected for public health interventions, most studies (18/46) used either a broad, societal (Barrett et al., 2015; Brown et al., 2007; Graziose et al., 2017; Kesztyus et al., 2017; Kesztyus et al., 2013; McAuley et al., 2010; Moodie et al., 2013; Ohinmaa et al., 2011; Segal et al., 2005; Wang et al., 2003) or modified societal perspective (Cradock et al., 2017; Gortmaker et al., 2015). The modified societal approach excluded productivity and patient costs. A healthcare perspective was used in 10 studies (Li et al., 2017; Mernagh et al., 2010; Rush et al., 2014; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wyatt et al., 2018), however four (Babey et al., 2014; Cecchini et al., 2010; Eckermann et al., 2014; Mora et al., 2015) did not report the perspective used (excluding the seven trials reporting costs). Three analyses took the school's perspective (Ekwaru et al., 2017; Ladapo et al., 2016; Lawlor et al., 2016), one of which combined it with a provider perspective (Lawlor et al., 2016) and four described programme (Wang et al., 2017), public health, public sector or social (Meng et al., 2013) perspectives.

What was included in a school perspective was explored. One study evaluating APPLE schools (Ekwaru et al., 2017) captured facilitator staff and manager salaries, cost of professional development, travel and research costs. These were described as programme costs. More detailed costs were reported in the evaluation of Students for Nutrition and eXercise (SNaX) (Ladapo et al., 2016). Teachers were intervention facilitators and their time to attend training and acting as peer leaders was costed. All materials (e.g. pens, handouts) and changes to the school environment (e.g. water filters) were included. Costs were classified as fixed (costs that do not change by number of pupils) and variable (costs that increase with pupils enrolling). Finally, the Active for Life Year 5 (AFLY5) intervention took a provider and school perspective (Lawlor et al., 2016). The length afforded by a health technology assessment report meant that the methodology and costs were highly detailed. Costs included were research staff and teacher time and expenses to attend training, time for staff to generate programme materials and the additional cost of an AFLY5 lesson, compared to a normal lesson.

As public health interventions incur costs and benefits across multiple sectors, this review examined whether costs were explicitly reported by sector (education, healthcare, government/local authority, criminal justice, parent/family), or were reported in enough detail that the applicable sector could be inferred. Twenty-one analyses did neither (Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Dollahite and Hosig, 1998; Eckermann et al., 2014; Gortmaker et al., 2015; Kesztyus et al., 2017; Kesztyus et al., 2013; Manger et al., 2012; Mernagh et al., 2010; Rush et al., 2014; Segal et al., 2005; Waters et al., 2017; Wyatt et al., 2018). Twelve reported costs by sector (Adab et al., 2018a; Cradock et al., 2017; Lawlor et al., 2016; Li et al., 2017; Moodie et al., 2011; Moodie et al., 2013; Ohinmaa et al., 2011; te Velde et al., 2011; Vale et al., 2012; Wang et al., 2003), and a further nine only included costs from one sector (Babey et al., 2014; Ekwaru et al., 2017; Hendy et al., 2011; Ladapo et al., 2016; Meng et al., 2013; Reznik et al., 2015; Tran et al., 2014; Wang et al., 2017; Wu et al., 2011). Four studies reported sufficiently detailed costs that could be attributed between sectors (Flores, 1995; Graziose et al., 2017; McAuley et al., 2010; Mora et al., 2015).

3.3.2.2 Costs

Costs of school-based interventions might include staff costs, materials and the opportunity cost of providing it. What costs were considered as programme costs was examined, where reported in sufficient detail. Costs of materials such as prizes or rewards (Flores, 1995; Hendy et al., 2011; Mernagh et al., 2010), workbooks (Kesztyus et al., 2017; Kesztyus et al., 2013), transport (Meng et al., 2013; Wu et al., 2011) and physical activity equipment (Barrett et al., 2015; Cradock et al., 2014; Cradock et al., 2017; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Mernagh et al., 2010; Mora et al., 2015; Wang et al., 2017) were included. The costs of facilitators and coordinators to organise and deliver the programme were reported by 11 studies (Barrett et al., 2015; Brown et al., 2007; Ekwaru et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Meng et al., 2013; Vale et al., 2012; Wang et al., 2017; Wang et al., 2003; Waters et al., 2017; Wu et al., 2011). The cost of teaching staff attending training was included in nine studies (Brown et al., 2007; Cradock et al., 2014; Cradock et al., 2017; Lawlor et al., 2016; Moodie et al., 2011; te Velde et al., 2011; Wang et al., 2003) and these costs included teachers' wages and additional staff required to teach in their absence. Training expenses incurred were also costed in many cases (Babey et al., 2014; Barrett et al., 2015; Cecchini et al., 2010; Flores, 1995; Hendy et al., 2011; Kesztyus et al., 2017; Kesztyus et al., 2013; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mora et al., 2015; Segal et al., 2005; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang et al., 2017; Wang et al., 2003; Wyatt et al., 2018), whilst in some cases the programme costs were clearly funded by the school (Adab et al., 2018a; Brown et al., 2007; Ekwaru et al., 2017; Graziose et al., 2017; Ladapo et al., 2016; Lawlor et al., 2016), in 21/46 analyses it was ambiguous (Babey et al., 2014; Barrett et al., 2015; Cecchini et al., 2010; Flores, 1995; Hendy et al., 2011; Kesztyus et al., 2017; Kesztyus et al., 2013; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mora et al., 2015; Segal et al., 2005; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang and Michael, 2015; Wang et al., 2003; Wyatt et al., 2018). Of the studies that did report the funder, they included public health (Wu et al., 2011), local or government funds or grants (Cradock et al., 2014; Cradock et al., 2017; Dollahite and Hosig, 1998; Eckermann et al., 2014; Gortmaker et al., 2015; Mernagh et al., 2010; Moodie et al., 2011; Rush et al., 2014) and combinations of education authorities, government and charities (Brown et al., 2007; Ohinmaa et al., 2011; Waters et al., 2017). The cost of developing the intervention was included in only six studies (Adab et al., 2018a; Dollahite and Hosig, 1998; Li et al., 2017; Meng et al., 2013; Moodie et al., 2013; te Velde et al., 2011; Wang et al., 2017), with one intervention already available (Wang et al., 2017).

Most analyses clearly included costs incurred directly by the school (34/46) (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Cradock et al., 2017; Dollahite and Hosig, 1998; Eckermann et al., 2014; Ekwaru et al., 2017; Flores, 1995; Graziose et al., 2017; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Moodie et al., 2011; Moodie et al., 2013; Mora et al., 2015; Ohinmaa et al., 2011; Reznik et al., 2015; Segal et al., 2005; te Velde et al., 2011; Wang et al., 2017; Waters et al., 2017; Wyatt et al., 2018). These were largely the cost of teachers' time for training, implementation and the cost of materials required for the intervention. Twenty-four studies in total included teacher's wages (Adab et al., 2018a; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Cradock et al., 2017; Graziose et al., 2017; Kesztyus et al., 2017; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Li et al., 2017; Meng et al., 2013; Moodie et al., 2011; Moodie et al., 2013; Segal et al., 2005; te Velde et al., 2011; Wang et al., 2017;

Wang et al., 2003; Wyatt et al., 2018), and one explicitly excluded these as external staff delivered the intervention (Mernagh et al., 2010). Costs reportedly incurred by the government or local authority were again programme costs (Cradock et al., 2017; Dollahite and Hosig, 1998; Mernagh et al., 2010; Ohinmaa et al., 2011), as well as the implementation and oversight of policies (Cradock et al., 2014; Cradock et al., 2017; Dollahite and Hosig, 1998; Gortmaker et al., 2015), capital costs of equipment such as kitchens (Eckermann et al., 2014) and training of staff at district or national level (Cradock et al., 2014; Gortmaker et al., 2015; Graziose et al., 2017; Vale et al., 2012).

Although this review focuses on school-based preventative interventions, the inclusion of healthcare costs was examined. The financial benefits of public health interventions are likely to be observed in the future, beyond data that can feasibility collected in a clinical trial of a school-based intervention. Indeed, the majority of healthcare costs were the estimated future costs of treating illness resulting from poor lifestyle and long-term healthcare predicted to be consumed by the general population. These were included in 17 studies (Barrett et al., 2015; Brown et al., 2007; Cradock et al., 2017; Gortmaker et al., 2015; Graziose et al., 2017; Li et al., 2017; Mernagh et al., 2010; Moodie et al., 2013; Rush et al., 2014; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang et al., 2003). Some interventions were delivered in part by health service staff, therefore labour costs were included (Mora et al., 2015; Waters et al., 2017; Wyatt et al., 2018). Finally, healthcare providers funded a minority of interventions (Mernagh et al., 2010; Moodie et al., 2013).

A minority (4/46) (Adab et al., 2018a; Cradock et al., 2017; Lawlor et al., 2016) reported parent/family costs incurred and were all published recently (since 2015). One study reported that these accounted for less than 1% of costs

(Cradock et al., 2017) and another 2% of costs (Cradock et al., 2017)yet did not specify what this encompassed. One study included parent time for doing homework activities (Lawlor et al., 2016) and another included parent costs in sensitivity analysis (Adab et al., 2018a). No studies included a cost for children's time participating in the intervention.

One study reported detailed unit costs, in addition to the annual incremental costs to each sector and its percentage of the total costs for a school community. The sectors included were health, local government, education, recreation and sport, commercial and other (Moodie et al., 2013). The health sector incurred the greatest proportion of costs (60%), and education the lowest (1.5%). This was because the multicomponent intervention was delivered by the health sector and was predominantly labour costs.

3.3.2.3 Outcomes

Of the primary analyses reported, the QALY was the outcome measure used most frequently (12/46) (Adab et al., 2018a; Brown et al., 2007; Ekwaru et al., 2017; Graziose et al., 2017; Mernagh et al., 2010; Rush et al., 2014; Vale et al., 2012; Wang et al., 2003; Wyatt et al., 2018). It was reported as the secondary outcome in three studies (Li et al., 2017; Segal et al., 2005). Only one study generated QALYs using primary utility data, using the CHU9D in a cluster RCT (Adab et al., 2018a). The remaining analyses either used utility estimates for adult chronic disease health states (Ekwaru et al., 2017; Graziose et al., 2017; Vale et al., 2012), HUI3 estimates identified in existing literature (Mernagh et al., 2010; Rush et al., 2014) or adjusted QALYs based on 2002 National Health Interview Survey (NHIS) Activity Limitations data (Brown et al., 2007; Wang et al., 2003). Change in BMI was the outcome used in nine studies (Gortmaker et al., 2015; Hendy et al., 2011; Manger et al., 2012; Meng et al., 2013; Mora et al., 2015;

Segal et al., 2005; Waters et al., 2017), with a further four studies estimating other weight related outcomes such as cases of obesity prevented (Kesztyus et al., 2017; Tran et al., 2014), centimetres of weight circumference or weight gain prevented (Kesztyus et al., 2013; McAuley et al., 2010). Seven used trial data (Hendy et al., 2011; Kesztyus et al., 2017; Kesztyus et al., 2013; McAuley et al., 2010; Meng et al., 2013; Mora et al., 2015; Waters et al., 2017), whilst the remainder identified estimates of weight change from literature using systematic or non-systematic methods (Gortmaker et al., 2015; Segal et al., 2005; Tran et al., 2014)(. Change in physical activity levels was another frequently used primary outcome (11/46) (Babey et al., 2014; Barrett et al., 2015; Cradock et al., 2014; Cradock et al., 2017; Flores, 1995; Lawlor et al., 2016; Reznik et al., 2015; Wang et al., 2017; Wu et al., 2011). Measures of metabolic equivalent (MET) hours (Babey et al., 2014; Barrett et al., 2015; Cradock et al., 2017; Wang et al., 2017; Wu et al., 2011) and time in MVPA (Cradock et al., 2014; Lawlor et al., 2016) were used in several studies, all of which were interventions involving a physical activity component. Used less frequently were outcomes such as change in food choices (Dollahite and Hosig, 1998), attitudes to physical activity (Flores, 1995), quantity of fruit and vegetables eaten (Eckermann et al., 2014; Ladapo et al., 2016; Waters et al., 2017), television viewing (Segal et al., 2005), enjoyment (Eckermann et al., 2014) and parents' knowledge (Dollahite and Hosig, 1998). These were predominantly secondary outcomes and those used in partial economic evaluations.

The incremental analyses of the full economic evaluations were examined. Of the CUA, the cost per QALY gained ranged from $275 (Graziose et al., 2017) to $307,552 (Mernagh et al., 2010). The intervention with the lowest cost per QALY was a food and nutrition education programme comprising 24 lessons delivered by teachers over one year (Graziose et al., 2017). The authors judged that the intervention would be cost-effective if it were to be delivered to all year 5 children in New York. The least cost-effective intervention reporting cost per QALY was APPLE (Mernagh et al., 2010). This is a 2-year, resource intensive multi-component intervention. Both of these analyses were model-based, using a lifetime horizon, discounted costs and outcomes and included the costs of future healthcare consumption. Secondary CUA of a multicomponent Australian obesity-prevention intervention reported an even larger ICER of $406,137 per QALY (Segal et al., 2005). Results of the primary CEA were $505 per BMI unit reduction. Neither CUA of interventions delivered in the UK would be judged as cost-effective, with ICERs of $32,166 (£46,083) (Adab et al., 2018a) and $37,204 (£52,400) (Vale et al., 2012) reported. The first was a trial evaluated a multi-component obesity intervention (Adab et al., 2018a) and the other, a model of a school food policy (Vale et al., 2012). While the obesity prevention intervention showed limited effectiveness and was not deemed cost-effective, authors of the policy intervention reported a need to consider the complex relationship between improving nutritional content of school-provided food, shifting preferences to packed lunches and potential impacts on inequalities.

Several CEA reported incremental results in BMI units, enabling some comparison of the results. Cost per decrease in BMI ranged from $6.10 (Gortmaker et al., 2015) to $504.56 in primary analyses (Segal et al., 2005). Where BMI was the secondary outcome, an ICER as high as $2825 was reported (Cradock et al., 2014). The most cost-effective intervention was federal nutrition standards for school meals and was judged to be cost saving (Gortmaker et al., 2015). The least cost-effective of $371 was discussed previously, having the greatest cost per QALY also (Segal et al., 2005). The Active School Day had the greatest ICER ($2825), with the intervention involving at least 150 minutes of MVPA per school day. Whilst estimated MVPA increased, the estimated impact on BMI was small. Conceptualising the cost per MET hour or MVPA unit is challenging as this review found that cost effectiveness ratios are small compared to results in BMI or QALY units. The majority of the analyses using measures of physical activity evaluated several interventions and only reported cost-effectiveness ratios, without conducting incremental analyses. All interventions aimed to increase physical activity, and all were judged to be cost-effective. The lowest ICER was $0.01 per MET hour gained (short physical activity breaks) (Babey et al., 2014), and highest $1.05 (Active School Day) (Cradock et al., 2014).

3.3.3 Public health considerations

3.3.3.1 Broader outcomes

As public health interventions might have benefits beyond the immediate participants and primary outcome, the consideration of broader outcomes was examined. Thirty-four studies did not analyse or discuss this issue (Adab et al., 2018a; Babey et al., 2014; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2014; Cradock et al., 2017; Ekwaru et al., 2017; Flores, 1995; Gortmaker et al., 2015; Graziose et al., 2017; Hendy et al., 2011; Kesztyus et al., 2013; Ladapo et al., 2016; Lawlor et al., 2016; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mernagh et al., 2010; Moodie et al., 2013; Mora et al., 2015; Ohinmaa et al., 2011; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang et al., 2017; Wang et al., 2003; Wu et al., 2011; Wyatt et al., 2018), however four studies acknowledged this in the publication (Kesztyus et al., 2017; Segal et al., 2005). The impact on families was considered by four studies, including analysing family food habits. One study estimated the impact of the intervention on lowering blood pressure in all family members (Li et al., 2017) and two assessed changes in parents' food knowledge or habits (Dollahite and Hosig, 1998; Waters et al., 2017). Two studies also reported benefits to the wider community (Eckermann et al., 2014; Moodie et al., 2011). A ROI analysis evaluated the involvement of community volunteers and their contribution to the economic viability of the programme (Eckermann et al., 2014).

3.3.3.2 Fidelity

Intervention fidelity is an important consideration, particularly in behavioural and educational programmes. Schools may not deliver interventions consistently and some individuals may not participate as intended. The majority of analyses did not account for any variation in fidelity (37/46) (Adab et al., 2018a; Babey et al., 2014; Brown et al., 2007; Cradock et al., 2014; Cradock et al., 2017; Eckermann et al., 2014; Ekwaru et al., 2017; Flores, 1995; Gortmaker et al., 2015; Graziose et al., 2017; Hendy et al., 2011; Kesztyus et al., 2013; Ladapo et al., 2016; Li et al., 2017; Manger et al., 2012; McAuley et al., 2010; Meng et al., 2013; Mernagh et al., 2010; Moodie et al., 2011; Mora et al., 2015; Ohinmaa et al., 2011; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012; Wang et al., 2017; Wang et al., 2003; Wu et al., 2011). One study involved schools developing their own approaches, therefore fidelity would be difficult to account for (Kesztyus et al., 2017). Three trials conducted and reported process evaluations, although no findings were used to inform the economic analyses (Dollahite and Hosig, 1998; Lawlor et al., 2016; Wyatt et al., 2018). Methods of accounting for fidelity included estimating the proportion of children that fully participated in the intervention (Cecchini et al., 2010), estimating 71% of teachers implemented the intervention (Cradock et al., 2017), estimating 72% compliance with the intervention (Barrett et al., 2015) and assuming that only 50% benefitted

(Moodie et al., 2013). All of these were model based analyses, and the sources of the estimates were unclear. One trial-based analysis compared results at high and low levels of fidelity, using findings from a RE-AIM evaluation (Reznik et al., 2015).

3.3.3.3 Modelling of behaviour change

Public health interventions would usually require individuals to maintain behaviour changes in order to have an enduring impact on health. Whether the analyses considered this was evaluated. Eleven of the 46 studies did incorporate the maintenance of effect into the analysis, all of which used modelling (Graziose et al., 2017; Li et al., 2017; Mernagh et al., 2010; Moodie et al., 2013; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011). One study assumed a 50% decrease in adherence to the intervention (Li et al., 2017). A 50% relapse rate by those becoming normal weight was estimated in two analyses (Segal et al., 2005), and one estimated that 50% would relapse by aged 40 (Graziose et al., 2017). One study assumed a 30% maintenance of effect (Moodie et al., 2013), whilst another did a threshold analyses and determined maintenance needed to be 70% for the intervention to be cost-effective (Rush et al., 2014). Five studies, all of which used the same model, assumed that BMI improvement decayed by 1% per annum after the first five years (Mernagh et al., 2010).

3.3.3.4 Time horizon and discounting

The health benefits of public health interventions might not be observable until long after the intervention has been delivered. The time horizon of the economic evaluation is therefore important, as is the use of discounting costs and outcomes. A lifetime horizon was used by 13 studies (Cecchini et al., 2010; Ekwaru et al., 2017; Graziose et al., 2017; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Rush et al., 2014; te Velde et al., 2011; Tran et al., 2014; Vale et al., 2012), and a year/school year in ten (Adab et al., 2018a; Babey et al., 2014; Kesztyus et al., 2017; Kesztyus et al., 2013; Lawlor et al., 2016; Meng et al., 2013; Ohinmaa et al., 2011; Wang et al., 2017; Wu et al., 2011; Wyatt et al., 2018). A horizon of less than one year was used in six studies (Cradock et al., 2014; Dollahite and Hosig, 1998; Flores, 1995; Hendy et al., 2011; Ladapo et al., 2016; Reznik et al., 2015), and ranged from five weeks (Ladapo et al., 2016) to five months (Dollahite and Hosig, 1998), although these were mostly partial economic evaluations of trials that reported costs.

The discounting approach used by the studies was examined. Both costs and outcomes were discounted and at the same rate in 22 studies (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Ekwaru et al., 2017; Graziose et al., 2017; Li et al., 2017; McAuley et al., 2010; Mernagh et al., 2010; Moodie et al., 2011; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011; Vale et al., 2012; Wang et al., 2003; Waters et al., 2017). In five studies, costs were discounted but whether outcomes were also discounted was indeterminable (Cradock et al., 2017; Gortmaker et al., 2015). Costs were discounted in the majority of studies that used a time horizon of over one year (27/30) (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Cradock et al., 2017; Ekwaru et al., 2017; Gortmaker et al., 2015; Graziose et al., 2017; Li et al., 2017; McAuley et al., 2010; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011; Vale et al., 2012; Wang et al., 2003; Waters et al., 2017). A discount rate of 3% was used in 15 studies (USA (Barrett et al., 2015; Brown et al., 2007; Cradock et al., 2017; Gortmaker et al., 2015; Graziose et al., 2017; Wang et al., 2003) Canada (Ekwaru et al., 2017) Australia (Moodie et al., 2011; Moodie et al., 2013) the Netherlands (te Velde et al., 2011) multiple countries (Cecchini et al., 2010)) , and 3.5% in seven (UK (Adab et al., 2018a; Vale et al., 2012) and New Zealand (Mernagh et al., 2010; Rush et al., 2014)). This is the recommended rate for USA cost-effectiveness analyses (Sanders et al., 2016). A minority (5/27) used a rate of 5%, which were analyses conducted in either New Zealand (McAuley et al., 2010) or Australia (Segal et al., 2005; Waters et al., 2017) and all discounted outcomes at 5% also. New Zealand recommend a discount rate of 3.5% for costs and benefits, whereas in Australia it is 5% (Attema et al., 2018). The UK recommends a discount rate of 3.5% for both costs and outcomes, however for public health interventions a rate of 1.5% can be used in sensitivity analysis (NICE, 2014). This was not adopted in the UK based studies (Adab et al., 2018a; Vale et al., 2012). Outcomes were discounted in 22 analyses (Adab et al., 2018a; Barrett et al., 2015; Brown et al., 2007; Cecchini et al., 2010; Ekwaru et al., 2017; Graziose et al., 2017; Li et al., 2017; McAuley et al., 2010; Mernagh et al., 2010; Moodie et al., 2011; Moodie et al., 2013; Rush et al., 2014; Segal et al., 2005; te Velde et al., 2011; Vale et al., 2012; Wang et al., 2003; Waters et al., 2017), with the seven analyses using a rate of 3.5% for costs and outcomes (Adab et al., 2018a; Mernagh et al., 2010; Rush et al., 2014; Vale et al., 2012).

3.3.3.5 Equity subgroup analysis

Whether the economic evaluations accounted for or explored the impact of biological or social variation was examined. Although these factors were considered in many analyses by including covariates such as age and gender, only two reported results by subgroup, both by age group (McAuley et al., 2010; Rush et al., 2014). Equity is also an important factor in public health. Four of the interventions evaluated were targeted at deprived groups (Dollahite and Hosig, 1998; Flores, 1995; Reznik et al., 2015; Waters et al., 2017), such as being delivered in low income or rural areas. One study included lessons in Spanish for

Hispanic pupils (Reznik et al., 2015). As the interventions were targeted, no subgroup analyses were conducted or comparison with less deprived groups reported. Nine studies did conduct subgroup analysis and presented the findings in a way that could be used to make recommendation for different groups (Brown et al., 2007; Graziose et al., 2017; Lawlor et al., 2016; Mernagh et al., 2010; Mora et al., 2015; Rush et al., 2014). These included ethnic groups such as Hispanic, Black and Maori participants. Other disadvantaged groups analysed were the less educated and immigrants (Mora et al., 2015). Some studies using the ACE-Obesity model modelled uptake differences for the lowest socio-economic groups (Mernagh et al., 2010; Rush et al., 2014).

3.4 **Discussion**

3.4.1 Summary of principal findings

This review identified 46 full or partial economic evaluations of school-based interventions for obesity prevention or to reduce risk factors for obesity. CEA and CUA were the approaches most commonly employed, with no CBA identified. Outcomes were predominantly weight (BMI, waist circumference), measures of physical activity (MET-hours, MVPA) or QALYs. A minority collected primary data to generate QALYs, however. Costs were rarely reported by sector; although the appropriate sector could be identified where detailed costs were reported. Labour costs tended to contribute most. Whilst all the interventions were school-based, the intended funder of the intervention was ambiguous in many economic evaluations. Good practice in regard to accounting for fidelity (e.g. estimate 71% of teachers implement it), equity (e.g. present estimates of cost-effectiveness for deprived groups) and maintenance of effect/behaviour change (e.g. assume 30% maintenance of effect) were identified and could be feasibly replicated where data allows.

3.4.2 *Strengths and limitations*

This review focused on a number of different types of interventions targeting different behaviours, delivered in a school setting. Previously, most literature has focused on one behaviour with no restriction on the setting. This has enabled a focused review of the methods relevant to schools, instead of an interrogation of the characteristics of cost-effective interventions for childhood obesity for example. Furthermore, no date limitations applied, and a wide variety of sources were searched, including education databases. Whilst these did not contain any unique records, it suggests that commonly used medical databases adequately index public health interventions delivered outside of the health sector. A number of relevant clinical trial protocols were identified, and searches were undertaken to ascertain whether associated economic findings were since published. non-English language publications were excluded for pragmatic reasons. This certainly excluded one relevant analysis; however the intervention was refined, re-evaluated and an economic evaluation subsequently published in English.

The broad range of eligible interventions may have implications for the homogeneity of methods used. The most frequently used outcomes were QALYs, physical activity and weight related measures and the majority of interventions were multi-component. A generic outcome such as the QALY is probably most suitable for interventions with several interacting components, as opposed to an outcome that robustly measures just one aspect, such as physical activity. For physical activity interventions, change in MET hours or MVPA is a logical and widely used outcome, which should be observable within a short time period. Its standardisation means that long-term changes could be modelled readily, drawing on evidence from existing studies. Dietary interventions on the other hand may have less standardised outcomes, such as change in fruit and vegetable consumption for example. These are less amenable to modelling yet change in BMI could be a longer-term indicator of success. The costs measured might also depend on the type of intervention. Physical activity interventions could require initial investment in equipment, but little long-term expenditure or delivery requiring specialist knowledge. Behavioural interventions might require expertise, therefore have initial and ongoing training requirements due to staff turnover, in addition to ongoing resource for delivery. Expertise may be drawn on from other sectors, such as health care for example. Despite this, the economic evaluations identified included similar costs (e.g. staff time) and there was no tendency for modelling to be used for particular interventions.

Since this review was conceptualised two similar reviews have been published (Batorova and Sørensen, 2019; Oosterhoff et al., 2018). Despite Oosterhoff et al. (2018) specifying similar populations, interventions, comparators and outcomes; this review identified 11 more full economic evaluations. Only three of these were published since Oosterhoff et al. (2018) conducted their searches (March 2017). The review of school-based lifestyle interventions reported key challenges and potential solutions. Important issues included a lack of clarity regarding intended audience, comparators, time horizon and scope of costs and outcomes, extrapolation of effects and the use of CEA preventing comparisons. These are mostly domains included in quality assessment checklists (Drummond et al., 2005; Husereau et al., 2013), thus are not novel challenges in a school setting. This review identifies similar variation in methods, although many might be due to the challenge of this unusual environment, where priorities of funders and schools are not aligned and the QALY holds less relevance. Batorova and Sørensen (2019) had a narrower focus on physical activity interventions only and those using modelling methods. All eight models were included in the review and criticisms of the analyses predominantly align with the findings. Concerns included the estimation of long-term effectiveness, assumptions regarding the maintenance of effect and the omission of children's time.

3.4.3 Findings in context

The majority of studies included in this review were published after the Weatherly et al. (2009) recommendations for the conduct of public health economic evaluations were published. The recommendations were generated based on a review of published economic evaluations. The proportion of studies conducting CUA seems to have increased since 2009 (27% v 39%), yet the proportion based on RCT evidence is lower (38% v 28%). This might be due to the advancement of modelling methods or expansion of expertise in this method, as the use of microsimulation modelling appears to have accelerated in the last ten years. The Australian ACE-Obesity microsimulation model was the basis for several analyses (Cradock et al., 2017; Gortmaker et al., 2015; Mernagh et al., 2010; Moodie et al., 2011; Rush et al., 2014), having inspired an American adaptation (Childhood Obesity Intervention Cost-Effectiveness Study, CHOICES). The ACE-Obesity model is accompanied by an extensive technical document that details the rationale for the methods used for the chronic disease prevention model. This transparency has allowed the replication of the analyses in different settings, and some of the methodological choices in some analyses have been highlighted positively, for example assuming a decay of effect and less than perfect implementation fidelity. While this approach could provide a framework for evaluating school-based interventions, many of the model parameters are from adult studies or extrapolate small changes in BMI to DALYs. Producing appropriate data to populate these models is therefore important. Nevertheless, one step recommended by Squires et al. (2016) in their guidelines for developing models of public health interventions is to assess the availability of existing models (Squires et al., 2016) to reduce the opportunity cost of developing new ones.

This review evaluated how the studies addressed the reported challenges from this review and other studies. Half of the analyses did not even discuss equity. Using Cookson's categorisations (Cookson et al., 2017), a minority of studies conducted a basic equity impact analysis by presenting results by deprived subgroup. None used equity trade off analysis or used equity weights. Some studies did attempt to account for variation in fidelity and a decay in effect over time. Where process evaluations were conducted alongside trials, in some cases the findings were used to inform intervention fidelity in the economic evaluation. If resources allow, process evaluations could be commissioned with the combined aim of assessing the implementation, delivery, uptake and real-world cost-effectiveness of the intervention.

3.4.4 Recommendations for future research

The findings of this review could help inform the conduct of future economic evaluations in a number of ways. Firstly, a variety of outcomes were used in the economic evaluations reviewed, including physical activity, QALYs, DALYs and indicators of weight. Whilst measures of physical activity might not appeal to national decision makers such as NICE, they could resonate with those implementing and funding physical activity interventions at a local level. In addition, physical activity might not have an immediate effect on health outcomes so might be the most appropriate short-term indicator of success, particularly when physical activity alone is targeted. The long-term implications for health could then be modelled. Collecting quality of life data that would satisfy decision makers would be worthwhile where possible too. Preference based measures are now available for children, therefore these could be implemented and CUA presented alongside CEA findings. Only one study collected CHU9D data (Adab et al., 2018a). The openness of NICE to CCA suggests that collecting a range of outcomes and presenting them in a disaggregated way could be particularly appropriate in this setting. This was an approach previously recommended (Weatherly et al., 2009). Broader impacts on the family and community could then be presented for consideration by the decision maker. Regarding the decision maker, this should be made clear in the analyses. The ambiguity surrounding the eventual funder of the interventions makes it challenging to understand the real-world cost-effectiveness and judge the sustainability. As socio-demographic data is usually collected in a trial, results could be relatively easily analysed by subgroup if it is postulated that the interventions will reduce or increase inequities.

This review found that both trial and model based economic evaluations have been used to assess school-based interventions for preventing obesity or addressing its risk factors. As would be expected, the time horizon of the trial-based studies was short, preventing the understanding of the long-term cost-effectiveness of interventions. In the UK, data linkage opportunities are emerging, particularly between the NCMP and NHS data. The enduring impact of interventions on BMI might be possible to monitor in the future and used to inform economic evaluations. This would not address the paucity of data on the quality of life impacts, used to generate QALYs however. Collecting preference-based utility data directly from children should be prioritised as relying on adult values or estimating QALYs from BMI data is unlikely to accurately reflect the true impact of interventions in this population. Future research could explore why this preference-based data is not currently being collected in this setting, and perhaps what outcomes are important in a school setting.

3.5 **Conclusion**

This review has identified that methods used to conduct economic evaluations of school based public health interventions are varied. There are many recommendations for how to conduct economic evaluations in this context, yet the adoption of guidelines and methodological recommendations are not widespread. However, due to the multisectoral nature of public health interventions it is perhaps unrealistic that all will be appropriate to a school setting.

Health and wellbeing activities divert time and resources away from what may be perceived as schools' main goal of achieving academic attainment (Littlecott et al., 2018). This is despite evidence that health improvement initiatives do not have a detrimental effect on academic performance in secondary school and they could even have a positive effect (Littlecott et al., 2018; Watson et al., 2017).

The benefits that schools would hope to observe by implementing health promotion programmes are unlikely to be comparable to the outcomes expected from NHS commissioned services delivered in primary and secondary care The NHS focuses on improving or maintaining individuals' health and wellbeing, normally by treating pre-existing ill health (Department of Health and Social Care, 2015). The expected benefits are well defined and with established approaches to measurement. Costs are usually expended from the NHS budget and immediate and future costs and health benefits also observed within the health sector (e.g. blood pressure treatment might prevent future strokes). There may be broader cost implications, for example in social care, but these could be estimated and included in evaluations of cost-effectiveness. Cost-effectiveness is an important consideration and decision making is informed by economic evidence. The opportunity cost of health care is also well-established in the UK through the NICE decision making framework with clear guidance on threshold willingness to pay values to judge cost-effectiveness of new or existing interventions

As discussed in Chapter 1, local authorities have responsibility for commissioning a range of public health services. These include community interventions for obesity prevention and alcohol and drug misuse services. Like the NHS services mentioned previously, those delivering the services would likely be funded with the goal of achieving pre-defined outcomes. These might be measured robustly, and the providers' performance judged against them. While there is less dependency on evidence of cost-effectiveness, costs of commissioning and indicators of benefits are considered. Benefits could be both long and short term and dispersed across society and this is recognised as an inevitable consequence of providing these services.

Public health interventions delivered in schools are not aimed at treating ill health necessarily and take a population level approach. As discussed in Chapter 1, schools' primary goal is educational attainment, yet they incur the opportunity cost of delivering public health interventions (Bonell et al., 2014). Although it is recognised that health and wellbeing contribute to children's development and is an important component of education delivered by schools, as evidenced by OFSTED requirements (OFSTED, 2019), this is not reflected in the metrics that schools are most frequently judged against. The benefits of interventions that schools provide may manifest well beyond the children's education, as may any financial benefits, for example health care resources reduced as a consequence of obesity prevention. In contrast, wellbeing public health interventions may have more immediate benefits and implications for academic attainment, although the primary aim would not be to increase a child's SATs score. Public health interventions may have additional benefits for school staff and families also.

Costs incurred to the education sector and beyond may not be adequately captured using the 'traditional' approach to evaluation associated with NICE technology assessment. These approaches would only include healthcare resources, which would neglect resource use incurred by schools, such as expenditure on equipment or staff wages. Published trial based economic evaluations identified in this review have included school staff wages for training only (Cradock et al., 2017; Ladapo et al., 2016), and some explicitly state the assumption that implementation of the intervention would be covered by staff wages (Cradock et al., 2017; McAuley et al., 2010), ignoring issues of opportunity cost. Ladapo et al. (2016) took a 'school' perspective, and it appears that the school funded the programme and training costs. In many other cases, however, the responsibility for funding the programme is ambiguous (Manger et al., 2012; McAuley et al., 2010).

Chapter 2 ended by highlighting several factors that appear to distinguish the school setting from healthcare, and it could be hypothesised that QALYs and BMI are not aligned with the transitional nature of schools or academic benefits that are potentially more desired in this setting. Furthermore, the finding that the intended funder of the intervention is usually unspecified indicates that the audience for the analyses is uncertain. The ambiguity of funding of school-based public health interventions in England noted in Chapter 1 perhaps explains this for UK based studies. Identifying who the decision maker is would enable an assessment of the suitability of the analyses for their needs, and the development of appropriate evidence. There is a body of implementation science literature examining the implementation and sustainability of health promotion in schools

(Darlington et al., 2018; Day et al., 2019; Friend et al., 2014), but this does not provide insights into decision making and evidence requirements before implementation. Qualitative research has indicated that funding is a barrier to the sustainability of school-based interventions (Day et al., 2019), yet a Delphi study Morton et al. (2017) indicated that when considering interventions, cost is the least important consideration for stakeholders (stakeholders being local authority/public health, education professionals/governors, parents and students). In the context of a physical activity intervention; mental health, wellbeing and academic attainment were the most desirable outcomes, with physical activity rated fifth of six. While this Delphi is informative, whether prioritisation would involve all these stakeholders, or their roles are unknown. It is unlikely they would all have an awareness of contextual factors such as available funding or academic needs and the feasibility of including all stakeholders in school-level decision making is uncertain. These results also conflict with the assumption that academic attainment is the school's focus and warrant the need for further research.

In a healthcare setting the decision maker and their evidence requirements are generally well known, accordingly the methods to choose and estimate the appropriate costs and outcomes are well developed. In a school setting however, neither appears to be true. These are issues that cannot be understood by examining policies and regulations, nor conducting a systematic review. Rather, qualitative research is necessary to examine these issues. Qualitative research uses an inductive, theory building approach. It is not a hypothesis testing exercise, rather it is an exploratory process whereby individuals' perspectives are elicited to generate an understanding of a topic (Braun and Clarke, 2006). These perspectives might differ between different individuals and interviewing a range of school staff would facilitate an understanding of these issues. In this case conducting a Delphi study, like Morton et al. (2017) would be inappropriate as the intention is not to achieve consensus, but rather to explore in an open way issues such as the decision making process used by participants and their own understandings of the costs and benefits of school based public health interventions.

3.6 Questions this thesis addresses

It has emerged that schools are a complex setting in which to conduct economic evaluations of public health interventions. Methods are varied in existing studies, suggesting there is no established approach and simply applying methods used in health is inappropriate. Key unanswered questions that should be pursued to inform relevant economic evaluations of school based public health interventions and which will be explored in this thesis are as follows:

- What is the decision-making structure currently employed in schools?
- What are the important outcomes in this setting and how can they be measured?
- What are the costs of implementing public health interventions in schools?

These questions will be considered in the context of health promotion initiatives delivered in primary schools and Chapter 4 describes the Birmingham Daily Mile trial, which forms a case study in which these questions are explored. This is followed by a qualitative study with school staff involved in the trial, which explores decision making and priority setting in schools and is used to provide further context to understand the role and design of economic evaluation within a school setting.

CHAPTER 4 THE DAILY MILE AND BIRMINGHAM DAILY MILE TRIAL

The Daily Mile initiative, evaluated within the Birmingham Daily Mile trial is a case study in which the research reported in Chapters 5, 6 and 7 are nested. This chapter describes the Daily Mile, the rationale for the trial, the methods used and the clinical effectiveness results.

4.1 The Daily Mile

The Daily Mile (The Daily Mile Foundation, 2018b) is an initiative developed and first implemented in a primary school in Stirling, Scotland in 2012. Initially designed to improve pupil's fitness, it is a simple and inclusive activity that increases children's physical activity in school. In addition to improving physical health, further possible benefits include improved wellbeing and academic achievement. The Daily Mile involves children doing an extra 15 minutes of activity by running or walking around a track within the school grounds. The 15 minutes reflects a distance of approximately one mile. Teachers can choose to do The Daily Mile at any time during the school day and in almost any weather, however it is not supposed to replace PE, break times or take place before or after school (The Daily Mile Foundation, 2018c). This leaves lesson time as the only opportunity in which to undertake the Daily Mile.

The UK Government's Childhood Obesity Plan (HM Government, 2018a) recommends that every primary school adopts the Daily Mile as part of a strategy to maintain a healthy weight. This affects a potential population of 4.7 million children aged 5-11, attending 16,700 state-funded primary schools (Department for Education, 2019b). Physical inactivity is a contributor to the energy imbalance responsible for excess weight (Sahoo et al., 2015), therefore it is hypothesized that the Daily Mile could provide a simple solution to the childhood obesity crisis. Despite its endorsement in UK Government Policy, there has been no robust evidence of either its effectiveness or cost-effectiveness. One small, non-randomised repeated measures pilot study was conducted in two Scottish primary schools, reporting improved fitness, measured using shuttle runs, but no significant effect of the Daily Mile on BMI (Chesham et al., 2018). Two qualitative studies conducted alongside this pilot, reported positive views on the health benefits but challenges in the Daily Mile's implementation (Malden and Doi, 2019; Ryde et al., 2018). The Daily Mile's implementation remains a recommendation and not a mandatory requirement.

4.2 **The Birmingham Daily Mile Trial**

4.2.1 Rationale for the trial

Whilst unlikely to cause harm, the Daily Mile is at the expense of the opportunity cost of missing lesson time, with potentially no benefit. RCTs are a method by which interventions can be tested rigorously to establish whether they are effective and cost-effective (Sibbald and Roland, 1998). To this end, the Birmingham Daily Mile cluster RCT was conceived to test whether the Daily Mile has an effect on children's BMI, wellbeing and whether it is cost-effective (Breheny et al., 2018).

The published protocol upon which this chapter is based describes the trial methodology (Breheny et al., 2018) and was published in advance of the analysis.

4.2.2 Methods

4.2.2.1 Design

The Birmingham Daily Mile trial was a pragmatic cluster RCT conducted in 40 state-funded primary schools in Birmingham, England. Cluster RCTs randomise groups (clusters) to an intervention or control arm, in this case the cluster being schools and the intervention the Daily Mile. This design is used when it is unfeasible to randomly allocate individuals to an intervention and when it is important to minimise contamination. As school based health promotion interventions are generally delivered at the school level, cluster RCTs are usually the most appropriate design in this context (Puffer et al., 2005). Indeed, this was the case with the Daily Mile.

While cluster RCTs may be the most appropriate design, there are associated statistical issues that necessitate larger sample sizes than would usually be required to achieve the same statistical power (Puffer et al., 2005). Individuals within clusters tend to be similar, which violates the assumption that individuals are uncorrelated. For this reason, clustering must be taken into account in the analysis. This requires more complex statistical approaches such as multilevel modelling, and not all software functionalities can be adapted for cluster trial analysis.

Ethics approval was obtained from the University of Birmingham Research Ethics Committee (reference number 16–0064). This study is registered with ISRCTN (12698269).

4.2.2.2 Participants

All state-funded primary and junior schools (children aged four-eleven years) located in South Birmingham with at least 20 children in school years three (aged seven-eight years) and five (aged nine-ten years) at baseline were eligible to participate. Outcomes were measured only from children in these two year-groups, although the whole school participated in the Daily Mile in the intervention arm. Children with a disability preventing them from running or walking for 15 minutes were excluded, as were those unable to have their height and weight measured at baseline. Written consent for outcome measures was obtained from parents/guardians and verbal assent obtained from eligible children.

4.2.2.3 Randomisation

After baseline data collection, schools were randomised in a 1:1 allocation ratio to either the intervention (Daily Mile) or the control condition. A constrained randomisation algorithm was used, which contained pre-specified school level covariates (school average BMI z-score (BMIz), percentage of pupils eligible for free school meals, and school size). This ensured allocation to the different arms was balanced on these characteristics. All analyses included these variables as covariates. Due to the nature of the intervention it was not possible to mask school staff, children, family members and project staff to the intervention allocation.

4.2.2.4 Intervention

Schools were provided with information regarding The Daily Mile and directed to the Daily Mile website (The Daily Mile Foundation, 2018a) for further guidance and resources. Whilst advised as a daily activity, the frequency and duration were at the school and class teacher's discretion. The control arm received no active intervention. Schools continued with usual health and wellbeing activities, and were requested not to implement new health or physical activity initiatives for the duration of the study. Currently the amount of physical activity that primary schools should provide in the UK is not mandated. At least two hours a week has been recommended by OFSTED (OFSTED, 2013), although more recently, the UK Government Childhood Obesity Strategy states that schools should provide 30 minutes of moderate to vigorous activity daily (HM Government, 2018a).

4.2.2.5 Outcomes

The primary outcome for clinical effectiveness was BMIz at the 12-month follow-up. Secondary outcomes collected at four months were BMIz, fitness (linear track test), and body fat percentage. Secondary outcomes at 12-months were fitness, body fat percentage, child-reported quality of life (Child Health Utility 9D (CHU9D) (Stevens, 2009), child-wellbeing (MDI) (Schonert-Reichl et al., 2012) and teacher-rated academic attainment (overall attainment and attainment in math, reading and writing). There is ongoing work to develop a core outcome set for school-based trials of childhood obesity interventions (COMET, 2017). In the absence of this guidance and in light of the claims made regarding the benefits of the Daily Mile, the rationale for the chosen outcomes were as follows.

BMIz and body fat percentage are objective measures of adiposity. It had been claimed that the Daily Mile has the potential to be part of the solution to childhood obesity and is highlighted in the Government's Childhood Obesity Plan (HM Government, 2018a). BMI is used routinely in trials of childhood obesity prevention interventions and is a practical measure where more accurate imaging such as DEXA scans are unfeasible (Adab et al., 2018b). BMI is also directly linked to cardiovascular risk factors (Sardinha et al., 2016). The linear track test is a simple measure of endurance, which was anticipated to provide an indication of the impact of additional 15 minutes of exercise on children's fitness. Another reported benefit of the Daily Mile is the positive impact on children's academic attainment. As already discussed, academic attainment is an important metric in the assessment of school performance. This suggested it was a valuable measure to include and hypothesis to examine. In the absence of standardised tests for all participating age groups over the course of the trial (e.g. SATs), teacher ratings of academic progress were obtained.

The CHU9D is described in detail in Section 6.2.2.1.3, but briefly it comprises nine questions about aspects of children's lives which contribute to their quality of life. The CHU9D was chosen for inclusion as it is a validated preference-based measure (PBM) developed for children that allows QALYs to be calculated for use in CUA. It is brief and has been widely used in a clinical trial setting (Thorrington and Eames, 2015) and in the UK (Wolstenholme et al., 2018). There are few PBMs available for use in children, as mentioned in Section 2.9.3. The EQ-5D-Y was considered as an alternative, however there is no child-specific value set available and the adult EQ-5D value set is therefore the only option to generate utilities. There is evidence that the different wording and perspective (child or adult) impacts valuations, therefore this approach is not recommended (Kreimeier and Greiner, 2019; Kreimeier et al., 2018).

The Daily Mile is also considered a wellbeing initiative (The Daily Mile Foundation, 2018b), therefore this was also measured in the trial. In comparison to the CHU9D, the MDI is a longer 77-item questionnaire measuring social and emotional health and well-being in middle childhood (ages 6-12) (Schonert-Reichl et al., 2012). It is described in detail in Section 6.2.2.1.1. It has not been previously used in the UK but has been implemented in large population studies in Canada, where it was developed, and in Australia. In addition, the MDI has only been used for population monitoring and not in a clinical trial setting that is designed to measure change over time in response to an intervention. There are several children's wellbeing questionnaires available, although evidence to support their use is variable (e.g. psychometric performance and responsiveness over time and use in pre-adolescents). A more comprehensive discussion of existing wellbeing outcomes is provided in Section 6.1, but in essence there is currently a lack of suitable measures for use in clinical trials in primary school-aged children and no preference based measures of wellbeing for children. The Birmingham Daily Mile trial therefore provided an opportunity to explore the feasibility of the MDI in a UK setting for the first time and understand the impact of the Daily Mile upon child wellbeing. As the MDI is not preference-based it cannot currently be used in a CUA. This limits its use to a CEA, if an appropriate score can be calculated. Indeed, a summary score of wellbeing was developed for this purpose. This is described in Chapter 6, but was the sum of items that comprise the MDI wellbeing index. This also enabled a psychometric analysis of the MDI, in terms of improvement to wellbeing in comparison to utility-based quality of life measured using the validated CHU9D. This was as a first-step towards considering wellbeing as an outcome within an economic evaluation framework.

Outcomes and data collection timepoints are reported in Table 6. Copies of the MDI and CHU9D are presented in Appendix C and Appendix D, respectively.

This chapter describes results in relation to the primary outcome (BMIz). BMIz is calculated using LMSgrowth software (Cole and Pan, 2002) and based upon age and gender specific British 1990 growth reference data (Cole et al., 1995). Detailed descriptions of the other outcomes used in this thesis are provided in Chapter 6 and Chapter 7.

Table 6 Outcome data collection

	Timepoints		
Outcome	Baseline	4-months	12-months
BMIz	x	x	x
Body-fat	x	x	x
Fitness	x	x	x
CHU9D	x		x
MDI	x		x
Academic attainment	x		x

4.2.2.6 Study implementation

The intervention was delivered over 12 months (April 2017 – March 2018). Baseline measures were collected in February/March 2017 and schools were randomised in April 2017. First follow-up data were collected for selected outcomes at 4 months (July 2017), prior to the 6-week summer vacation. At this first follow up point, outcomes collected were limited to items needed for the BMIz, bodyfat percentage and fitness (linear track test) to limit the data collection burden for schools. Final follow-up data were collected for all outcomes, 12 months after the start of the intervention (March 2018).

4.2.2.7 Analysis

Analysis was by a modified intention to treat and used a complete case approach. Analysis of the primary outcome used a mixed-effect linear regression model with 12-month BMIz score as the dependent variable, with trial arm and baseline BMIz score as independent variables. School (cluster) was included as a random effect. In partially adjusted models, the analysis was adjusted for covariates used in the constrained randomisation (baseline school mean BMIz, percentage of children eligible for free school meals and school size). In the further adjusted models, additional adjustments were made for pre-specified pupil level covariates (age, gender, and ethnicity) and school level deprivation (UK Index of Multiple Deprivation (IMD) score based on school postcode). Significance was considered at the 5% level.

4.2.3 Results

A summary of the trial results is provided here, with the full results recently published (Breheny et al., 2020).

Participant flow through the study is presented in Figure 2. Of 108 eligible schools invited to participate, 40 schools consented to participate in the study. Parental consent for measurements was obtained from 2280 children at baseline. Three schools dropped out over the course of the study.

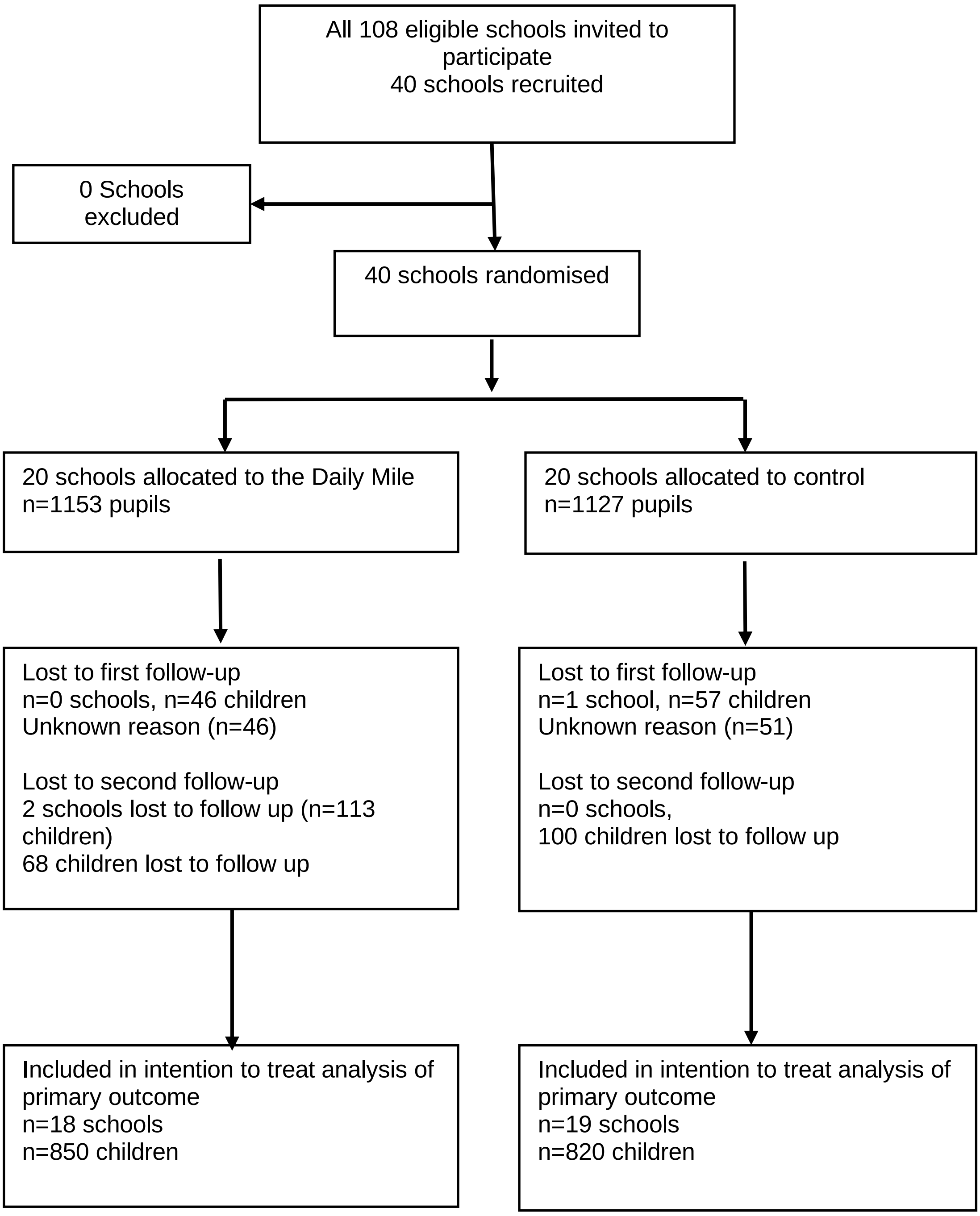

Figure 2 Participant flow

At 12-months (after start of intervention), an increase in mean BMIz from baseline was observed in both arms, and whilst the mean difference in BMIz score indicates a smaller increase in the intervention compared with the control arm, this was not statistically or clinically significant, MD = -0.036, 95% CI -0.085 to 0.013, p=0.146 (results adjusted for school size, % free school meals, school BMIz, school baseline outcome, participant baseline outcome). In pre-specified subgroup analysis, there was a significant interaction by gender with a modest and statistically significant intervention effect on BMIz for girls at 12 months in the partially-adjusted model (MD = -0.097, 95% CI -0.156 to -0.037, p=0.001).

4.2.4 Discussion

Analysis of the clinical results found that the Daily Mile did not have a clinically important impact on BMIz overall. There was however evidence of a modest intervention effect on BMIz for girls. This suggests that delivered in isolation, the Daily Mile is not a solution to tackling childhood obesity but could contribute to a system-wide approach. A discussion of the Birmingham Daily Mile is presented in Chapter 7, which reports an economic evaluation of the Daily Mile.

A question, which also has implications for evaluation, is whether the Daily Mile is a complex intervention. Complex interventions pose a number of challenges for the evaluation of effectiveness and cost-effectiveness. The basic definition in the current, although soon to be updated MRC guidance (Craig et al., 2008), specifies that a complex intervention is one that contains several interacting components. Reflecting on this, it would appear that the Daily Mile is a simple intervention. In the context of the UK Government's childhood obesity policy (HM Government, 2018a), there is an implicit causal link between increased physical activity and the prevention of excess weight, which is a biologically plausible relationship. The expanded characteristics for consideration might suggest otherwise, however. These include multiple possible outcomes and adaptable interventions. The Daily Mile has several potential outcomes, highlighted in the anecdotal reports of its effectiveness. These include improved academic attainment and wellbeing. A link between running around a playground and these outcomes is less clear. This perhaps suggests a backstep is required to explore this, prior to claiming these relationships with the Daily Mile. In the case of the Birmingham Daily Mile, there was little harm in supplementing anthropological measures with these outcomes, although interpreting the results as an absence of effect is perhaps premature. Furthermore, its implementation is somewhat flexible. While it is acknowledged that there is no clear boundary between simple and complex interventions, a more stringent application of the MRC guidelines for systematically evaluating complex interventions might have been beneficial. As the Daily Mile was already an established intervention, the initial development and feasibility/pilot testing phases were redundant. The final 'Implementation' phase of evaluation is still relevant though. This includes dissemination and long term follow-up (Craig et al., 2008).

The Birmingham Daily Mile trial aimed to answer the policy question of whether the Daily Mile is an effective and cost-effective obesity prevention initiative. The primary outcome chosen was BMIz, which is an objective, practical measure (Adab et al., 2018b) and the most frequently used outcome in trials of obesity prevention interventions (Brown et al., 2019). This would enable the findings to be compared to existing evidence and the meaningfulness of the results interpreted in the context of established minimal clinically important differences. BMIz could also potentially indicate unintended consequences that could be explored in further research, such as an increase in weight due to compensatory sedentary behaviours.

BMIz is a logical primary outcome for a fully powered, RCT of a single, simple intervention aiming to prevent obesity. Whether it is an appropriate outcome for a trial of a complex intervention is less certain. The proposed mechanism by which the Daily Mile influences obesity levels is by increasing physical activity, which should reduce the energy imbalance that causes excess weight. Measuring physical activity levels may have provided a better indication of how the Daily Mile is impacting on children's activity and behaviour change. Children may do less activity in the evening to compensate for their ‘Daily Mile’ for example. It could also contribute to the interpretation of the gender differences in effectiveness. Measuring physical activity could also indicate the fidelity of how the Daily Mile was implemented. Additional impacts that the Daily Mile could have influenced are improved mental health, and potentially spillover effects to other family members. These were not measured in the trial but could have provided a better indication of the impact of the Daily Mile.

If the evidence generated from the trial were to be used to inform resource allocation decisions between sectors (e.g. housing) or between different initiatives (e.g. smoking prevention), a CEA using BMIz would not be an informative analysis. Powering the trial and prioritising data collection based on BMIz instead of a generic HRQL outcome such as the CHU9D does not facilitate a robust CUA or decisions regarding allocative efficiency. While outcomes such as CHU9D and wellbeing (MDI) were collected and could be reported in a CCA, interpreting the meaningfulness of the results is difficult. For example, what a meaningful difference in wellbeing is is unknown. A CCA could still provide an informative overview of the costs and benefits of the Daily Mile for decision makers, yet this must be balanced against the resource implications of measuring a large number of outcomes.

4.2.5 Conclusion

This chapter has described a school-based public health intervention, and a robust evaluation of its effectiveness. The Daily Mile and the Birmingham Daily Mile trial will be used as a case study for the remainder of this thesis. The clinical results raise some interesting questions, such as whether the opportunity cost to boys is justified, given the clinical benefit observed in girls. These will be reflected upon in Chapter 7, where an economic evaluation of the Daily Mile is reported.

The following chapter reports a qualitative study with school staff which explores issues regarding opportunity costs and appropriate outcomes. The relevance of a difference in BMIz across a school year and BMIz as an outcome are unknown to a school population, and on what basis a school adopts an initiative like the Daily Mile is uncertain.

CHAPTER 5 PERCEPTIONS OF DECISION MAKING AND ECONOMIC EVALUATION AMONGST SCHOOL STAFF. A QUALITATIVE STUDY

5.1 Introduction

As discussed in Chapter 1, schools are a setting regularly utilised to deliver public health functions and interventions. These can be mandatory, such as the NCMP, or they could be optional, such as obesity prevention initiatives identified in the systematic review reported in Chapter 3. The review highlighted the availability of economic evidence, although whether this evidence is used at a school level is unclear. How these discretionary initiatives are prioritised is currently unknown, including if or what costs and outcomes would be considered by the individuals with decision making responsibilities. As reported in Chapter 2, cost and outcome data are critical components of economic evidence. Identifying the most appropriate inputs for the appropriate decision maker/makers are important contributions to generating useful economic evidence.

A recent qualitative study by Jessiman et al. (2019) sought to explore decision making regarding health promotion in academy trusts. The sample included only 'elite' participants, who were those with considerable power and influence over decision making, seeming to assume that other school staff have no decision-making role. Sections 1.4.9 and 2.10.1.3 highlight the Academy System in the UK. The Academies Act 2010 perhaps introduced some clarity surrounding leadership structures and what decision-making responsibilities schools have, yet the actual process of decision making 'on the ground' remains unclear and there is little formal examination of this in a local authority maintained school context and with individuals tasked with implementing interventions.

There is limited qualitative research exploring school staff's views on the expectations upon them to deliver public health interventions and how these staff decide on the priority that should be given to these health interventions in the context of the limited time and other resources available to staff, the educational objectives that they are judged upon, and the available evidence about the success of these public health interventions. As indicated in Section 3.5, studies have explored implementation (post-decision making) and sustainability of interventions (Darlington et al., 2018; Day et al., 2019), finding for example that funding was a moderating factor. It could therefore be assumed that costs are also a consideration pre-implementation. Indeed, a qualitative study in the UK attempted to understand current physical activity provision in schools and identify barriers to implementation of physical activity schemes, with the aim of informing policy makers about what to commission (Arnold et al., 2016). Interviews with headteachers revealed the most frequently reported barriers to be a lack of time and funding constraints for specialist provision. Academic attainment was seen as more important than health outcomes, with priority particularly directed towards literacy and numeracy. A qualitative investigation into the acceptability of health promotion programmes in primary schools in Ireland interviewed 31 teachers (Bennett et al., 2016). Staff recognised the importance of health promotion and agreed that schools are a logical setting for this, however concerns included unrealistic expectations with respect to including health promotion alongside a demanding curriculum. A recent study explored the barriers and facilitators to delivering the Daily Mile (Malden and Doi, 2019). Interviews with 12 teachers found that time required was challenging, although they were positive about the health benefits. These studies highlight the challenges of implementing physical activity interventions in a school context, particularly in relation to schools' priorities. They do not however explore the decision-making process involved, the outcomes that schools desire beyond academic attainment or the costs involved in achieving these outcomes. Alternative outcomes may contribute to more persuasive evidence if they align with the perspectives of school staff. In addition, if economic evaluations encompass the true cost implications of delivery in the school setting the results could be more informative. Informative to whom is still an outstanding question, however. Chapter 3 highlighted the limited costs included in economic evaluations of school-based interventions, yet the suitability or appropriateness of these from the view of the school has not been explored.

Whilst schools may provide a captive audience to deliver public health interventions, schools' objectives and funding arrangements may not align with those typically considered in economic evaluation. This qualitative study aimed to explore what costs and outcomes are important to the providers and decision makers in this context, how decisions are made and how interventions are prioritised in this setting. It was intended that the findings would help inform economic evaluations in a school setting and develop an understanding of a schools' perspective in economic evaluation.

5.2 **Methods**

The study reported was conducted within The Birmingham Daily Mile cluster RCT (Breheny et al., 2018) and as detailed in Chapter 4. Approval for this qualitative element of the study was granted by the University of Birmingham Research Ethics Committee (ERN_17-0171). Approval documentation is provided in Appendix E.

5.2.1 *Sample and Recruitment*

Semi-structured interviews were conducted with teachers and individuals involved in school decision making at primary schools participating in the Birmingham Daily Mile trial. The inclusion criteria were intentionally broad in order to elicit views from those responsible for implementing interventions and individuals more distanced from day-to-day implementation but with relevant roles (e.g. curriculum oversight, leadership, pastoral input).

School staff were the target sample, but the involvement of other stakeholders was considered during the design of the study. Views regarding broader inclusion criteria, such as governors, was sought from an advisor working in the education sector, but their inclusion was deemed unnecessary and they believed that governors would be challenging to recruit. A separate qualitative study was conducted alongside the work reported in this thesis (Frew and Breheny, 2019a). This involved interviews with individuals working in a local authority, and no decision-making responsibility influence over schools was identified. One participant did note that economic evidence presented to schools should be targeted and persuasive, given their bottom line is academic attainment. What constitutes persuasive and targeted evidence is however unknown. It was believed that parents and children could have provided a valuable contribution to the study, although this was beyond the resources available. The different methodology required for children would constitute an additional study, alongside stringent requirements regarding ethics and safeguarding.

Only staff in schools allocated to the intervention arm were eligible to participate. This was due to the focus on the costs and outcomes associated with providing a public health intervention within primary school environment. A purposive sampling approach for maximum variation was planned. This is a method of sampling where participants are selected by the researcher based on a variety of criteria (Jupp, 2006; Sandelowski, 1995). Criteria included school and individual characteristics. School characteristics were OFSTED rating, number of pupils and eligibility for free school meals. Participant characteristics were whether they held a leadership or teaching role. It was intended that participants would be recruited until saturation of themes around priorities, outcomes and costs of The Daily Mile had been achieved. However, due to anticipated challenges of recruiting participants, all 20 schools in the intervention arm were invited to participate.

The staff member coordinating the intervention at each school was first approached by email to participate. Follow-up emails and phone calls were made if no response was received. The email contained a participant information sheet (Appendix F) detailing the aim of the qualitative study and what it would involve. A snowballing approach was then used, whereby participants were asked to recommend other individuals within, or working with, the school that they believed would provide a helpful contribution to the study.

5.2.2 Data collection

Semi-structured interviews were conducted face-to-face or telephone. Face-to-face interviews were conducted at the participants' schools and telephone interviews were used if face-to-face interviews were not possible. Locations included offices, staff rooms and available classrooms. All participants provided informed consent (Appendix G) and completed a short participant characteristics form (Appendix H). The interview broadly followed a topic guide that included questions regarding implementation of The Daily Mile, any costs incurred and the prioritisation of health and wellbeing interventions in schools (Appendix I). The topic guide was reviewed and refined by members of the project team. As not all participants were aware of the outcome measures being collected, visual aids were used for familiarisation, regardless of previous experience using them. These included the domains captured by the quality of life and wellbeing measures. These were described verbally during telephone interviews. The interviews were audio-recorded with the participants' permission.

5.2.3 Data analysis

Audio data were transcribed verbatim and then checked for accuracy. NVivo 11 (QSR International, 2017) was used to organise and manage the data. A constant comparative method was used to analyse the data. Transcripts were reviewed and coded, with codes reflecting interpretations of the data's meaning. New data were compared, initially to previous data and, as analysis progressed, to the properties of emerging themes. Themes are higher order concepts that reflect common meaning of characteristics in the data (Corbin and Strauss, 2015; Glaser and Strauss, 2004). A second individual reviewed a sample of transcripts to compare themes. Three detailed analytic accounts were developed (Coast and Jackson, 2017). These described the data and made connections between themes and subthemes. The accounts reported interviews grouped into leadership staff (staff who were on the school senior leadership team), teachers responsible for their own class and other staff (e.g. staff with pastoral roles and teachers delivering physical education only). These descriptive accounts were subsequently drawn together to generate one account synthesising findings from all interviews. Data are presented as a narrative summary of relevant findings, supplemented by illustrative quotes. Quotes were selected to be representative of participant views and where possible, to provide a succinct conceptualisation of the theme. Quotes are presented verbatim with the use of ellipses to represent missing text; phrases such as 'you know' , 'umm', 'err', 'like', or repeats of words that do not add to meaning are excluded without use of ellipses. Participants are referred to using an identifying (ID) number.

5.3 **Results**

5.3.1 Participants and their roles

The following results report the findings from 14 interviews with staff holding a number of different roles and responsibilities (Table 7). Participants comprised leadership staff (N=6), teachers (N=4) and other staff (including safeguarding leads and PE specialists) (N=4). Some leadership staff also had part-time teaching roles. School work experience ranged from 3 to 33 years, with leadership staff having most experience. Interviews were conducted in seven schools, with the number of participants from each school ranging from one to four.

Table 7 Participants' roles and interview settings

ID number	School ID	Position	Year teaches	Setting and notes
1	1	PE teacher only	All	Unoccupied office
2	2	Executive headteacher	N/A	Office shared with ID3. Conducted with ID3.
3	2	Head of school	N/A	Office shared with ID2. Conducted with ID2
4	2	Teacher	6	Unoccupied classroom, with outside noise audible
5	2	Teacher	4	Unoccupied classroom, with outside noise audible
6	3	Pastoral lead	N/A	Personal office
7	3	Assistant head/teacher	N/A (5 previous year)	ID6 office, with ID6 remaining in the room. ID6 was unobtrusive and made occasional remarks. Recruited at short notice so was unprepared and apprehensive about audio recording
8	3	Teacher	4	ID6 office. Recruited at short notice so was unprepared and apprehensive about audio recording
9	4	Deputy head	5	Personal office with several interruptions requiring recording to be paused
10	5	Assistant head/teacher	6	Empty classroom after school. Cleaners audible
11	6	SENCO and safeguarding lead.	N/A	Telephone interview
12	7	Headteacher	N/A	Personal office
13	7	PE teacher only	All	Staffroom, with occasional interruptions
14	6	Teacher and PE coordinator	3 (5 previous year)	Telephone interview

5.3.2 Interview process

Interviews were conducted between November 2016 and March 2017. Mean interview duration was 28 minutes and the majority of interviews were conducted face to face (N=12). Table 7 reports the location of each interview and notable details that might have impacted the discussion. ID2 and ID3 were interviewed together in their shared office. ID6 remained in their office while ID7 was interviewed, interjecting occasionally in agreement or to assert that they had raised a similar issue in the earlier discussion. Outside noise was audible during several interviews, including cleaners and children passing by. Some interviews were interrupted, but in only one case was the interview paused.

There was a range of enthusiasm for participating in the interview and the Daily Mile trial. ID7 was currently Assistant Head and a fitness instructor. Their enthusiasm for physical activity and its influence on the decision to participate in the trial emerged frequently. As a PE specialist, ID13 was very focused on sports and spoke at length about what extra-curricular sports they offer. ID6 was very enthusiastic about The Daily Mile and was very keen to chat. In contrast, ID8 appeared quite 'frosty' and was not enthusiastic about the Daily Mile. ID11 was personally very keen on The Daily Mile as a concept, although reported facing resistance from parents and teachers.

The following results describe the main themes that arose during the interviews. Figure 3 provides a broad outline of these themes. Shaded boxes relate to the topics explored and included in the discussion guide. The remaining shapes reflect the themes that emerged out of the data. Pentagons include more detailed findings, with some examples provided.

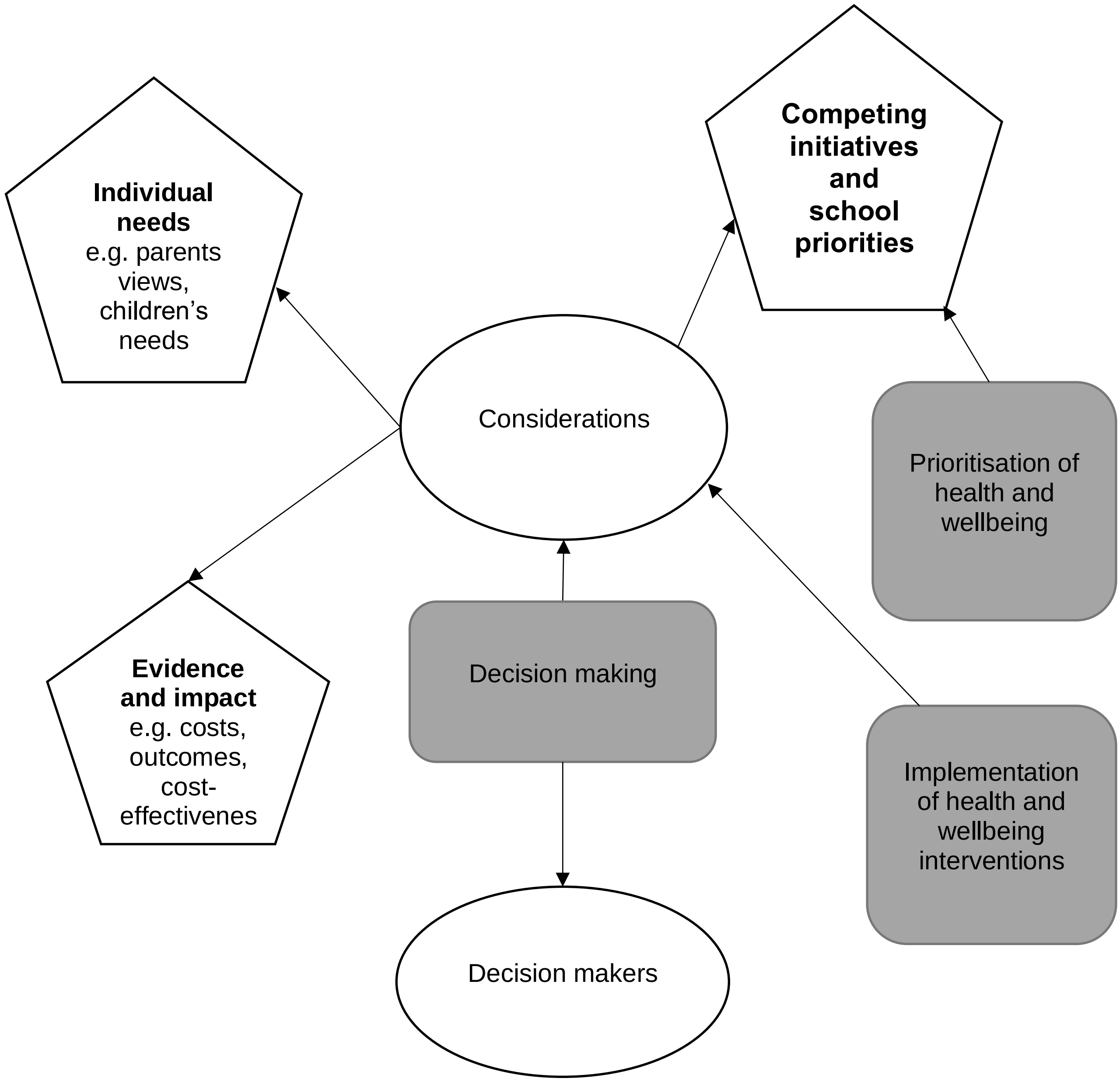

Figure 3 Outline of themes

5.3.3 Themes

5.3.3.1 The decision maker

The participants were asked how it was decided that their school would participate in the trial and who made decisions about implementing similar initiatives. The majority of participants (including those with leadership roles) believed it was the role of the headteacher to make these decisions. Some staff with lesser leadership responsibilities (e.g. deputy headteachers) reported having some authority, although with headteachers' oversight.

I'm the headteacher, so normally if there's an initiative or anything like that if I'm enthusiastic about it, it normally gets done [ID12, Leader, School 7]

Me, obviously anything like this, I'd run by my head as well. They trust me with any of the initiatives I've done so far, so, yes, it's down to me, for good or bad. [ID9, Leader, School 4]

Teachers in Schools 2 and 6 spoke about participation in hypothetical health and wellbeing initiatives being whole school decisions, with commitment required from all teachers for it to go ahead. Speaking about the trial however, teachers at School 2 were clear that the definitive decision to take part in the trial was made by the headteacher and they complied without argument. Despite most non-leadership staff discussing being isolated from decision making, headteachers reported welcoming their ideas and considering potential opportunities for new initiatives. Some staff supported this, mentioning that they could highlight opportunities for leadership staffs' consideration.

… we've been told, from the top, that this is what we're going to do and so everyone's doing it [ID4, Teacher, School 2]

At the end of the day the head needs to have the final say on whether we do or whether we don't. We put the suggestions to him, and if I get a green light then fine [ID13, PE teacher, School 7]

The role of school governors in decision making was raised in the discussion. It appeared that they had little input or influence. At three schools (schools 2, 4 and 6), staff in leadership and pastoral roles thought governors were aware of health and wellbeing initiatives being implemented yet believed there was no requirement for their consultation or approval. Despite this, leadership staff believed governors were interested in their impact and wanted to see evidence to justify investments and demonstrate benefits.

> *So they're aware we're doing it but we didn't ask permission, no, but they would be interested in the impact of it, definitely. [ID2, Leader, School 2]*

> *I think any impact you can show. I'd quite like to go to governors and show them that this is what we're investing in, this is part of our curriculum and this is the impact it's had [ID9, Leader, School 4]*

5.3.3.2 Decision making considerations

What is taken into consideration when staff are assessing the possibility of introducing a new initiative was discussed. Tangible evidence was thought by leaders and teachers to aid decision making. Interviewees highlighted anticipated or desired benefits, both in relation to the Daily Mile and other health and wellbeing interventions. The importance of costs was discussed, and the costs of the Daily Mile investigated in more depth.

5.3.3.2.1 <u>Evidence and impact</u>

The majority of participants spoke about the need for evidence of effectiveness, and/or cost-effectiveness to support decision making. Teachers thought that leadership staff wanted to observe quantifiable improvements that could be attributed to new initiatives. Leadership staff also expressed a desire for school level evidence, but highlighted that broader uptake could be improved if larger studies demonstrated benefit. In the absence of evidence, participants at two schools discussed trying out initiatives and generating their own evidence. Their intentions included assessing impact and understanding process outcomes that could inform the refinement of initiatives.

> *So we trial it for six weeks and then we look at the end of the six weeks. Has it made the impact? Has it not? What is, what is it that has? Or if it hasn't, then we'll look at different ways to adjust it and change it so it does [ID7, Leader, School 3]*

> *I think they wanted to see the impact really. They want to see the improvement of the figures in a sense is what they're looking for [ID1, PE teacher, School 1]*

5.3.3.2.2 Cost-effectiveness

The possibility of using economic evaluation to inform school decision making was discussed with the participants. A number of participants recognised budget constraints, and some drew the link between using budgets and achieving benefit.

> *Obviously, we're quite tight on budget but again, it's all dependent on what the children need and if it's going to benefit them [ID7, Leader, School 3]*

> *If something comes in and it's going to cost too much, whether it be transport costs or whatever then, obviously it has to be looked at as to whether it does actually benefit the child [ID6, Pastoral role, School 3]*

In reference to CEA, several teaching staff were not concerned about the cost of initiatives, although they tended to see this as the responsibility of senior leadership.

> *I can imagine from senior leadership that that would be a factor if you could prove that this amount of money would have this effect… if they can see that it might appeal. If I'm honest, it's … seems a bit irrelevant to my classroom practice but that's because I'm not making the decisions [ID4, Teacher, School 2]*

> *I'm always conscious of things costing money and how much they cost, that's when I hold my hands up and say 'well look, you're in charge boss!' [ID13, PE teacher, School 7]*

Several leadership staff expressed a desire for evidence of cost-effectiveness, essentially describing what an economic evaluation could demonstrate. They reflected on the need to spend public funds wisely and show that initiatives have observable results to justify implementation.

> *I think you'd have to really prove its effectiveness and the impact of it because it's public funds and it's like anything we do; we have to measure impact, and I know this is a time thing, so I suppose, in a way, it is a cost, isn't it? [ID2, Leader, School 2]*

> *So, for example, if we were trying to reduce BMI, and we wanted to pay for somebody to come in and do a sporting activity over a period of time, we would want to see a measurable decrease [ID10, Leader, School 5]*

The current climate, you've got to make sure that we are using all the money as effectively as possible and we have to measure it against the impact that it's going to have [ID12, Leader, School 7]

5.3.3.3 Costs of public health interventions

The costs directly incurred when implementing interventions were discussed with participants, mostly in the context of the Daily Mile. When alternative initiatives were mentioned, their funding sources were highlighted, such as sports premium and parent contributions.

5.3.3.3.1 Costs of Daily Mile

When asked what the Daily Mile had cost the school, most staff initially asserted that it had not cost the school anything. Many cited running as a free or cost-neutral activity. Teachers tended to report that it was an additional activity, or the school day was adjusted to accommodate it, therefore the children did not miss out on anything. It was also suggested that the Daily Mile has cost savings, as opposed to imposing costs on a school. This was attributed to a decrease in the use of supply teachers due to reduced teacher absence.

I don't think so. Running's great, isn't it? It's a free kind of form of exercise. They don't need any equipment [ID4, Teacher, School 2].

In fact, it's probably had benefits because I think it's upped the level of fitness of the staff and I actually think it's been cost positive because I think the amount of supply that we've had to have because of staff being off has dropped slightly over the last year. [ID12, Leader, School 7]

5.3.3.3.1.1 Material costs

Some participants reported costs they perceived as relatively minor. These included the costs of displays erected to track progress and integrate The Daily Mile with other lessons, generating certificates ,entering data, the purchase of pedometers for children to record their steps and an increase in first aid supplies used due to the collisions between children and accidents that occurred when the Daily Mile was first initiated. Not all schools incurred all costs.

> *I mean possibly a very small cost of some wrapping paper or stuff to make it look pretty… Just things we've got at school really [ID1, PE teacher, School 1]*

> *We've used it a little bit in maths to look at the range of steps. We've had pedometers for a little while, although they didn't last that long [ID9, Leader, School 4]*

> *The First Aid Box in the outside classroom. It's been revamped [ID3, Leader, School 2]*

Whilst these are relatively small costs, one school (School 7) received £10,000 of Lottery funding for a running track to be installed. A purpose-built track was something that several staff aspired to, although not all staff agreed this should have been the priority for additional spending.

> *So we went for a £10,000 charity lottery fund and had a track put round, a kind of rubberised track [ID12, Leader, School 7]*

> *We did look at having a specific track done but really our playground needs replacing before we think about that and that*

is probably around £70,000 we haven't got. [ID9, Leader, School 4]

5.3.3.3.1.2 Opportunity cost

Although some teachers spoke about there being no cost attributable to the Daily Mile, several, predominantly leadership staff, recognised that there was an opportunity cost in terms of time taken out of academic lessons. They did not seem to conceptualise that a monetary value could be ascribed to this time. At one school, the opportunity cost of the Daily Mile was perceived as missing attendance at the school assembly. Whilst not an academic lesson, the latter was perceived to have an important pastoral role in providing recognition and encouragement to the children.

> *We haven't had to spend anything on this. So it's 10-15 minutes a day; it hasn't cost us anything, except time [ID10, Leader, School 5]*

> *No well it, the cost as in monetary it doesn't cost, the cost as in time out of class [ID8, Teacher, School 3*

> *They were missing assembly, so they were missing out on that but we did it as much as we could… It's a celebration assembly where there's a writer of the week and a mathematician of the week and there's attendance certificates, it's a big celebration of their work and what they do, so it's nice, they really can't miss that because it's boosting them and if a whole class has done well, they all stand up and they get cheered and for some kids, that's a big thing because sometimes they don't get a lot of praise at home, so it's missing out on that [ID6, Pastoral role, School 3]*

5.3.3.3.1.3 Opportunity cost of the Daily Mile, in practice

The Daily Mile is a pragmatic initiative, although it is intended to be conducted daily for 15 minutes. The Daily Mile's implementation fidelity, and actual opportunity cost was explored during the discussions with staff. There were contrasting reports from those with different responsibilities. Those who were not responsible for teaching their own class generally believed that the Daily Mile was not completed daily; they had ascertained this through speaking with children and/or teaching staff. While some teachers acknowledged that they did not run the sessions on a daily basis for reasons such as competing curriculum activities or lack of time, others were confident that it was implemented as intended.

> *Occasionally I've had to remind staff that they have to do it on a daily basis because I go round and talk to the children about it and that's the best way of finding out if they are doing it and occasionally it has slipped in certain year groups… generally speaking they do it on a regular basis [ID9, Leader, School 4]*

> *Every day well um the day we don't do it is the day they have PE because I kind of think well they're having PE anyway so they should be doing the fitness. I can't say hand on heart we have done it every day because, so we have workshops…….. so there are curriculum things that stop us doing it but I would say 95% of the time we're doing it [ID8, Teacher, School 3]*

> *I mean it's a rare day that we don't manage it [ID4, Teacher, School 2]*

Staff were asked how long the Daily Mile typically took to conduct on a day to day basis, with inconsistent implementation and policies reported. Implementation was varied in regard to whether the time spent getting to the playground was included within the 15 minutes and whether the full mile had to be completed before returning to the classroom. Some staff reported reducing the time outside to account for the time to change and return to the classroom, with the classroom's proximity to the playground reported as a consideration.

> *Some of them finished quicker than others, so then they would come in, kind of staggered really [ID7, Leader, School 3].*

> *We start and then I tend to make it, 12 minutes-ish and then we'll come in, … overall it's 15 minutes but a few minutes of that is coming back into the building and taking coats off and having a drink after break and all that faffing around that comes with being a 9 year old [ID5, Teacher, School 2]*

> *So it's make sure that they're out there for at least 15 minutes and to get as close to a mile as possible [ID1, PE teacher, School 1]*

Whether the children ran or walked the Daily Mile was discussed. It seemed that in general they were encouraged to run, although teachers experienced resistance from some children for reasons including a lack of motivation. Maintaining motivation was raised as a key requirement for long-term success of the initiative. It was highlighted that children saw staff as role models and would imitate them if they were seen to be walking.

> *Some of the children are really motivated to set themselves targets of – 'I'm going to run for three laps' or 'Next week, I'm going to run for four laps and walk for one' and some children are*

– 'It's not my favourite thing' and then it's quite hard to motivate them and make them run (ID4, Teacher, School 2)

I keep like pushing them to try and run but they just walk around and I'm encouraging them to run a little bit but– because I think, if they see adults running, they'll want to run with them. If they see them walking, they would rather walk (ID7 Teacher, School 3)

One feature of the Daily Mile that intends to reduce the time burden and improve ease of implementation is that children do not change their shoes or clothes. In this respect, it seemed there was no consistent policy across schools. Where children did change their shoes, it was most frequently justified by girls' footwear being inappropriate for running. Few staff reported children changing their clothes, although understandably coats were sometimes worn in winter.

No, they'd just go out as they are, so some would argue footwear might be an issue erm but when they go out in the playground at playtime they don't change their footwear for playtime and lots of them are just running around at playtime anyway (ID13, PE Teacher, School 7).

I can see them running round in their shoes, but especially girls with their little ballet slipper things that they have, it's not ideal, so we do encourage them to change and that can take time… so it all impacts on the learning and what's going on that day (ID6, Pastoral lead, School 3).

5.3.3.3.1.4 Staff costs

Aside from the regular supervision of the Daily Mile, staff time was also consumed doing tasks related to the initial set-up and ongoing facilitation. These tasks included conducting risk assessments, inputting data and creating classroom displays. Risk assessments were not seen as onerous, although they were not mentioned by staff at all schools. Inputting data tended to be conducted by supernumerary staff such as trainees.

> *I carried out risk assessment… we just used the same risk assessment we would for a P.E. lesson…so, half an hour of my time, and probably 15 minutes of P.E coach time. [ID9, Leader, School 4]*

Staff were asked who usually supervised the Daily Mile. It appeared that teachers were rarely alone. Typically, teaching assistants (TAs) were also in attendance, with other teachers sometimes coincidentally outside with their own classes. Policies regarding a minimum of two adults supervising a class for safety were mentioned, as were limits on student teachers supervising independently. Regulations were not consistently reported within schools, however.

> *No, it's me. I mean I've got student teachers at the moment who come out with us but not on their own [ID4, Teacher, School 2]*

> *No there's two of us that go out so sometimes the classes go together if it suits. Sometimes we go out separately and we have two adults so if something happens to a child that adult can deal with it while the other adult is still out watching them [ID8, Teacher, School 3]*

5.3.3.4 Outcomes

The outcomes that staff would hope to observe as a result of implementing new health and wellbeing interventions were discussed, as were the anticipated benefits of the Daily Mile specifically. The themes largely related to inspiring enduring behaviour change in order to improve long-term health outcomes and wellbeing benefits.

5.3.3.4.1 *Long-term health and weight*

Many staff were concerned about the potential long-term impact that obesity and inactivity might have on the children's health. Several participants referred to issues with obesity in their school, based on observations of children and routinely collected data. Some also reflected on their own weight difficulties. Leaders spoke about their school's NCMP results, which in some cases motivated participation in the trial. One teacher was aware that their school compared poorly to other schools in the city.

> *The way our ethos is, it's about getting the children healthy, out of this cycle of unhealthiness and obesity which, looking at some of the children already, it's a major concern what they're going to be like in four or five years' time [ID9, Leader, School 4]*

> *I'm overweight now, but was never overweight as a child and I'm concerned when I see how overweight some children are as children, as to what they're long-term outcomes are [ID2, Leader, School 2]*

> *You get the school's data for obesity and the BMI data, well we're really low. We're one of the worst schools in Birmingham [ID10, Leader, School 5]*

5.3.3.4.2 <u>Enthusiasm and enjoyment</u>

An outcome mentioned frequently and that aligned with staffs' interests in children's long-term health was enjoyment and generating a legacy of participation in physical activity. While increasing physical activity was important, staff of all roles spoke about how enthusiasm and enjoyment were also required. Perhaps coincidentally, these were reported as benefits of the Daily Mile. Several staff spoke about how children were now doing extra activities such as Parkruns or had a new-found interest in running. Staff believed that not developing good habits in childhood would increase the likelihood of being inactive in adulthood and short-term increases in physical activity were seen as only an intermediate outcome. Staff with strong interests in physical activity appeared to believe that exposure to different opportunities and allowing the children to find their own strengths and passions was critical for developing enduring healthy behaviours.

> *I think it's more important that children enjoy it than feel they have to do it…. I think, maybe when the people are making policies about improving children's fitness levels, we should be looking at children's enjoyment [ID10, Leader, School 5]*

> *I think you've got to bring in some enjoyment levels of children because to lead an active lifestyle or healthy lifestyle they've got to enjoy some form of exercise. If they don't enjoy it they're not going to do it when they're older [ID1, PE teacher, School 1]*

> *I mean I'm guessing when they're running a mile in school the key is to try and increase their fitness, but really the lasting benefit is only really if by doing that they're actually more fit outside of school because that's what the game is isn't it? That*

when they leave education they're still going to the gym, playing football, running, playing tennis. That's the end game isn't it so that's what I think would be a good measure [ID8, Teacher, School 3]

5.3.3.4.3 *Knowledge*

Staff spoke about the school's role in enabling children to understand the rationale for physical activity and the knowledge of what a healthy lifestyle is. Some leadership staff broadened the scope to include educating parents too. They acknowledged that educational interventions can have limited impact due to a lack of control over children's lifestyle beyond the school setting, for example children's diet. Some highlighted a need for parental involvement in interventions. The Daily Mile was praised as it was fully overseen by the school, although it was also criticised for not having an educational component explaining the underlying purpose.

It's just getting them exposed to it and parents in to actually see what is healthy for their children and what's good for them [ID7, Leader, School 3]

That's a school issue that, the teaching of it, we need to make sure they need to understand what they put in their body results in how they feel and what they get out [ID9, Leader, School 4]

If you're doing something regularly like the Daily Mile, it will have a direct impact much more effectively than preaching about what they might eat because I don't think the children have control of that [ID12, Leader, School 7]

5.3.3.4.4 Physical activity

A recurring theme in the interviews was that staff were concerned about children's lack of physical activity and sedentary lifestyles. Staff at Schools 1 and 3 criticised the current PE curriculum. In the past PE involved vigorous activity, whereas now the focus is on learning skills. In contrast, at School 4 it was believed that the provision of two PE lessons a week was inadequate, and could not compensate for children's inactive lifestyles. Another contributor to inactivity was raised by pastoral staff. They expressed concern that children no longer walked to school, despite living nearly. An issue cited by the majority of participants was a shift in the culture of leisure activities, from sports to inactive computerised hobbies such as games consoles and tablets. The importance of participation in sport for physical and emotional wellbeing was also highlighted.

> *The problem is I see that the PE they're doing at the moment is a lot of co-ordination, balancing … Their PE is more sedentary [ID8, Teacher, School 3]*

> *They then get picked up at the school gate in a car. So they don't actually have any physical activity in their normal day [ID11, Pastoral role, School 6]*

> *I can remember at primary age playing football and athletics but nowadays it's not happening so much because they're on their iPad's or they're in their bedroom on their X Box's and things like that. A lot of the children don't have that side to them and I think it's not only good for their physical wellbeing but I think it's their character building and their emotional wellbeing [ID9, Leader, School 4]*

5.3.3.4.5 Confidence and self-esteem

Confidence and self-esteem were mentioned as both being observed benefits of the Daily Mile and outcomes that participants would hope to influence by implementing health and wellbeing interventions. Being able to observe progress reportedly boosted self-esteem and developed children's confidence in running and other activities previously perceived as unachievable. Confidence was also seen as being intertwined with wellbeing, which might increase engagement with sport.

> *I have seen improvements in confidence. Some of the children who have now taken part in running things previously always said 'I can't do that' [ID13, PE teacher, School 7]*

> *In terms of wellbeing, it will be to look at that confidence and whether or not they might then start to engage in more sport [ID7, Leader, School 3]*

5.3.3.4.6 Wellbeing and mental health

Improving wellbeing and mental health were reported as outcomes that staff of all roles were interested in. Staff at School 6 spoke about how improving wellbeing was a priority for their school, alongside improving mental and physical health. Teachers believed that when promoting implementation of new initiatives, the mental health and academic benefits should be underlined rather than physical outcomes. There was some concern from leaders about how wellbeing as a concept could be measured and quantified in a research setting. Leadership staff tended to be unfamiliar with the questionnaires used in the study however as they did not facilitate data collection within a class. Confounding influences affecting wellbeing were mentioned as potentially problematic and this was also mentioned in regard to academic attainment, with potential contamination from other initiatives and outside influences.

> *It would be difficult to say about wellbeing, because of how measurable that is. I don't know how you'd measure the wellbeing before and after but I know you've done the questionnaires though [ID12, Leader, School 7]*

> *We were selling it on the benefits of physical activity and long-term health benefits but actually, if it was packaged up as the mental health, and the readiness to work and academic outcomes I think it would be easier to convince staff that this was a worthwhile thing to do [ID11, Pastoral role, School 6]*

Improving staffs' wellbeing alongside children's through participation in health and wellbeing initiatives was seen as desirable. 'Spillover' staff wellbeing benefits were reported as a result of the Daily Mile for example. Wellbeing and mental health were also believed to resonate with the wider population. Participants were uncertain whether children could grasp wellbeing as an outcome and economic outcomes for example were seen as inaccessible to both children and staff however. No participants reported children having a decision-making role, although their cooperation might be enhanced if objectives are pertinent.

> *I think the economic one is the most difficult one. I think the wellbeing one would make sense to people [ID2, Leader, School 2]*

I think qualitatively, it's having an, an impact on their wellbeing… I would hope for the wellbeing to have improved but again, it's, so many variables [ID2, Leader, School 2]

5.3.3.4.7 <u>*Academic attainment, health and wellbeing*</u>

Academic attainment was highlighted by many as the school's main priority, although several participants noted that good health and wellbeing were prerequisites for academic attainment. Participants from School 7 in particular both spoke about their interdependence. Following an explanation of how cost-effectiveness could be reported to schools, the cost of an improvement in health or academic attainment was seen as potentially useful. Doubts were raised about how the direct impact on academic attainment could be inferred however, due to the vast amount of initiatives schools undertake.

So I mean obviously, I have to decide whether I feel that the intervention is going to be of any benefit to the children, mainly academically, I mean, because I think that's our reason to be here but equally for the wellbeing as well and I think the two go hand in hand (ID12, Leader, School 7)

At the end of the day, they can be as bright as anything but if they can't get off their sofa because they're too big then they've got real issues (ID9, Leader, School 4)

I think they're quite keen to focus on the physical and mental health, for its own sake but also for its effect on children learning (ID4, Teacher, School 2)

5.3.4 Other considerations

5.3.4.1 Parents

A number of participants mentioned parents as being a consideration, or hindrance to providing public health interventions such as the Daily Mile. One striking issue was parents' concerns that children shouldn't be forced to run for 15 minutes or that they were doing too much exercise. Leadership staff seemed to dismiss these concerns when challenged by parents, asserting that running was part of the curriculum for example. Some teachers believed it was not an extraordinary departure from normal playtime although others appeared more sensitive to parents' views, referring to wanting to have them on board and not taking young children out in the rain.

> *For example, one parent complained because she felt her son was doing too much exercise, uh, because he already cycled to school. So, we pointed out that it is part of our curriculum and he had to do it [ID9, Leader, School 4]*

> *So we could agree, because you want the parents on board, you don't want the parents saying I don't want my child doing this and you are thinking, well they ran around at lunchtime so why, what's the problem with them doing this [ID13, PE teacher, School 7]*

5.3.4.2 Children's gender and age

Children's gender was reported as a factor impacting the fidelity of the Daily Mile and as a barrier to successful implementation of physical activity interventions in general. The reluctance of girls to participate was highlighted by several staff. Reasons seemed to be either an aversion to physical activity, a preference for socialising, or it not being considered 'cool'. Older girls were reportedly most resistant to running, instead walking and chatting. Many participants thought older children, not just older girls showed reluctance to run in comparison to the younger ones. Reception and year 1 for example were reportedly enthusiastic, with engagement declining on a continuum until year 6. More generally, several staff spoke about how initiatives can be more suited to particular age groups, which would be considered when making decisions. Young children for example are physically disadvantaged when running long distances and lack the experience to pace themselves. This did not limit their participation in the Daily Mile however.

> *Quite a lot of the girls, in particular in year 6, don't enjoy doing any exercise, and that's one of the battles I've had really is get the girls to take part more. They will say, 'I don't want to do it' and then they'll talk to their friends, and they're linking arms, and having a little chat [ID10, Leader, School 5]*

> *Further down the school is definitely more enthusiastic about it. I've watched lower year groups and year two and three will run most of it. When you get up to year six they're almost really reluctant to do it… I'd say there's a good body of probably a third that think it's alright just to walk around and have a chat (ID9, Leader, School 4).*

> *Some things are more suitable for younger children, some things are more suitable for older ones like Bikeability That's not necessarily suitable for nursery, so it depends on the activity(ID6, Pastoral lead, School 3)*

5.3.5 Prioritisation of Health and Wellbeing

The challenge of introducing new health and wellbeing initiatives was discussed, predominantly in relation to the Daily Mile. Teachers were largely negative, citing an already crowded curriculum restricting the opportunity for additional activities. Many did appreciate their value, with some suggesting it is achievable if staff are passionate about it. In contrast, leadership staff acknowledged the demands of the school day but believed that some initiative and flexibility would enable the inclusion of short health and wellbeing initiatives feasible.

> *I really like the idea of it. I find it quite difficult to fit it in. There's a lot of pressure to being in year 6; a lot of pressure on fitting in the curriculum, and obviously doing lots of SATs preparation, let alone finding 15 minutes to go and run (ID4, Teacher, School 2)*

> *I would make sure it was done because I'm so passionate about it but I can't speak on behalf of the other teachers who are in there teaching because obviously, we do have a really tight timetable (ID7, Teacher, School 3)*

> *It's just putting it into your timetable… I think sometimes it's just a state of mind some of these changes and once its embedded then it's just normal curriculum (ID9, Leader, School 4)*

The fundamental aim of school being about learning academic subjects (e.g. English, maths, science) was raised by several participants. This was highlighted by staff in all roles, appearing to be a widespread sentiment. Participants spoke about how policy makers were too focused on academic outcomes, particularly English and Maths. Even those without teaching responsibilities spoke about how maths would be prioritised over the Daily Mile for example. One leader mentioned that if the opportunity cost of the Daily Mile had a detrimental impact on history knowledge it would be accepted, whereas if maths performance were to decline it would not be. Staff at School 6 thought that being healthy, mentally well balanced and having good wellbeing were important, yet these were not metrics that teacher's performance was judged on. Teachers and pastoral staff expressed dissatisfaction at this, with one noting that if targets related to health or exercise, schools might prioritise these differently.

> *Well, nobody cares if children haven't got the history knowledge but they do care if they don't do well in Maths (ID10, Leader, School 6)*

> *I mean I think if it came down to these children have to understand this bit of Maths, then people would say, 'Okay, well, we're not going to run today' (ID4, Teacher, School 2)*

> *I say rather cynically, any organisation will respond to what they are going to be measured on. They're not going to be measured on how healthy their kids are at – by year 6, or what weight they are, or how many hours of exercise they do. They're going to be measured on whether they get their SATs (ID11, Safeguarding lead, School 6).*

5.4 **Discussion**

The aim of this qualitative study was to explore the decision-making process regarding the implementation of public health interventions in schools, outcomes schools value, prioritisation compared to other school services and the associated resource implications. Having gained insight into the tangible and intangible costs, and important outcomes, the schools' perspective is now better understood and can be used to inform methods for economic evaluation in this setting. The Daily Mile was used as a case study, with participants seeming to believe it to be a largely free initiative. There was recognition that there was an opportunity cost of lesson time, although the academic subjects that schools were assessed on tended to be 'protected' from this. Leadership staff reported a need for evidence of effectiveness to justify resources expended, reporting tight budgets and alluding to CEA. Benefits that schools would hope to observe included improved wellbeing and a legacy of participation and enjoyment of physical activity. Whilst academic attainment was reported as their primary objective, staff acknowledged its inter-dependence with good health and wellbeing.

5.4.1 Reflections on the research

The interviews were conducted over a four-month period, and mostly increased in duration as they progressed. This was likely due to a number of factors. Firstly, the interviewer became more confident and familiarity with the topic guide improved. This enabled more active listening and more probing of topics, without being overly concerned about the next topic or what remained. The recognition of closed questions and their avoidance also improved. Some participants were clear that their time was limited at the outset; therefore, the interviewer was aware and guided the discussion closer to the content in the topic guide. Few interviews were ended prematurely, with most drawing to a natural close. Experience also enabled the framing of questions to evolve. This included changing terminology to be more comprehensible to staff and altering how questions were posed by providing more context. As the staff were unfamiliar with economic terms for example, the presentation and description of economic evaluation was adapted after early interviews. Attempting to describe how an ICER could be generated from evidence was changed to merely highlighting the possibility of presenting the cost per unit of relevant effect. Challenges like this were discussed with supervisors and solutions arrived at collaboratively.

It became clear that some topics were not relevant to all participants. For example, some staff had little exposure to, or knowledge about, the Daily Mile. In later interviews, participants were asked to consider hypothetical health and wellbeing initiatives in order to explore decision making and costs in addition to or in the absence of Daily Mile experience. As schools regularly introduce new initiatives, participants did not find this challenging. This also strengthens the findings of these interviews, perhaps allowing some generalisation beyond the Daily Mile as a single case study. Findings that staff found it challenging to implement initiatives in the busy curriculum and recognised the opportunity cost would be applicable to many initiatives. Participants' preferences for enduring behaviour change as the outcomes as opposed to fitness and BMI would likely be relevant to initiatives beyond those increasing physical activity. Not all findings would be generalisable however, for example the specific costs associated with The Daily Mile, such as painting a track and purchasing additional first aid supplies.

The sample comprised 14 participants, which might be considered small. Some schools only contributed one participant, which limited the opportunity to compare views within schools and the views of one individual could not be assumed to be reflective of a whole school. Some schools did have multiple participants with different responsibilities, which enabled such comparisons to be made. These comparisons suggested that views elicited were not extreme, isolated perspectives. A purposive sampling approach was intended, although in practice all potential participants were approached. Despite this, all levels of school Ofsted rating were represented; therefore, the findings are not reflective of just high achieving schools. It might have been that health and wellbeing initiatives are not prioritised in high performing schools due to the academic opportunity costs, which manifests in better academic performance. Alternatively, lower achieving schools may be striving to improve academic outcomes, thus also having a lesser commitment to implementing such initiatives. Academy sponsored and local authority schools were also equally represented, so findings were not representative of one school funding model. In addition, a range of staff roles and differing levels of experience were included.

It became evident in the interviews that familiarity with the Daily Mile trial outcomes was dependent upon the staff members' role. Teachers had mostly administered the questionnaires with their classes, and pastoral staff seemed to have had informal discussions with teachers regarding their completion. Leadership staff were mostly unfamiliar with the trial outcomes. Visual aids were provided to facilitate the discussion, although the length of the MDI limited the amount of detail that could be conveyed. This perhaps hindered participants' ability to comment on what outcomes were and were not being collected. Variability in participants familiarity allowed a variety of views, from those with experience of using the measures and those with 'fresh eyes'. Participants were probed about whether they thought any outcomes were missing from the Daily

Mile trial. Detailed feedback from all participants regarding the questionnaires could have facilitated interpretation of the study results more and been used to support the validity of the measures in this setting. This was not the intended aim of the study, however.

On reflection, one group perhaps omitted from this study were school pupils, although it is less clear how their perspective might contribute. Children have minimal input into school decision making and their influence is even more distant than teachers'. Their views on desirable outcomes of health and wellbeing initiatives like the Daily Mile and those collected in the trial, for example the MDI, could have been useful. They could have assisted in interpreting the findings of the Daily Mile trial, understanding mechanisms and reasons for missing data. In addition, they could have expressed what works and does not work for them in relation to the intervention, thus informing future successful intervention design. Conducting interviews with children is challenging, requiring the adaptation of qualitative research techniques used with adults and development of new approaches. This was methodologically beyond the scope of this thesis, although could be an area for future research. Parents are also stakeholders who could have provided an important contribution. Head teachers raised that parents' views were a consideration when prioritising public health interventions, and parental engagement has been highlighted as a potential factor that moderates the success of interventions for obesity prevention (Brown et al., 2019). It could also be that parents have expectations regarding the outcomes that should be targeted and their own role in contributing to the funding of public health interventions delivered in schools.

Recruiting participants was challenging. Several schools approached did not respond to multiple contacts, using both email and telephone. Balancing repeatedly contacting potential participants and maximising study recruitment was considered, as well as maintaining positive relationship with the schools and not jeopardising trial participation. A minority replied but declined, citing a lack of time. Some individuals agreed to take part initially, but either cancelled or did not arrange a time. Several participants were recruited using a snowball approach, which was most successful after the initial contact had been interviewed. This might suggest that the process was not as burdensome or worrying as initially thought, which was then communicated. Indeed, participants 'rounded up' to be interviewed while I was at a school were initially unaware about possible audio recording and apprehensive about participating. Explaining the study and obtaining informed consent alleviated their concerns.

Lessons could be learned about recruiting and interviewing participants in this setting. Information sheets could be simplified or adapted to alleviate the concerns about participant burden and the topics being discussed. Incentives were not provided to participants. This eliminated the possibility that individuals were motivated for financial reasons. It does raise the possible issue that only enthusiastic staff and those interested in research were willing to participate. Again, as views were not overwhelmingly positive this is likely to have not occurred. Topic guides could be restricted to only the most pertinent of prompts, as to optimise data collection in the short time teachers appear to be available.

The interviews were relatively informal, with participants apparently prepared to speak openly and not censor their comments. In fact, several were openly critical of the Daily Mile and the government prescribed metrics and curriculum. The topics discussed were not especially sensitive or personal. Two interviews were conducted with multiple individuals in the room, and another was in a staffroom which was utilised throughout. There was no noticeable shift in tone when others were present.

5.4.2 Findings in context

Whilst there have been many economic evaluations conducted of school-based health and wellbeing interventions, there is currently no qualitative evidence exploring schools' views on costs incurred and evidence valued. Arnold et al. (2016) and Bennett et al. (2016) both found that academic attainment is prioritised by schools but is accompanied by an understanding of the importance of health promotion, and these align with the findings of this study. This qualitative study took more exploratory view than Jessiman et al. (2019), and did not identify decision makers a-priori, guiding the inclusion criteria. Nevertheless, these studies arrived at similar conclusions regarding the role of senior leadership, regardless of academy or local authority-maintained status.

Various funding sources such as the curriculum budget, Sports Premium and charities were mentioned in this study as supporting initiatives, thus highlighting the challenge of identifying where the costs lie when providing public health interventions in schools. The lack of clarity and consistency regarding funding sources, and the inclusion of staff wages in published cost-effectiveness analyses could be due to these ambiguities.

Similar to the healthcare system and public health services, financial resources are limited in education in England (Belfield et al., 2018). Economic evaluation would be valuable as there is a need to demonstrate value for money when implementing initiatives. Potential benefits seen as important by the schools included wellbeing, enjoyment and increased participation. A Delphi study (Morton et al., 2017) examining criteria for prioritising school environment-focused physical activity interventions also found that mental health and wellbeing were the most valued outcome, above physical activity. Literature has highlighted the important role of enjoyment for the success of physical activity interventions in adolescents and mid to older age adults (NIHR Dissemination Centre, 2019). Moreover, physical literacy which includes enjoyment has been associated with levels of physical activity in young children. In all groups, this does not appear to be measured as an outcome in trials (NIHR Dissemination Centre, 2019). These other outcomes reflect the schools' desire to have a long-term impact on the children's mental health and lifestyle choices in the future. It could be that there is evidence linking participation in sport at a young age and long-term health outcomes or resource use, which could be used in exploratory analysis to model the benefits in a way that appeals to decision makers in this setting. Nevertheless, enjoyment and increased participation are outcomes that could be collected in future trials.

5.4.3 Conclusion

These interviews revealed that from their perspective, change in wellbeing is an outcome that school staff value when implementing new initiatives. This raises the question of whether wellbeing should be used as an outcome to provide evidence to the school decision makers that an intervention is a worthwhile investment of both time and financial resources. Furthermore, the UK government recognises that wellbeing is instrumental to health policy (Department of Health, 2014). The potential to include wellbeing in economic evaluation has been recognised, although its operationalisation is reportedly hindered by challenges with measurement (Brazier and Tsuchiya, 2015). The wellbeing adjusted life year (WELBY) has been mooted, although its compatibility with the QALY framework and current decision making paradigms is also problematic (Brazier and Tsuchiya, 2015). There are also efforts to generate a new preference-based measure for adults suitable for measuring health and non-health benefits for use in both health and social care (Peasgood and E-QALY team, 2018). The findings of the extending the QALY (E-QALY) programme are yet to be published, however.

Although the views expressed here about both financial and opportunity costs are important and should be taken into account in future economic evaluations, the financial costs for this particular evaluation were very small. This leaves two major issues for further study. The first concerns the opportunity costs in terms of lost learning time, and potential impact on academic attainment. This is a complex issue and there is currently work going on in this area (Andronis et al., 2019). The second concerns the challenges associated with measuring wellbeing. Given the views expressed by school staff about the importance of wellbeing and the limited and relatively untested methods available for measuring wellbeing in children, using a wellbeing measure in economic evaluation requires further information on the psychometric properties and feasibility of using such measures. The next chapter of this thesis therefore moves on to explore these issues using the MDI, as a precursor to incorporating it in an economic evaluation of The Daily Mile intervention.

CHAPTER 6 AN EVALUATION OF THE MIDDLE YEARS DEVELOPMENT INSTRUMENT (MDI) IN A UK POPULATION SAMPLE, FOR USE IN ECONOMIC EVALUATION

6.1 Introduction

Wellbeing has been noted as a potential or complementary alternative outcome to HRQL in economic evaluation by policymakers (Peasgood et al., 2019) and local authority decision makers (Frew and Breheny, 2019a). There is ongoing work into broadening economic evaluation beyond the QALY (Brazier and Tsuchiya, 2015), particularly for multi-sector interventions. Wellbeing also emerged as a relevant outcome in school-based economic evaluations in Chapter 5 by school staff yet it was not reported in any studies identified in Chapter 3. Given this interest, if there was a suitable measure of wellbeing for use in primary school-aged children available, there is the opportunity to include it in economic evaluation.

It was highlighted in Chapter 2 that measuring self-reported outcomes in children is problematic. Changing reading ability, comprehension and sense of self as children age mean that one single wellbeing scale covering the whole of childhood is perhaps unrealistic. The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) (Tennant et al., 2007) is widely used in adults and adolescents, although it has not been used in children younger than 13 and there is no child version. Children's wellbeing measures are available, although the evidence of their psychometric performance and ability to show change over time is largely incomplete. These are important to demonstrate if a scale is to be used in research. Furthermore, none have been used in economic evaluation previously. The Personal-Wellbeing Index – School Children (PWI-SC) (Cummins and Lau, 2005) is a 12-item measure of subjective wellbeing for children aged over 11. While available in many languages, the evidence supporting its use is sparse with the exception testing of the scale structure. The Stirling Children’s Wellbeing Scale (Liddle and Carter, 2015) also comprises 12-items and measures emotional and psychological wellbeing. In contrast with the PWI-SC, it can be used in children as young as eight, although the evidence of its use over time is also unknown.

PHE provides examples of how children’s wellbeing could be measured (PHE, 2015). The only outcomes suitable for those younger than adolescents are the Heubner’s Student Life Satisfaction Scale (SLSS) (Huebner, 1991) and the Children’s Society Good Childhood Index (Rees et al., 2010). The SLSS is a brief seven item measure of life satisfaction and the Children’s Society Childhood Index is a 10-item measure of happiness in several areas of life, including friends and school. There are a multitude of indicators of wellbeing, including health, relationships and the community (Lippman et al., 2011). Whether short measures such as the Children’s Society Childhood Index can capture all of these, or whether measuring life satisfaction alone is sufficient is unknown. Whether one questionnaire item can adequately measure one dimension of wellbeing is also uncertain. A multi-dimensional measure of wellbeing may be the most appropriate way to measure this construct.

The MDI (Schonert-Reichl et al., 2012) is a 77-item questionnaire (Grade 4 version) that assesses many different aspects of a child’s wellbeing such as connectedness and relationships at home and school. The MDI was developed in Canada, for the purpose of population monitoring of health and wellbeing in middle childhood years (aged 6-12). The developers propose that data can be used to identify actions to be taken at a school or community level to address children's needs or could be used by policymakers to plan investments and policy changes (University of British Columbia). Despite its intended use to facilitate local and national decision making (Schonert-Reichl et al., 2012; University of British Columbia), the MDI is not currently used in research examining impact at an individual level, nor change over time. Thus far, the MDI has been evaluated in Canada (Schonert-Reichl et al., 2012), Australia (Gregory et al., 2018) and Italy (Castelli et al., 2018). The literature reporting its development and subsequent validation has predominantly focused on the structure of the sub-scales and their associations with each other (Castelli et al., 2018; Gregory et al., 2018; Schonert-Reichl et al., 2012), and its performance in population groups (e.g. age and gender).

Determining whether a measure is valid and reliable requires the collation of evidence demonstrating a number of properties. Reliability and validity also involve an interaction between the measure (e.g. the MDI), the population completing it (e.g. UK children) and the situation (e.g. clinical trial) (Streiner et al., 2015). Thus far, MDI evidence has not been generated in a UK population, nor in a clinical trial setting. Construct validity of a scale refers to whether the questionnaire is an adequate reflection of the construct it is intended to measure (Mokkink et al., 2010). Put simply, does it measure what it is supposed to measure? In the case of the MDI, this would be whether the Optimism subscale measures optimism, and not an unrelated attribute such as mobility. Construct validity can be tested by exploring whether it performs according to theoretical expectations (Carmines and Zeller, 1979). A framework of testable hypotheses can be developed, that describe links and relationships between the construct and determinants of it (Cronbach and Meehl, 1955). This can be operationalised by testing whether populations that would be expected to differ do report different scores (known groups validity) and by testing whether scores correlate with other measures that are theoretically related (convergent validity). Gregory et al. (2018) evaluated the construct validity of the Grade 7 MDI by examining its association with the EPOCH Measure of Adolescent Well-being, which assesses five positive psychological characteristics. They concluded that the MDI had good convergent validity, however the Grade 7 MDI was used, which has questions targeted specifically at older children. No studies have examined the relationship between the Grade 4 MDI and other outcomes as previous research relied upon testing associations with other MDI subscales (Castelli et al., 2018; Schonert-Reichl et al., 2012).

Reliability of a measure is the degree to which the same or similar results are obtained on different occasions or using similar tests. One way of measuring scale reliability is internal consistency. This is an assessment of whether all items in a scale are homogeneous, measuring the same construct (Streiner et al., 2015). Good internal consistency indicates scores can be interpreted as a good reflection of the scale's content, thus repeated testing is less prone to error (Streiner et al., 2015). If removing an item from a scale improves measures of homogeneity or it shows little association with the scale score, it suggests it is not consistent with the other items (Fayers and Machin, 2016). Evidence of the MDI scales' internal consistency is mixed (Castelli et al., 2018; Gregory et al., 2018; Schonert-Reichl et al., 2012).

In a context where the effects of alternative interventions are being compared, it is important for a measure to detect change over time, when change has actually occurred in the construct being measured (Mokkink et al., 2010). There is no research on the responsiveness of any of the MDI versions. A measure unable to detect change would be of limited use, and one cause of this could be a limited distribution of scores, or 'clumping'. If individuals are already scoring at the highest or lowest levels, their scores have no potential to improve or decline over time, respectively. Other threats to a scale's use include feasibility. Its length would suggest that the MDI is a very comprehensive measure but whether this justifies the additional response burden and potential non-completion is uncertain.

The Birmingham Daily Mile trial was a cluster RCT and has already been described in detail in Chapter 4. Primary schools were randomised to either their usual health and wellbeing activities, or the Daily Mile for a duration of 12 months. HRQL and wellbeing questionnaires were administered at baseline and at 12-month follow-up, alongside anthropometric (BMIz, bodyfat), fitness and educational attainment outcomes. HRQL was measured using the CHU9D (Stevens, 2009). This is a short questionnaire, comprising nine items which can be used to calculate a preference-based utility score for use in economic evaluation. It has been used widely in studies in the UK (Kwon et al., 2019; Wolstenholme et al., 2018). Evidence of the validity and reliability of a measure should be established in the context of how it is expected to be used. There is a paucity of evidence regarding the MDI's validity and reliability in the UK and no longitudinal administration in a comparative situation testing responsiveness. Broadening the scope of economic evidence beyond HRQL requires robust measures, with evidence of these properties.

6.1.1 Aims

An economic evaluation from the schools' perspective should include outcomes deemed be important in this setting. Qualitative interviews reported in Chapter 5 indicated that wellbeing was such an outcome. In the absence of a widely adopted and responsive wellbeing measure, the aims of this analysis were to examine the suitability of the MDI as a measure of wellbeing, for potential use in economic evaluation. Specific aims are:

1. Assess the psychometric properties of the Grade 4 MDI in a UK population in school years 3 and 5.
2. Evaluate the acceptability of the MDI as a measure of wellbeing for use in economic evaluations of interventions targeted at children.

6.2 Methods

6.2.1 Data collection

6.2.1.1 Participants

Participants were children in years 3 and 5 at baseline (aged 7-8 and 9-10), attending one of 40 schools participating in the Birmingham Daily Mile trial. Eligible participants for this analysis were allocated to either the intervention or control arm and completed the MDI at baseline.

6.2.1.2 Procedure

Child-reported questionnaire data were collected electronically using tablet computers. Children completed the measures independently, under the supervision of teachers who were able to provide assistance if required. Questionnaire order was standardised for all participants. Socio-demographic data were obtained from children's school records. Reports of participants' academic attainment were collected from the children's class teacher.

6.2.2 Measures

The Birmingham Daily Mile trial (Breheny et al., 2018) collected a range of outcome measures at baseline, 4 months and 12 months. The MDI was only administered at baseline and 12-month follow-up, thus only data from these timepoints was used. Socio-demographic data collected included age, gender and ethnicity. Deprivation data were obtained at the school level, using the IMD aligning with the school's postcode. Finally, wellbeing and quality of life were measured using the MDI and CHU9D, respectively. The questionnaires were completed electronically, and most questions were forced-choice. Individual respondent data were not recorded if the survey was abandoned prior to completion. This resulted in no item level missing data.

6.2.2.1.1 MDI

The MDI measures children's social and emotional health and well-being in middle childhood (ages 6-12). This is a time when children experience important changes that establish their identity and impact their adolescent and adult development (National Research Council, 1984). The MDI is a self-report, multiple choice questionnaire asking children about their thoughts, feelings and experiences.

The measure assesses five dimensions of child development:

1. Social and emotional development (optimism, happiness, self-esteem, sadness)
2. Physical health and wellbeing (general health, nutrition, sleep)
3. Connectedness (presence of supportive adults and sense of belonging with other children)
4. Use of after-school time (time spent participating in organised and unorganised activities)

5 School experiences (including school belonging, future goals, bullying)

Sixteen subscale scores can be generated, calculated as the mean of all items within the scale. Scores can range from 1-5, which is consistent across all subscales with the exception of the general health item (range 1-4). Higher scores reflect better wellbeing, with the exception of the sadness and bullying scales, which are reverse scored.

A wellbeing index can be calculated, in addition to subscale scores. This is a combination of optimism, satisfaction with life, self-esteem, sadness, and general health item. Scores are categorised into three levels of wellbeing: Thriving, medium to high wellbeing or low wellbeing. A sum score suitable for statistical analysis can also be produced using these items. This can range from between 15 and 74, with higher scores reflecting better wellbeing. Table 8 details the MDI domains and scales.

Table 8 MDI domains and scales

Domain	Scale	Scale description	Example item
Social and emotional development (SED)	Empathy	Being able to take another person’s perspective and feel what they are feeling	I care about feelings of others
	Prosocial behaviour	Voluntarily choosing behaviours that help and benefit others	I helped someone who was hurt
	Optimism	The mindset of having positive yet realistic expectations for the future and making efforts to achieve them	I have more good times than bad times
	Self-esteem	How children judge their abilities and their self-worth.	I like being the way I am

Domain	Scale	Scale description	Example item
	Satisfaction with life (SWLS-C)	Contentedness and happiness with life	Things in my life are excellent
	Sadness	Frequency of overwhelming sadness	I feel unhappy a lot of the time
	Worries/anxiety	Frequency of excessive or compulsive worrying	I worry a lot that people won't like me
Connectedness	Adult relationships (home)	The presence of adults who care for them, believe they will be a success and listen to them	There are adults that I can talk to
	Adult relationships (at school)		
	Adult relationships (in the neighbourhood)		
	Peer belonging	Feelings of fitting in, acceptance and close friendships	I feel I belong with other kids
	Friendship intimacy		I have a friend I can tell everything to
School experiences	Academic self-concept	Beliefs about academic ability and what they are like as a student	Even if it's hard I can learn it
	School support	The ambiance of the school environment, including staff and child interactions and how children treat each other	People care about each other at school
	Bullying	Intentional and repeated aggressive behaviour to cause harm or embarrassment	I experience cyber bullying
	School belonging	Comfort in the learning environment	I feel like I belong in this school

Domain	Scale	Scale description	Example item
Wellbeing	Wellbeing sum score	Optimism Satisfaction with life Self-esteem Sadness General health	

6.2.2.1.2 <u>Satisfaction with life scale – child (SWLS-C)</u>

The SWLS-C is an instrument that measures global life satisfaction and is nested within the MDI. It is a five-item measure with a five-point Likert scale response format. Options range from 'Agree a lot' to 'Disagree a lot'. Originally a measure for use in adults, a paediatric version was created by modifying the item wording.

6.2.2.1.3 <u>CHU9D</u>

The CHU9D is a generic HRQL measure for children aged between 7 and 11 years (Stevens, 2009). It is a self-report questionnaire that comprises nine questions (dimensions), each with five response options. Children respond based on how they are feeling today. The questions address worry, sadness, pain, tiredness, annoyance, problems with school work, their daily routine and their ability to join in with activities.

The CHU9D was developed with the intention of being used as a preference-based measure for use in paediatric economic evaluations. Each question has a tariff, obtained from UK (Stevens, 2012) or Australian (Ratcliffe et al., 2012) valuation studies. The Australian valuation set was obtained from an adolescent sample using best-worst-scaling methods, in contrast to the UK study which used a standard gamble task and was completed by adults. The analyses reported here uses the UK value set (Stevens, 2012) due to its reflection of UK population's preferences. Utility weights can be calculated for each child's response, enabling the measure to be used to calculate QALYs for CUA. Scores can range from 0.33 to 1, with higher scores reflecting better HRQL.

6.2.2.1.4 Academic attainment

Teachers rated how children were performing against their age-related expected level in reading, writing and maths. They were rated as either at the expected level, above, below or well below it. A score of 1-4 was given for each subject, which was then summed to generate an overall attainment score.

6.2.3 Analysis

6.2.3.1 Sample characteristics

Socio-demographic characteristics (age, gender, IMD group, year group, ethnicity) were tabulated for the whole sample, and by category of MDI missing data (baseline missing/12 month follow-up missing). Categorical variables were summarised using frequencies and percentages. Continuous variables were summarised by appropriate statistics, depending on the distribution of the data (mean, standard deviation (SD) or median).

6.2.3.2 Descriptive statistics

Descriptive statistics (mean, standard deviation, median, minimum and maximum) are presented for the MDI scales at both baseline and follow-up.

Floor and ceiling effects at baseline were examined. A high percentage of participants scoring the best or worst score would suggest that the measure is unable to discriminate between children with known different levels of wellbeing and could be unable to detect improvement or worsening. If children are reporting the maximum score at baseline, they have no potential to improve. This may mean that the scale does not detect change when it has occurred.

6.2.3.3 Internal consistency reliability

Internal consistency of the MDI scales (detailed above) was assessed using Cronbach's alpha (α) and item-total correlations (Streiner et al., 2015). Item total correlations are the correlation between the item and the scale total, omitting that item. It is recommended that item-total correlations are greater than 0.2 (Kline, 1979) to indicate homogeneity. An α threshold of 0.7 was used as a threshold to indicate good item interrelatedness (Tavakol and Dennick, 2011), although high α values approaching 1.0 may indicate item redundancy (Boyle, 1991). α is dependent on the number of items in the scale, therefore as some scales have few items, results were interpreted with caution. α was also re-calculated when eliminating each item from the scale. This can indicate which items contribute to low reliability (Streiner et al., 2015).

6.2.3.4 Construct validity

6.2.3.4.1 Convergent validity

Convergent validity is the association between the measure being evaluated and outcomes measuring the same or overlapping constructs (Fayers and Machin, 2016). Hypothesised relationships are generated a-priori and the associations tested statistically. To assess the convergent validity of the MDI, hypotheses were formed by a process used previously by (Al-Janabi et al., 2013; Coast et al., 2008b). Four investigators independently assessed whether they thought there would be an association between MDI scale scores and items/scores from other outcomes. These outcomes included CHU9D items and overall utility score and teacher-rated academic attainment in a composite score of performance in reading, writing and maths. If three or more investigators hypothesised that a relationship existed, these were tested statistically. Spearman's rank correlations were calculated between baseline scores. Spearman's rank was chosen due to the skewed nature of CHU9D and MDI data. Cohen's interpretation of correlation coefficients was used (large >0.5, moderate 0.5-0.3, small 0.1-0.3, insubstantial <0.1) (Cohen, 1988).

6.2.3.4.2 <u>Known groups/divergent validity</u>

Known groups or divergent validity compares the outcome measure in groups that you would expect to differ. Known groups were selected based on existing literature and hypothesised differences. As measuring wellbeing is complex and largely unexplored in children, known groups based on evidence of differing HRQL was considered. Mixed linear regression models were conducted comparing baseline mean scores on the Wellbeing Sum Score and MDI scale scores between known groups. Mixed models were used due to the clustered design of the trial. The group was the predictor and MDI score the dependent variable. Covariates included age, gender, socioeconomic status, ethnicity (where appropriate). Statistical significance was considered at the $p<0.05$ level.

The Office for National Statistics (ONS) reports that good health and learning/development are determinants of wellbeing (Office for National Statistics, 2018). Socioeconomic status has been associated with wellbeing (Patalay and Fitzsimons, 2016) and differences in child reported quality of life have been reported for gender (Breslin et al., 2012; Sanders et al., 2015; Williams et al., 2005) and ethnicity, as well as close relationships between quality of life and wellbeing (Mukuria et al., 2016; Pinto et al., 2017). Based on this evidence, the following known groups were explored.

- **Child-rated general health** (excellent/good v fair/poor) (child rated health is included in Wellbeing Sum Score, therefore this hypothesis was not be tested on the MDI Wellbeing Sum Score)
    - Children reporting poor health would report lower wellbeing

- **HRQL**. Children were split based on whether their CHU9D score was above or below the median index score.
    - Children with low quality of life would report lower wellbeing
- **Academic attainment** (whether the child is performing at or below the expected level). Analysis was conducted for reading, writing and maths.
    - Children performing below the expected level would report lower wellbeing
- **Gender** (male v female)
    - Females would report lower wellbeing than males
- **Age** (year 3 v year 5)
    - Children in year 3 (aged 7-8) would report higher wellbeing than year 5 children (aged 9-10), as wellbeing reportedly declines with age.
- **Socio-economic status** (deprived or not deprived, using IMD score from child's school postcode)
    - Children in deprived areas would have lower wellbeing
- **Ethnicity** (white or non-white)
    - Non-white children would have lower wellbeing.

6.2.3.5 Responsiveness

Responsiveness is the ability of a measure to detect a change over time, where a change in the participant's condition has occurred. Of interest is whether the measures are sensitive to change in health and wellbeing between baseline and 12-month follow-up. Responsiveness can be evaluated using three methods. An external indicator (anchor) that categorises participants into groups that reflect changes in their health can be used and between group differences tested. Anchors should be associated with the scale's construct, demonstrated using correlations or justified by theoretical importance (Wyrwich et al., 2013). Change scores between baseline and follow-up can be compared between groups using analyses of variance. Alternatively, standardised response means (SRMs) and effect sizes can be calculated, and statistics compared between groups (Wyrwich et al., 2013).

6.2.3.5.1 <u>Anchor method</u>

For each MDI scale, change in index score was calculated between baseline and follow-up for each participant. Mean difference was compared between groups improving, worsening and those not changing using mixed linear models (significance $p<0.05$), with confidence intervals. Covariates included school, age, gender, study arm, ethnicity, and socio-economic status. Anchors chosen were change in general health and change in quality of life. Academic attainment was considered, however Spearman's correlations found that no MDI scales were associated with this construct.

- Change in General Health
    - Change scores were calculated for the 'General Health' question. Children were classified as no change, improved or worsened.
- Change in HRQL
    - Participants were classified as improved/no change or worsened quality of life, using the CHU9D utility score. Improvement or decline was defined as a change of half a standard deviation of change score.

6.2.3.5.2 Distribution methods

SRMs and effect sizes were calculated for the MDI scale scores and Wellbeing Sum Score. SRM is mean change score divided by the standard deviation of change of the group. Effect size is the change score divided by the standard deviation at baseline. Suggested interpretation of effect sizes (Cohen's d) are 0.20 small change, 0.50 moderate change and 0.80 large change (Cohen, 1988). Effect sizes were compared between anchor groups, with the expectation that effect sizes will be greater for participants experiencing change.

6.3 **Results**

6.3.1 Sample Characteristics

The complete Daily Mile dataset comprised 2,280 children. Of these, 527 did not provide any MDI data at either baseline or second follow-up. These children were removed from the 'MDI Validation Dataset'. In addition, 201 children did not provide MDI data at baseline. These children would not contribute any data to the MDI validation analyses and were excluded also. The MDI Validation Dataset therefore comprises 1,552 children, of which 788 are complete cases at baseline and follow-up. Table 9 presents the sample characteristics for the MDI Validation Dataset. The data is reported for the whole sample, and by complete/incomplete cases.

Just over half of the sample completed data at both timepoints. More children with complete data were of white ethnicity and attended schools in areas of low deprivation compared to those with partial data. There was no difference in baseline MDI or CHU9D scores between those with complete and partial data, with other characteristics also similar. Overall, there were more males than females and more children in year 5 than year 3.

Table 9 Sample characteristics

		Data group		Whole sample
		Complete case	Partial (baseline only)	
N (%)		788 (50.8)	764 (49.2)	1552
Child Age in years, mean (SD)		8.9 (1.0)	8.9 (1.1)	8.9 (1.0)
Gender N (%)	Female	376 (47.7)	348 (45.5)	724 (46.6)
	Male	412 (52.3)	416 (54.5)	828 (53.4)
Ethnicity N (%)	White	454 (57.8)	372 (48.7)	826 (53.3)
	South Asian	117 (14.9)	120 (15.7)	237 (15.3)
	Black African/ Caribbean	68 (8.7)	77 (10.1)	145 (9.4)
	Other/not specified	146 (18.6)	195 (25.5)	341 (22.0)
Deprivation group N (%)	High deprivation	671 (85.2)	535 (70.0)	1206 (77.7)
	Low deprivation	117 (14.8)	229 (30.0)	346 (22.3)
Year group N (%)	Year 3	374 (47.5)	352 (46.1)	726 (46.8)
	Year 5	414 (52.5)	412 (53.9)	826 (53.2)
Baseline Wellbeing sum, mean (SD)		58.2 (11.0)	58.1 (10.9)	58.2 (11.0)
Baseline CHU9D, mean (SD)		0.837 (0.16)	0.835 (0.17)	0.836 (0.16)

6.3.2 *Descriptive statistics*

Table 10-Table 12 present descriptive statistics for the MDI scales at baseline and follow-up. The proportion scoring the maximum and minimum for each scale (floor and ceiling and effects) are also reported. For the SED domain, four scales (Empathy, Optimism, Self-esteem and Satisfaction with life) had median scores greater than 4. This suggests data are skewed towards the top of the scale, which reflects better wellbeing. This is reflected by the number reporting the maximum, with between 22% and 40% of participants reporting the ceiling value. The Sadness and Worries/Anxiety scales are reverse scored, but do not show as large ceiling effects. In the connectedness domain, all scale scores show evidence of a ceiling effect, ranging from 34% to 55%. Data from the School Experiences domain also suggests a tendency to report scores from the better end of the scale. Across all the scales, there is little evidence of floor effects.

Table 10 Descriptive statistics for scales in Social and Emotional Development domain

	Social and Emotional Development Domain													
	Empathy		Prosocial behaviour		Optimism		Self-esteem		Satisfaction with life (SWLS-C)		Sadness		Worries/anxiety	
	BL	F	BL	F	BL	F	BL	F	BL	F	BL	F	BL	F
Mean	4.20	4.26	3.32	3.16	3.98	3.98	4.27	4.28	4.01	4.10	2.65	2.54	3.11	2.91
SD	0.95	0.84	1.21	1.14	1.05	0.99	0.94	0.91	1.00	0.97	1.12	1.05	1.35	1.40
Median	4.67	4.67	3.33	3.00	4.33	4.33	4.67	4.67	4.20	4.40	2.67	2.33	3.33	3.00
Min	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Max	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Floor N (%)	26 (1.68)	7 (0.89)	48 (3.09)	20 (2.54)	42 (2.71)	17 (2.16)	31 (2.00)	12 (1.52)	34 (2.19)	12 (1.52)	222 (14.30)	109 (13.83)	219 (14.11)	162 (20.56)
Ceiling N (%)	518 (33.38)	250 (31.73)	276 (17.78)	83 (10.53)	418 (26.93)	194 (24.62)	617 (39.76)	308 (39.09)	336 (21.65)	204 (25.89)	56 (3.61)	12 (1.52)	221 (14.24)	99 (12.56)
N	1552	788	1552	788	1552	788	1552	788	1552	788	1552	788	1552	788

BL - Baseline F - 12-month follow-up

Table 11 Descriptive statistics for scales in the Connectedness domain

	Connectedness domain									
Scale	Adults at school		Adults neighbourhood		Adults at home		Peer belonging		Friendship intimacy	
	BL	F	BL	F	BL	F	BL	F	BL	F
Mean	3.29	3.30	3.03	3.06	3.45	3.54	4.12	4.17	4.33	4.39
SD	0.76	0.74	1.00	0.98	0.75	0.70	1.03	0.99	1.00	0.96
Median	3.33	3.50	3.33	3.33	3.67	4.00	4.33	4.50	4.67	5.00
Minimum	1	1	1	1	1	1	1	1	1	1
Maximum	4	4	4	4	4	4	5	5	5	5
Floor N (%)	35 (2.26)	10 (1.27)	159 (10.24)	70 (8.88)	36 (2.32)	11 (1.40)	34 (2.19)	13 (1.65)	36 (2.32)	19 (2.41)
Ceiling N (%)	539 (34.73)	268 (34.01)	537 (34.60)	264 (33.50)	759 (48.90)	430 (54.57)	537 (34.60)	288 (36.55)	772 (49.74)	421 (53.43)
N	1552	788	1552	788	1552	788	1552	788	1552	788
BL - Baseline F - 12-month follow-up										

Table 12 Descriptive statistics for scales in the School experiences domain and Wellbeing score

	School Experiences Domain								Wellbeing summary score	
Scale	Academic self-concept		School support		Bullying		School belonging		Wellbeing sum	
	BL	F	BL	F	BL	F	BL	F	BL	F
Mean	4.32	4.37	4.17	4.00	1.93	1.71	4.10	4.05	58.2	58.9
SD	0.92	0.83	0.98	1.00	1.15	0.91	1.16	1.15	11.0	11.1
Median	4.67	4.67	4.33	4.33	1.50	1.50	4.50	4.50	60	61
Minimum	1	1	1	1	1	1	1	1	15	22
Maximum	5	5	5	5	5	5	5	5	74	74
Floor N (%)	34 (2.19)	7 (0.89)	32 (2.06)	10 (1.27)	538 (34.66)	292 (37.06)	84 (5.41)	36 (4.57)	1 (0.01)	3 (0.38)
Ceiling N (%)	651 (41.95)	329 (41.75)	558 (35.95)	229 (29.06)	70 (4.51)	12 (1.52)	681 (43.88)	325 (41.24)	47 (3.03)	29 (3.68)
N	1552	788	1552	788	1552	788	1552	788	1552	788

BL - Baseline F - 12-month follow-up

6.3.3 Convergent validity

Spearman's correlation coefficients between MDI scales and the Wellbeing Sum Score are reported in Table 13. The Bullying scale did not correlate at least moderately with any other MDI scale. The largest correlation observed was between the School Support and School Belonging scales (ρ=0.64), closely followed by Optimism and Self-esteem (ρ =0.63). The SWLS was moderately associated with the majority of MDI scales, with the exception of Sadness, Anxiety and Bullying.

6.3.3.1 Social and emotional development domain

It was expected that the Empathy and Prosocial behaviour scales would correlate most highly with each other, compared to correlations with other scales in the SED domain. Although the correlation between these scales was moderate (ρ =0.31), stronger relationships were observed between Empathy and Optimism (ρ =0.40), and between Self-esteem (ρ =0.41) and SWLS (ρ =0.34). As expected, the correlations between Optimism, Self-esteem and SWLS were large (>0.5) and higher than with other constructs in the domain. Sadness and Anxiety were expected to have large correlations and have coefficients greater than associations with other scales. A moderate correlation was observed (ρ =0.45), and for both scales was the only relationship of notable size.

6.3.3.2 Connectedness domain

It was expected that peer relationship scales (Peer belonging and Friendship intimacy) should correlate more highly than the other relationship scales (e.g. relationships with adults). This was confirmed, with the correlation (ρ =0.55) being larger than with other scales in the domain. The Peer belonging scale had large relationships with scales in the school experiences domain also (ρ >0.5), with the exception of the Bullying scale. Against expectations, the correlation between the

Adults at School and Peer Belonging scales (ρ =0.46) was larger than the association between the Adults in the neighbourhood and Adults at home scales (ρ =0.38). These scales still showed a moderate relationship, however.

6.3.3.3 School experiences

It was anticipated that all School experiences scale scores would correlate highly (ρ>0.5) and have a negative association with Bullying. The Academic Self-concept and School Support scales were negatively associated with Bullying, however this association was small. Academic self-concept and School support and School belonging were strongly related (ρ>0.5). Bullying was not moderately associated with any scale; with the largest association being with the Anxiety scale (ρ =0.23).

Table 13 Spearman's correlations between MDI scales

		MDI Scale number															
Scale number	**MDI scale**	**1**	**2**	**3**	**4**	**5**	**6***	**7***	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15***	**16**
1	Empathy	1.00															
2	Prosocial behaviour	0.31	1.00														
3	Optimism	0.40	0.28	1.00													
4	Self-esteem	0.41	0.30	0.63	1.00												
5	SWLS-C	0.34	0.30	0.58	0.59	1.00											
6	Sadness*	0.01	0.01	-0.20	-0.21	-0.17	1.00										
7	Anxiety*	0.14	0.13	-0.05	-0.05	-0.05	0.45	1.00									
8	Adult relationships (at school)	0.30	0.29	0.37	0.41	0.39	-0.14	-0.01	1.00								
9	Adult relationships (in the neighbourhood)	0.19	0.27	0.25	0.27	0.30	-0.05	-0.01	0.48	1.00							
10	Adult relationships (home)	0.32	0.23	0.34	0.41	0.42	-0.19	-0.06	0.59	0.38	1.00						
11	Peer belonging	0.30	0.26	0.42	0.45	0.46	-0.18	-0.10	0.46	0.36	0.43	1.00					
12	Friendship intimacy	0.27	0.23	0.30	0.34	0.34	-0.13	0.01	0.39	0.29	0.33	0.55	1.00				
13	Academic self-concept	0.33	0.25	0.41	0.46	0.43	-0.16	-0.02	0.41	0.34	0.45	0.52	0.41	1.00			
14	School support	0.28	0.21	0.37	0.37	0.39	-0.15	0.00	0.41	0.37	0.37	0.50	0.39	0.54	1.00		
15	Bullying*	-0.06	0.14	-0.16	-0.18	-0.16	0.24	0.23	-0.12	-0.01	-0.18	-0.16	-0.10	-0.13	-0.17	1.00	
16	School belonging	0.26	0.23	0.40	0.44	0.43	-0.18	-0.03	0.48	0.34	0.42	0.57	0.43	0.55	0.64	-0.20	1.00

Values are Spearman's correlation coefficients (ρ)
Highlighted scales indicate correlations >=0.295
* Scales are reverse scored

6.3.3.4 Outcome of hypotheses generating exercise

Of 204 potential associations, 67 scales/domains were hypothesised to be conceptually related (Table 14) by at least three of the four independent raters. Of these, forty were hypothesised by all four raters to be related. No associations were hypothesised between 59 constructs, which related predominantly to CHU9D domains of Pain, Tiredness, Sleep and Daily Routine. Seventy-eight relationships were hypothesised by only 1 or 2 raters, which meant they did not meet the threshold for testing. The MDI Prosocial behaviour and Empathy scales were not anticipated to be related to any measure of HRQL or academic attainment. All remaining MDI scales were expected to be associated with the CHU9D utility. The Wellbeing Sum Score was expected to correlate moderately with academic attainment and all HRQL domains except Daily Routine. All scales in the School Experiences domain were expected to be associated with academic attainment. Relationships with Adults at school, SWLS-C, Self-esteem and wellbeing sum score were also hypothesised to be related to academic attainment.

6.3.3.4.1 <u>Results</u>

Spearman's correlation coefficients of tested associations are reported in Table 14. Academic attainment was not associated with any MDI domains, with the highest coefficient being 0.16 (Academic self-concept). In total, eight hypothesised relationships were confirmed. Four were between the Wellbeing Sum Score and CHU9D. Wellbeing showed a small association with all CHU9D domains (>0.2), although only the Sadness, School work and Joining in activities coefficients exceeded 0.3, for a moderate effect. The association between the CHU9D utility and Wellbeing sum score was the strongest relationship of any tested (ρ =0.44), confirming the interrelationship between wellbeing and HRQL.

The remaining four confirmed hypotheses will be discussed in relation to their respective MDI domains.

In the SED domain, four hypotheses were confirmed. These related to an association between Self-esteem, SWLS-C, Sadness and Anxiety and HRQL (CHU9D utility). All associations were in the expected direction, with better social and emotional development associated with greater HRQL. Correlations between the MDI and CHU9D Sadness and Worry/Anxiety scales and domains did not exceed the threshold for a moderate effect. Small associations were observed between MDI Sadness and CHU9D Sadness and Worry however. The hypothesised associations between academic attainment and Self-esteem and SWLS-C were not confirmed.

For scales comprising the Connectedness domain, expected associations were mostly with the CHU9D Worry and Sadness domains and the CHU9D utility score. No hypothesised relationships were confirmed. The Peer belonging scale showed a small association greater than ±0.25 with the Sadness, Joining in activities and CHU9D utility score however.

In the School Experiences domain no tested hypotheses were confirmed, although small associations were observed (±-0.2) between all scales in the domain and HRQL (CHU9D utility score).

Table 14 Results of hypotheses generation and testing of hypotheses

CHU9D (HRQL)

N – Count of raters hypothesising relationship r – Correlation coefficient if tested

	Worry		**Sadness**		**Pain**		**Tired**		**Annoyed**		**School work**		**Sleep**		**Daily routine**		**Joining in activities**		**Utility score**		**Academic attainment**	
MDI scale	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ	N	ρ
Empathy	1		0		0		0		2		1		0		0		2		2		0	
Prosocial behaviour	0		0		0		0		0		0		0		0		2		2		2	
Optimism	4	-0.26	4	-0.24	0		0		3	-0.22	1		0		1		3	-0.25	4	0.36	2	
Self-esteem	4	-0.21	4	-0.27	0		0		2		3	-0.23	1		0		2		4	0.34	3	0.11
SWLS-C	4	-0.18	3	-0.23	2		2		4	-0.19	3	-0.22	2		2		4	-0.27	4	0.31	4	0.05
Sadness*	3	0.21	4	0.20	1		2		2		2		2		1		3	0.16	4	-0.33	2	
Worries/anxiety*	4	0.14	4	0.14	0		0		1		2		1		0		3	0.10	4	-0.21	1	
Adult school relationships	3	-0.14	2		0		0		0		3	-0.15	0		0		1		4	0.22	3	0.07
Adult neighbourhood relationships	3	-0.10	2		0		0		0		0		0		0		1		3	0.15	0	
Adult home relationships	4	-0.12	3	-0.15	0		0		1		2		1		2		1		4	0.24	1	
Peer belonging	4	-0.19	4	-0.25	0		0		1		1		1		0		4	-0.26	4	0.29	2	
Friendship intimacy	4	-0.17	4	-0.14	0		0		1		0		1		0		2		3	0.17	0	
Academic self-concept	1		1		0		0		0		4	-0.21	0		0		0		3	0.29	4	0.16

CHU9D (HRQL)

N – Count of raters hypothesising relationship r – Correlation coefficient if tested

	Worry		Sadness		Pain		Tired		Annoyed		School work		Sleep		Daily routine		Joining in activities		Utility score		Academic attainment	
School support	2		2		0		0		0		4	-0.15	0		0		3	-0.22	3	0.24	3	0.06
Bullying	4	0.17	4	0.21	0		0		2		3	0.10	2		0		1		4	-0.21	3	-0.13
School belonging	3	-0.16	3	-0.20	0		0		1		2		0		0		4	-0.20	4	0.26	2	
Wellbeing sum score	4	-0.29	4	-0.31	4	-0.22	4	-0.20	4	-0.27	3	-0.31	3	-0.25	2	-0.27	3	-0.31	4	0.44	4	0.13

MDI – higher scores are better (except Sadness and Anxiety scales) CHU9D – lower scores are better

Academic attainment – higher scores are better

Shaded cells reflect hypotheses explored and coefficients generated that exceed +/-0.3.

6.3.4 Known groups validity

Based on evidence of factors that impact children's wellbeing and quality of life, it was anticipated that scores on the MDI would differ between certain groups of participants. Differences were tested for each of the 15 MDI scales and the Wellbeing Sum Score. Scale scores are means, which could range between 1 to either 4 or 5. Absolute differences were therefore small, with the exception of the Wellbeing sum score. Results tables are presented in Appendix J.

6.3.4.1 Gender

Across all 16 scales, females reported scores reflecting better wellbeing. The majority of between group differences were significant, with the exception of the SWLS-C ($p=0.14$), Adults in the neighbourhood ($p=0.07$) and Peer belonging ($p=0.24$) scales. All scales in the School Experiences domain showed significantly higher scores for females. All mean differences were small, with the maximum being -0.37 (Anxiety, $p<0.001$). This reflects less than half of a point difference on the Likert scale response.

6.3.4.2 Ethnicity and deprivation

No significant differences were observed between white and non-white participants for any MDI scale. Neither group consistently reported better scores, with differences ranging from 0 (Adults at home, $p=1.00$)) to 0.11 (Anxiety, $p=0.13$). No statistically significant differences were observed between children attending schools in deprived areas compared to those in affluent areas, although children in affluent areas reported better wellbeing across all MDI scales.

6.3.4.3 Age

Previous research has found that wellbeing declines with age. Neither year group consistently reported better wellbeing. Twelve comparisons were statistically significant. In the Social and Emotional wellbeing domain, year 5 children reported significantly better Satisfaction with life ($p=0.02$) and Self-esteem (p=0.01), but higher Anxiety (p<0.001). In the Connectedness domain, year 3 children had better relationships with adults in the neighbourhood (p<0.001), whereas year 5 children had better relationships with adults in the home (p<0.001). Older children reported significantly better close friendships (p=0.04). Finally, in the school experiences domain, year 5 children reported significantly better Academic Self-concept (p<0.001) and less Bullying (p<0.001), but less School Support (p<0.001). Again, differences were small, with the largest being just over a quarter of a point difference on the Likert scale (-0.29, Anxiety)

6.3.4.4 Academic attainment

In all scales except for Prosocial behaviour, mean scores were worse for those performing below the expected level. For all subjects, children performing below the expected level had significantly worse Empathy (p<0.001), Self-esteem (p<0.001), Friendship intimacy (p<0.001), Academic self-concept (p<0.001) and reported more bullying (p<0.001). Results for the remaining scales were less consistent across subjects however, with some comparisons not reaching significance.

6.3.4.5 Quality of life

It was expected that children with lower quality of life, measured using the CHU9D would report lower wellbeing. This relationship was observed across all MDI scales, with all scales except the Prosocial behaviour scale (p=0.09) reaching significance (p<0.001). Between group differences were greater than those observed in previous comparisons, with differences in Optimism (MD 0.58), Sadness (MD 0.66) and Peer belonging (0.51) reflecting differences greater than half a point on the MDI Likert scale response.

6.3.4.6 Self-rated health

Poor health was also expected to have a negative effect on wellbeing. The magnitude of the differences were similar to that of those between children with good and poor quality of life and all were in the expected direction. Differences ranged from 0.63 (Peer belonging) to -0.21 (Anxiety). All between group differences were statistically significant ($p=0.003$ to $p<0.001$).

6.3.5 Internal consistency reliability

Cronbach's Alpha (α) for each MDI scale is reported in Table 15, overall and by gender and year group. For the whole sample, α exceeded 0.7 for all scales except Sadness (0.65). For the whole sample all Sadness items correlated with the scale score > 0.2 and the exclusion of items did not increase α. This suggests that no single items fit the Sadness scale poorly. Examining the impact of deleting items on reliability, the only scale that appears to have poorly fitting items is the Relationships with Adults in the Neighbourhood scale. Removing any item would increase the reliability of the scale.

The only scale that performed notably differently in Gender and Age subgroups was Empathy. Internal consistency was less than 0.7 for Females (0.67) and year 5 children (0.68). For scales in the SED domain, internal consistency was lower for females for all scales except Anxiety. Differences by year group were not as evident. In the Connectedness domain, year 5 α values were higher for all scales compared to year 3 children. In this domain internal consistency was lower for females, except for the Adult Relationships at Home scale. All scales performed similarly for both Gender and year groups in the School Experiences domain.

Reliability for the Wellbeing sum score was good (0.86) but was slightly better in year 5 children compared to year 3. Exploratory analysis calculated reliability for the domains, including items from all appropriate scales. Reliability was good for all domains, ranging from 0.81 (School experiences) to 0.89 (Connectedness).

Table 15 Internal consistency reliability for MDI scales and domains

Domain	**Scale**	**α overall**	**α** Gender		**α** Year group	
			Females	Males	Year 3	Year 5
Social and emotional development	Empathy	0.71	0.67	0.72	0.74	0.68
	Optimism	0.73	0.70	0.75	0.75	0.71
	Self-concept	0.79	0.76	0.80	0.80	0.78
	Sadness	0.65	0.63	0.66	0.63	0.66
	Anxiety	0.8	0.83	0.78	0.79	0.81
	SWLS-C	0.84	0.82	0.85	0.84	0.83
Connectedness	Adults- school	0.89	0.77	0.75	0.74	0.78
	Adults - home	0.76	0.77	0.81	0.79	0.81
	Adults - neighbour	0.8	0.89	0.90	0.87	0.91
	Peer belonging	0.81	0.79	0.82	0.78	0.83
	Friendship intimacy	0.83	0.80	0.84	0.81	0.84
School experiences	Academic self-efficacy	0.83	0.82	0.84	0.83	0.83
	School support	0.82	0.81	0.82	0.81	0.82
	Bullying	0.88	0.87	0.88	0.89	0.86
	School belonging	0.79	0.80	0.79	0.75	0.83
Exploratory analysis	Wellbeing score	0.86	0.85	0.87	0.84	0.87
	Social and emotional development	0.84				
	Connectedness	0.89				

	School experiences	0.81

Table 16 Item-level reliability statistics for whole sample

Domain	Scale	Label	Item-total correlation	α if item deleted
Social and emotional development	Empathy	I feel sorry for other kids who don't have the things	0.55	0.59
		When I see someone being treated mean it bothers me	0.48	0.68
		I am a person who cares about the feelings of others	0.56	0.58
	Optimism	I have more good times than bad times	0.58	0.61
		I believe more good things than bad things will happen	0.56	0.63
		I start most days thinking I will have a good day	0.52	0.69
	Self-esteem	In general, I like being the way I am	0.59	0.75
		Overall, I have a lot to be proud of	0.67	0.67
		A lot of things about me are good	0.63	0.71
	Sadness	I feel unhappy a lot of the time	0.45	0.56
		I feel upset about things	0.48	0.51
		I feel that I do things wrong a lot	0.44	0.58

Domain	Scale	Label	Item-total correlation	α if item deleted
	Anxiety	I worry about what other kids might be saying about me	0.63	0.75
		I worry a lot that other people might not like me	0.68	0.70
		I worry about being teased	0.64	0.74
	SWLS-C	In most ways my life is close to the way I would want it to be	0.57	0.82
		The things in my life are excellent	0.71	0.78
		I am happy with my life	0.75	0.77
		So far I have gotten the important things I want in life	0.61	0.81
		If I could live my life over, I would have it the same way	0.55	0.83
Connectedness	Adults-school	School - adult that cares	0.59	0.69
		School - adult that believes	0.61	0.66
		School - adult that listens	0.58	0.70
	Adults - home	Home - adult that believes	0.63	0.74
		Home - adult that listens	0.67	0.69
		Home - adult to talk to	0.62	0.74
	Adults - Neighbour	Neighbourhood - adult that cares	0.78	0.85
		Neighbourhood - adult that believes	0.80	0.83
		Neighbourhood - adult that listens	0.78	0.85

Domain	Scale	Label	Item-total correlation	α if item deleted
	Peer belonging	I feel part of a group of friends that do things together	0.62	0.77
		I feel that I usually fit in with other kids around me	0.67	0.71
		When I am with other kids my age, I feel I belong	0.67	0.72
	Friendship intimacy	I have at least one really good friend I can talk to	0.66	0.78
		I have a friend I can tell everything to	0.71	0.73
		There is somebody my age who really understands me	0.67	0.77
School experiences	Academic self-esteem	I am certain I can learn the skills taught	0.72	0.74
		If I have enough time, I can do a good job on all my school work	0.69	0.77
		Even if the work in school is hard, I can learn it	0.67	0.79
	School support	Teachers and students treat each other with respect	0.66	0.75
		People care about each other	0.73	0.68
		Students in this school help each other	0.61	0.80
	Bullying	Physical bullying	0.76	0.83
		Verbal bullying	0.76	0.83
		Social bullying	0.76	0.83

Domain	Scale	Label	Item-total correlation	α if item deleted
		Cyberbullying	0.66	0.87
	School belonging	N/A (2 item scale)		

6.3.6 Responsiveness

6.3.6.1 Anchor method

Regression analyses were conducted to assess whether changes in children's MDI scores corresponded to changes in children's HRQL and self-reported health between baseline and 12 months. Children were categorised into three groups for each anchor, reflecting whether they improved, declined or reported no change. The analyses compared those who improved or worsened to those who did not change, with the expectation that change scores would be significantly different between those experiencing change and no change.

6.3.6.1.1 <u>Quality of life</u>

All MDI scales and the Wellbeing sum scores were compared, which resulted in 16 analyses (Table 17). Eight scales and the Wellbeing sum score showed significant between group differences. For the Self-esteem, SWLS-C, Relationships with adults at home and Peer belonging scales, significant differences were observed only between those whose HRQL declined and HRQL did not change. These results suggested that these MDI scales were responsive to a decline in HRQL. One scale showed responsiveness to only improvement in HRQL, this being the Relationships in the neighbourhood scale. Improvement in HRQL appeared to correspond to a reduction in relationships, however. Results for the Optimism and Sadness scales and the Wellbeing Sum Score indicate the scales are responsive both to an increase and a decline in HRQL. No scales in the School Experiences domain (Academic self-concept, School support, Bullying) indicated responsiveness to changes in HRQL.

6.3.6.1.2 General health

As with HRQL, the MDI scales were predominantly responsive to a worsening in the children's health (Table 17). Seven scales were responsive to worsening (Empathy, Optimism, Self-esteem, Sadness, Relationships with adults at school and the neighbourhood, Peer belonging) and one to an improvement (Academic self-concept) in self-reported general health. Worsening health consistently corresponded to worsening constructs of wellbeing compared to those whose health did not change, with the exception of relationships with adults at school and in the neighbourhood. In these cases, children whose general health declined had a greater improvement in relationship score. While responsiveness of the Wellbeing sum score to change in general health was not tested, three of the four scales of which it comprises was responsive to worsening, the exception being SWLS-C.

6.3.6.2 Distribution methods

6.3.6.2.1 Quality of life

Four MDI scales indicated a responsiveness to both improvement and decline in HRQL of which three were in the SED domain (Table 18). Effect sizes were small for both improvement and decline in the Optimism and SWLS-C scales, but small/medium and medium/large for the Sadness and Wellbeing Sum scores, respectively. Four scales were responsive to improvement in HRQL only (Anxiety, Neighbourhood relationships, Home relationships and Bullying), with effect sizes in opposing directions for the relationship scales. Neighbourhood relationships improved, but Home relationships declined with improving HRQL. The effect size for the Bullying scale indicated that when no change in HRQL occurred, bullying scores improved, in addition to when HRQL increased. The School Support scale was responsive to decline in HRQL only. Effect sizes for the remaining scales did not exceed 0.2, although the Empathy, School Relationships, Friendship intimacy and Academic self-concept were notably negligible.

6.3.6.2.2 General health

Two MDI scales were responsive to both improvement and decline in general health (Table 18). Worsened health corresponded to an improvement in Neighbourhood relationships, but a decrease in Peer belonging. The inverse relationships were observed for those whose health improved. Four scales were responsive to worsening health only, with worsening health corresponding to worse scores (Prosocial behaviour, Optimism, Self-esteem, School support). The SWLS-C, Friendship intimacy and Academic self-concept scales were responsive to an improvement in general health, with small effect sizes. While the Bullying scale indicated responsiveness to health improvement, the effect size for the no change group also exceeded 0.2. The Anxiety scale also appeared responsive to no change in health.

Table 17 Regression analysis evaluating responsiveness of MDI scales to change in HRQL and general health

			HRQL				General health			
			MD*	**Upper CI**	**Lower CI**	**P-value**	**MD***	**Upper CI**	**Lower CI**	**P-value**
Social and emotional development	Empathy	Decline	-0.04	-0.23	0.15	0.70	-0.27	-0.46	-0.08	0.01
		Improved	0.08	-0.11	0.28	0.39	-0.07	-0.26	0.13	0.49
	Prosocial	Decline	-0.04	-0.28	0.20	0.76	-0.14	-0.37	0.10	0.25
		Improved	-0.04	-0.28	0.20	0.76	0.02	-0.22	0.27	0.85
	Optimism	Decline	-0.26	-0.47	-0.06	0.01	-0.22	-0.43	-0.02	0.03
		Improved	0.31	0.11	0.52	0.00	0.10	-0.11	0.31	0.33
	Self-esteem	Decline	-0.31	-0.50	-0.12	0.00	-0.24	-0.42	-0.05	0.02
		Improved	0.09	-0.10	0.28	0.36	0.09	-0.11	0.28	0.37
	SWLS-C	Decline	-0.35	-0.56	-0.15	0.00	-0.19	-0.40	0.01	0.07
		Improved	0.18	-0.03	0.38	0.09	0.13	-0.08	0.34	0.21
	Sadness	Decline	-0.36	-0.58	-0.15	0.00	-0.31	-0.52	-0.10	0.00
		Improved	0.34	0.13	0.56	0.00	0.08	-0.14	0.30	0.49
	Anxiety	Decline	0.12	-0.13	0.38	0.34	0.14	-0.11	0.39	0.27
		Improved	-0.19	-0.45	0.06	0.14	0.00	-0.26	0.26	1.00
Connectedness	Adults school	Decline	0.10	-0.06	0.26	0.22	0.19	0.03	0.34	0.02
		Improved	-0.04	-0.20	0.12	0.61	-0.05	-0.22	0.11	0.51

			HRQL				General health			
			MD*	**Upper CI**	**Lower CI**	**P-value**	**MD***	**Upper CI**	**Lower CI**	**P-value**
	Adults neighbour	Decline	0.14	-0.07	0.35	0.19	0.33	0.12	0.53	0.00
		Improved	-0.28	-0.49	-0.07	0.01	-0.16	-0.37	0.05	0.14
	Adults home	Decline	-0.20	-0.35	-0.05	0.01	-0.12	-0.27	0.02	0.10
		Improved	0.05	-0.10	0.20	0.52	-0.08	-0.23	0.08	0.33
	Peer belonging	Decline	-0.25	-0.45	-0.05	0.01	-0.27	-0.47	-0.08	0.01
		Improved	0.04	-0.16	0.24	0.68	0.18	-0.02	0.38	0.07
	Friendship intimacy	Decline	-0.04	-0.25	0.16	0.68	-0.17	-0.37	0.03	0.10
		Improved	-0.04	-0.25	0.17	0.71	0.17	-0.04	0.37	0.12
School experiences	Academic self concept	Decline	-0.01	-0.20	0.19	0.95	-0.06	-0.25	0.12	0.51
		Improved	0.07	-0.12	0.26	0.47	0.30	0.10	0.49	0.00
	School support	Decline	-0.12	-0.33	0.09	0.25	-0.20	-0.40	0.00	0.06
		Improved	0.14	-0.06	0.35	0.18	0.20	-0.01	0.41	0.06
	Bullying	Decline	0.15	-0.07	0.38	0.19	0.19	-0.04	0.41	0.11
		Improved	-0.09	-0.32	0.14	0.45	-0.07	-0.30	0.16	0.57
	School belonging	Decline	-0.22	-0.45	0.01	0.06	-0.18	-0.40	0.05	0.12
		Improved	0.08	-0.15	0.31	0.48	0.40	0.17	0.63	0.00
Wellbeing	Wellbeing sum	Decline	-4.55	-6.60	-2.51	0.00				

	HRQL				General health			
	MD*	**Upper CI**	**Lower CI**	**P-value**	**MD***	**Upper CI**	**Lower CI**	**P-value**
Improved	3.31	1.26	5.36	0.00				

* Mean difference compared to no change MD – Mean difference CI – Confidence interval

Table 18 Effect sizes of change between baseline and follow-up

		HRQL			General health		
MDI domain	**MDI scale**	**Decline**	**No change**	**Improve**	**Worsen**	**No change**	**Improve**
Social and emotional development	Empathy	0.01	0.06	0.11	-0.15	0.15	0.06
	Prosocial	-0.18	-0.15	-0.17	-0.26	-0.14	-0.12
	Optimism	-0.27	-0.03	0.27	-0.2	0.01	0.12
	Self-esteem	-0.23	0.09	0.19	-0.2	0.06	0.17
	SWLS-C	-0.22	0.14	0.29	-0.1	0.11	0.23
	Sadness	-0.25	0.09	0.38	-0.17	0.13	0.17
	Anxiety	-0.06	-0.17	-0.31	-0.08	-0.22	-0.19
Connectedness	Adults school	0.08	-0.05	-0.1	0.18	-0.07	-0.16
	Adults neighbour	0.09	-0.01	-0.31	0.22	-0.09	-0.29
	Adults home	-0.09	0.17	0.22	0	0.17	0.09
	Peer belonging	-0.14	0.11	0.15	-0.2	0.08	0.26
	Friendship intimacy	0.01	0.05	0.04	-0.15	0.04	0.23
School experiences	Academic self-concept	0.02	0.02	0.12	-0.09	-0.02	0.28
	School support	-0.31	-0.19	0	-0.38	-0.18	0.06
	Bullying	-0.05	-0.21	-0.25	-0.04	-0.21	-0.24
	School belonging	-0.22	-0.04	0.06	-0.29	-0.12	0.25
	Wellbeing	-0.32	0.1	0.4	N/A	N/A	N/A

N/A as General health measure contributes to the Wellbeing score

6.4 Discussion

6.4.1 Summary of principal findings

These analyses found that several scales comprising the MDI are valid and reliable measures of optimism, satisfaction with life, self-esteem and overall wellbeing in children aged between 7 and 10. Internal consistency reliability was comparable or exceeded results of previous studies (Castelli et al., 2018; Schonert-Reichl et al., 2012) and the MDI could discriminate between children with different levels of general health and HRQL. Construct validity was supported when replicating previously reported within MDI scale associations, although hypothesised associations between many MDI scales and HRQL domains were not confirmed. Nevertheless, scales reflecting domains of social and emotional development and overall wellbeing were associated with HRQL. Finally, as children appeared to already be at a high level of wellbeing and of good health, the MDI was predominantly responsive to declines in HRQL and general health, but not improvement.

6.4.2 Strengths and limitations

This was the first use of the MDI in the UK and the first to assess the responsiveness of the MDI to change in health and HRQL. Analyses that replicated previously published development work (Castelli et al., 2018; Gregory et al., 2018; Schonert-Reichl et al., 2012) largely aligned with reported findings. The non-selective inclusion criteria and diverse sample suggest that the findings are applicable to a population of healthy UK children. The sample was also large, which allowed robust analyses of between group differences. The sample included children in two school years. Collecting data from children in years 3 and 5 allows the results to be applied to both younger and older primary school children. The CHU9D has been validated in children as young as seven (Stevens, 2012), which is comparable to this study. Moreover, the CHU9D is a widely used and validated outcome measure, meaning the examination of convergent validity used a robust and appropriate measure.

A strength is the large dataset used in these analyses. There was however large attrition of 50% between baseline and follow-up, reducing the data available to evaluate responsiveness. The data utilised were collected as secondary outcomes in a large RCT, so this may be related to the trial design and independent of collection of the MDI data. Research staff measured anthropometric outcomes (BMIz and body fat), whereas school staff administered the MDI and supervised the linear track test. The amount of missing linear track test data was comparable to the MDI, which could suggest the reason for missingness was related. Both required a notable time commitment, which considering the findings in Chapter 5, could be explained by time limitations. While motivation and investment in the trial may have been high at baseline, this could have depleted over the study period. There were also reports that the tablets used to collect MDI data timed out and deleted data after a period of inactivity. This meant that if completion was paused due to breaktime, the data was lost. Staff may have had a better understanding of the time to allow for completion at follow-up and planned it around the school day but were unaware of this issue. This could probably be resolved simply by altering the software. If the MDI was to be used in a future trial, the potential reasons for attrition should be explored prior to baseline data collection and solutions identified or its inclusion reconsidered. This should involve collaboration between those expected to administer the measure (e.g. researchers or teachers) and those completing the measure (children). Solutions could include researchers administering the measure outside of a lesson, only completing the minimum necessary questions or splitting up completion across a school day.

There were no item-level missing data, but this was due to the method of data collection. Being unable to proceed without a response might have resulted in selections that were not actually endorsed by the children and perhaps not fully understood. While it is well documented that missing data is lower using electronic data capture in children (Mangunkusumo et al., 2006; Young et al., 2009), the impact of forced choice on response processing is unexplored (Bowling, 2005; Tourangeau et al., 2000). It was beyond the scope of this study to assess comprehension of MDI items, and similar issues might have occurred using optional responses and paper completion. There is no indication that younger children's data is of lesser quality, so it could be assumed that it is equally feasible to use the MDI in both age groups.

The analyses reported were generated from a study evaluating an intervention, and thus the measures against which the MDI could be assessed were determined by that broader study, rather than as the result of a study designed specifically to evaluate the MDI. Whilst several measures collected in the RCT trial enabled the assessment of MDI psychometric properties, a study developed entirely to assess the psychometric properties of the MDI could have been strengthened by the inclusion of additional measures. Firstly, collecting a second measure of wellbeing such as The Children's Society Childhood Index (Rees et al., 2010) would have enhanced the evaluation of construct validity. The difficulty in defining wellbeing and the absence of a gold standard measure of children's wellbeing mean that simply adding another questionnaire would not entirely resolve this issue but it would have been informative in terms of exploring alternative conceptualisations of wellbeing. Furthermore, the MDI scales measure a number of constructs that were identified from the literature as contributing to wellbeing. None of the ‘core’ scales provides a wellbeing score as such. The ideal approach would be to examine the relationships between the MDI scales and measures of the scale construct, e.g. the Anxiety scale and a measure of child anxiety. It is unlikely that gold standard measures exist for all scales and administration would have a huge response burden. Gold standard measures may also be proxy completed or developed for clinical diagnoses.

The inclusion of an anchor measure of perceived change in wellbeing over 12 months could have supplemented the assessment of responsiveness. As previously mentioned, wellbeing is difficult to define and a debated concept. Asking children to report wellbeing change could be subject to many issues including a lack of understanding of the concept, different conceptualisations of wellbeing and difficulties recalling their previous state of wellbeing. This is not an issue with the MDI itself as none of the MDI scales directly refer to wellbeing. Including anchor questions related to each scale would further increase the response burden.

Many limitations of the study are related to the number of scales and length of the MDI. The Wellbeing sum score was intended to generate a summary score of wellbeing that could be feasibly used to assess change in wellbeing in a trial setting, acknowledging that 15 scale scores would be unacceptable as primary outcomes. The multiple testing would be highly likely to find some false positive and false negative findings. It could be that one MDI scale is prospectively identified as being particularly related to the hypothesised intervention effect, therefore it could be preselected as the primary outcome. The MDI sum score appeared to perform well psychometrically, so could also be chosen as a primary outcome. This does not overcome the administrative burden however, with the sum score comprising items from several MDI scales. The scale has also been validated in its complete form so administrating scales in isolation would not be advisable without additional validation work. The SWLS-C scale would be an exception however as this was an existing, validated measure.

In order to measure intervention effect using the MDI sum score or any MDI scale score, some conceptualisation of meaningful change or difference should be established (McGlothlin and Lewis, 2014). The importance of a one-point improvement in relationships with adults in the neighbourhood is not known for example. As this is the first use of the MDI in a trial setting, the study did not aim to determine meaningful change or between group differences. Interpreting differences between the known groups and responsiveness were perhaps hindered by the absence of such criteria as it is questionable to rely upon statistical testing only to determine true differences.

If the MDI were to be used in a CEA, the incremental cost of one-point improvement in wellbeing or sadness could be calculated. Decision makers' WTP for such a one-point change would be unknown (Eichler et al., 2004). As the MDI currently has no potential to be used in a CUA, this challenge can be likened to most non-preference-based measures (McIntosh, 2019). In a UK context, thresholds are predominantly pertinent for NICE decision making, although the decision maker may be a school's headteacher or local authority commissioners in the context of a school based public health intervention. Furthermore, NICE do propose that CCA and CBA are appropriate frameworks for presenting economic evidence of public health interventions (NICE, 2014). MDI results could therefore be presented alongside costs and other outcome data in a CCA. Establishing a WTP for wellbeing gain from children for use in a CBA would be an interesting, yet challenging endeavour.

6.4.3 Findings in context

The internal consistency results are similar to those observed in Australia (Gregory et al., 2018) and compare favourably to analyses in Italian/Swiss (Castelli et al., 2018) and Canadian (Schonert-Reichl et al., 2012) samples. Gregory et al. (2018) also found the Sadness scale to have low reliability, whilst Castelli et al. (2018) found the reliability of seven MDI scales to be poor (Empathy, Optimism, Self-esteem, Sadness, Adult relationships at home, Academic self-concept and Bullying). In the first MDI validation study, Schonert-Reichl et al. (2012) reported only three scales with poor internal consistency (Empathy, Optimism, Adult relationships at home). A recurring finding appears to be the poor internal consistency of the Sadness and Empathy and Optimism scales. These scales reside within the SED domain, which had superior performance compared to other domains in respect to responsiveness and known groups validity. The observation that, in Australian children, internal consistency across all the MDI scales was generally higher for older children and females (Gregory et al., 2018) was not replicated, suggesting that the MDI performs more consistently across population groups in the UK.

The analyses found clustering of scores at the highest level on several scales. MDI ceiling effects do not seem unusual however. Whilst only highlighted in one previous publication (Gregory et al., 2018), median scores reported in the remaining validation publications do not appear markedly different to those observed in the current study, with medians approaching the highest possible score on most scales (Castelli et al., 2018; Schonert-Reichl et al., 2012). This could be due to the samples comprising healthy and flourishing children. Generic HRQL measures are designed to be applicable to whole population groups, therefore it would be anticipated that ‘well’ individuals would score highly (Upton et al., 2005). The importance or value of improving wellbeing of children who already appear to be flourishing, and thus the necessity of measuring ‘improvement’ in these children is a difficult judgement for decision makers. Should interventions be targeted at those with lowest wellbeing, and should improving their wellbeing be valued the same as those that are approaching a good level of wellbeing (Layard, 2016)? As scores differed significantly between levels of self-reported health and HRQL, this would suggest that more variability could be observed in a less ‘well’ population. Notable ceiling effects are observed in CHU9D domains (Petersen et al., 2018) and PedsQL scores (Varni et al., 2007) in general population samples for example. Some MDI scales showed responsiveness to change, despite many demonstrating large ceiling effects at baseline. This perhaps explains the finding that scales were predominantly responsive to declines in HRQL and health in this predominantly healthy population. Reviews of paediatric outcome measures have noted the lack of published evaluations of responsiveness (Janssens et al., 2015; Solans et al., 2008). This study therefore contributes to this literature, but whether similar findings have been observed in other populations is unknown.

It was hypothesised that the MDI would be able to discriminate between different groups of children previously shown to differ in their wellbeing. Whilst most hypothesised differences were confirmed, ethnicity and deprivation were not. This may be due to deprivation being measured at the school, not individual level. Petersen et al. (2018) for example tested known groups validity of the CHU9D and PedsQL (a non-preference based measure of children’s HRQL) in Australian adolescents. Groups were self-rated general health and family affluence, with between group differences for gender also tested. As with the Daily Mile study, worse health was associated with worse wellbeing/quality of life, however affluence was also found to be related. In contrast to the evaluation of the MDI, affluence was measured at an individual level instead of school level. While previous research has found ethnic differences in CHU9D scores in a similar population (Adab et al., 2018a) ethnic groups were examined at a more detailed level than white/non-white. Exploring ethnic differences using a less arbitrary distinction might yield different results, albeit with less statistical power.

6.4.4 Recommendations for future research

Future research could examine the test-retest reliability of the MDI. Test-retest is usually measured over a short time frame, to explore whether MDI score is stable over time when no change in health or wellbeing has occurred. Children's wellbeing may be transient, impacted by what adults might perceive as minor incidents during a school day. The stability of the MDI across a short period could therefore be variable. Establishing meaningful differences in MDI scale scores is important if the MDI were to be used to evaluate the impact of an intervention, and in theory, this would be required for each MDI scale. Potential ways forward could be to use consensus methods to decide which MDI scores are the most suitable for use in economic evaluation. It is likely that preferences would be context or intervention dependent, for example the SWLS-C might be deemed the most suitable outcome in trials of mental health interventions. Individuals consulted could include health economists, clinicians, teachers, children and parents. Once this is established, individual's willingness to pay for this meaningful change could be established, enabling a CBA to be conducted. Another way forward could be to develop a preference-based utility scoring algorithm for the MDI, as has been done previously for the SF-12/SF-6D (Brazier and Roberts, 2004), Short-Bowel Syndrome Quality of Life Questionnaire (SBS-QoL) (Lloyd et al., 2014) and Aberrant Behaviour Checklist (ABC-C) (Kerr et al., 2015).The correlation between the CHU9D and wellbeing sum score was not large, suggesting they are not equivalent outcomes and administering the longer MDI would not be futile given the already preference-based composition of the CHU9D, although society's willingness to pay for a change in wellbeing would still need to be established.

6.4.5 Recommendations for use of the MDI in economic evaluation

Possible ways in which the MDI could be used in economic evaluation have been raised in this chapter, with many requiring additional research before their implementation. Table 19 summarises possible approaches available currently and potentially in the future, alongside the potential advantages and disadvantages. Due to the more widespread use of the SWLS-C this is the only MDI subscale presented, but most of the same limitations associated with a shorter scale apply. At present, recommended analyses would comprise either a CEA or CCA using the MDI or the SWLS-C. The MDI is more comprehensive and in the absence of any consensus on the most suitable subscale, using the Wellbeing Sum Score is the most appropriate approach.

Table 19 Potential uses of the MDI in economic evaluation

Scale	Type of economic evaluation	Proposed use	Advantages	Disadvantages
MDI Wellbeing sum score	CEA	Cost per unit improvement in wellbeing	Scale is suitable for statistical analyses	Response burden. The majority of MDI scales must be administered
			Covers a range of contributors to children's wellbeing	No threshold to determine cost effectiveness and current absence of studies for comparison
				Meaning of a unit of wellbeing unknown
				Wellbeing sum score has not been used previously
	CCA	A comparison of scores between study arms presented alongside other outcome data	Results can be evaluated by the decision maker in the context of other outcome data. This could be suited to a complex intervention	Response burden. The majority of MDI scales must be administered
			No need for a cost-effectiveness threshold	Decisions regarding efficiency cannot be made

Scale	Type of economic evaluation	Proposed use	Advantages	Disadvantages
MDI wellbeing index	CEA	Cost to move from low wellbeing to thriving	Potentially more meaningful than a unit increase in wellbeing	Response burden. The majority of MDI scales must be administered
			Covers a range of contributors to children's wellbeing	No threshold to determine cost effectiveness and current absence of studies for comparison
	CCA	A comparison of proportions moving between categories presented alongside other outcome data	Results can be evaluated by the decision maker in the context of other outcome data. This could be suited to a complex intervention	Response burden. The majority of MDI scales must be administered
			Results could be compared with interventions in other sectors	Decisions regarding efficiency cannot be made
			No need for a cost-effectiveness threshold	
SWLS-C	CEA	Cost per unit increase in SWLS-C	SWLS-C is an accepted measure of satisfaction with life	Measure not as comprehensive as the full MDI
			Lower response burden compared to full MDI	No threshold to determine cost effectiveness

Scale	Type of economic evaluation	Proposed use	Advantages	Disadvantages
				The short scale means that children with already high wellbeing may be unable to improve
	CCA	A comparison of scores between study arms presented alongside other outcome data	Results can be evaluated by the decision maker in the context of other outcome data. This could be suited to a complex intervention	Measure not as comprehensive as the full MDI
			Lower response burden compared to full MDI	Decisions regarding efficiency cannot be made
Potential use following additional work				
MDI Wellbeing Sum Score MDI wellbeing index SWLS-C	CBA	Cost-benefit ratio presented, using a willingness to pay value for change in MDI outcome	Results could be compared with interventions in other sectors	Potential difficulty in putting a monetary value on change in wellbeing and obtaining WTP from children
			Interpretation is simpler than CUA for non-specialists and no need for a cost-effectiveness threshold	Response burden. The majority of MDI scales must be administered (excluding SWLS-C)

Scale	Type of economic evaluation	Proposed use	Advantages	Disadvantages
MDI Wellbeing Sum Score MDI wellbeing index SWLS-C	ROI SROI	Percentage return on investment presented, including MDI	Results could be compared with interventions in other sectors (only if ROI)	Potential difficulty in putting a monetary value on change in wellbeing and obtaining WTP from children
			Interpretation is simpler than CUA for non-specialists and no need for a cost-effectiveness threshold	Response burden. The majority of MDI scales must be administered (excluding SWLS-C)
MDI Wellbeing Sum Score SWLS-C	CUA	Cost per wellbeing adjusted life year (WALY), with utilities generated from the MDI using a preference-based scoring algorithm	Accepted outcome for economic evaluation	Response burden. The majority of MDI scales must be administered (excluding SWLS-C)
				Potential difficulty obtaining health state valuations from children
				No established cost per WALY threshold for evaluating cost-effectiveness

Scale	Type of economic evaluation	Proposed use	Advantages	Disadvantages
MDI Wellbeing Sum Score MDI wellbeing index SWLS-C	CEA	Cost per unit change in wellbeing, using an established WTP threshold and in with an understanding of what a meaningful change in score is	The MDI Wellbeing Sum Score is suitable for statistical analyses The response burden for the SWLS-C is low	Response burden. The majority of MDI scales must be administered (excluding SWLS-C)
				Resource required to estimate a meaningful change in wellbeing and generate a cost-effectiveness threshold

6.5 **Conclusion**

This chapter reports results that indicate that the MDI is a valid measure of wellbeing in a UK sample of children in years 3 and 5. MDI scales showed some association with the commonly used CHU9D, although the strength of the association and broader conceptual coverage suggests that results from the MDI would not replicate what could be obtained using this more brief measure of HRQL. Findings from the qualitative interviews reported in Chapter 5 suggested that a schools' perspective in an economic evaluation could include wellbeing as an outcome, as opposed to a QALY or unit of BMI. Chapter 7 reports an exploratory economic evaluation of school-based public health intervention, which broadens the perspective to include wellbeing.

From the candidate 15 scales, the MDI Wellbeing Sum Score was chosen as the outcome for the CEA. While the qualitative interviews did not explore participant's conceptualisations of wellbeing nor which MDI scales were particularly pertinent, the analyses conducted in this chapter and composition of the scales would suggest this is an appropriate approach. As the individual scales are short, a small variation in scores was observed for many. Furthermore, many scales showed ceiling effects. The majority of participants performing at the same level would not be amenable to demonstrating an incremental benefit between interventions. A composite score of several scales measuring constructs known to contribute to wellbeing, in addition to a child's assessment of their own health should provide an adequate indicator with variation between individuals. The Self-esteem, Sadness and Optimism scales were responsive to change in general health and quality of life, suggesting the overall score would also be responsive.

While the perspective of the economic evaluation can be broadened by altering the measure of benefit, the cost perspective could also be expanded. When exploring costs and prioritisation of public health interventions in schools, staff were aware that there is an opportunity cost to providing them in terms of children's time. A monetary value cannot currently be attached to that time, with potential implications including reduced academic attainment or less opportunity for pastoral activities. The economic evaluation reported in Chapter 7 which broadens the perspective to the school also attempts to address this opportunity cost, in addition to wellbeing. The intention of this approach is to provide an analyses which is useful to the school decision maker, and relevant to those implementing the intervention.

CHAPTER 7 CONDUCTING AN ECONOMIC EVALUATION FROM THE SCHOOLS' PERSPECTIVE

7.1 Introduction

This chapter reports an exploratory economic evaluation of the Daily Mile, conducted alongside the Birmingham Daily Mile cluster RCT. The objective of this analysis was to assess the cost-effectiveness of the Daily Mile compared to usual school health and wellbeing activities using a school perspective, with a comparison against the public sector perspective. This economic evaluation builds on the findings reported in Chapters 3, 5 and 6. Interviews reported in Chapter 5 revealed that the impact on wellbeing was thought to be an important outcome by school staff when considering public health interventions. A cost-per QALY or unit of BMI were not outcomes that resonated with those implementing the intervention. The opportunity cost of the supposedly free Daily Mile was also highlighted. In Chapter 6, it was shown that the MDI had good psychometric performance in a UK population. The trial methods and clinical effectiveness results were described in Chapter 4 and a detailed trial protocol has been published (Breheny et al., 2018).

The school perspective was informed by findings of qualitative interviews reported in Chapter 5. The primary decision makers (headteachers) were eager for evidence of effectiveness and cost-effectiveness to support decision making, although were unsure whether methods used in healthcare would be informative. Features of the newly developed school perspective compared to a public sector perspective, determined from the interviews reported in Chapter 5 are:

- The irrelevance of the QALY and preference for wellbeing as an alternative outcome
- A lack of preference-based measures of children's wellbeing means a CEA approach is necessary, instead of CUA
- The opportunity cost of providing the public health intervention is important to schools, therefore should be considered in the analysis
- Individual characteristics such as gender can have significant impacts on the implementation of interventions, therefore a simple equity analysis could be informative
- The impact on attitudes and behaviours (e.g. enthusiasm, long-term participation) may be more indicative of benefit than anthropometric measures. While measures such as BMI may decrease initially, enduring impacts may depend on changes in attitudes towards health and wellbeing and voluntary participation.

The schools' perspective, as described above, would appear to be a novel approach, at least within the context of interventions to prevent obesity or risk factors for obesity. The systematic review in Chapter 3 identified 38 analyses, although none would be consistent with the school perspective. No analyses included wellbeing as an outcome measure, which is perhaps understandable given the paucity of wellbeing measures available for children, discussed in Section 6.1. Measuring the impact on long-term behaviour change and attitudes to physical activity were also not included in any analyses. This could be due to the challenges of measuring these outcomes over a clinical trial's duration, its recognition as an important outcome by public health decision makers or being perceived as inappropriate for demonstrating immediate benefit and value. These outcomes remain an area for further research.

One aspect of the schools' perspective that was addressed in a number of economic evaluations identified in Chapter 3 was equity. Equity is an important consideration in public health (Marmot, 2010; Marmot et al., 2008; WHO, 2012) and how to incorporate it in economic evaluation is an area of methodological debate (Cookson et al., 2017). Indeed, it is noted as an issue in economic evaluations of public health interventions, discussed in Weatherly et al. (2009). Some studies conducted a basic equity analysis (Cookson et al., 2017), reporting findings by deprivation subgroup. No school staff interviewed identified differing efficacy by ethnicity or deprivation status as considerations when prioritising interventions, although gender differences in suitability and implementation fidelity were raised. Implementing an initiative with known benefits for one gender only could potentially widen inequities. Two economic evaluations identified in the systematic review reported gender differences in effect, with Wang et al. (2003) basing the economic analysis on females only and te Velde et al. (2011) reporting a beneficial effect in females (although this study did not report separate ICERs). This would suggest that the equity impacts of these findings were overlooked.

In an economic evaluation, all the relevant costs and outcomes should be included. If an analysis takes a societal perspective, time consuming healthcare (e.g. receiving treatment) is considered a cost. Time participating in a public health intervention also has an opportunity cost, which by extension could also be included. In the case of a school-based intervention this could include lost education time or playtime for example. This was highlighted in the interviews reported in Chapter 5 and including children's time in economic evaluation was the subject of a recent review (Andronis et al., 2019). No studies in the systematic review reported in Chapter 3 addressed the opportunity cost of providing the intervention.

Valuing children's time for consideration in economic evaluation is an area of current research interest and a recent publication (Andronis et al., 2019) noted the current paucity of methods and discussion of the issue, in contrast to adults. The question was approached in the context of individual treatment of ill health and not population level health interventions, which this thesis has addressed. Andronis et al. (2019) searched the Paediatric Economic Evaluation Database using terms related to school absence and leisure time, finding seven relevant studies. All were evaluations of medical interventions and only lost education was quantified, not leisure time. No studies costed time receiving treatment, with only school absence due to ill health accounted for. Three methods were used to estimate lost education. The first assumed that the school receives payment for attendance, therefore the cost of absence to the school was included. A second approach used the daily wage rate of the primary caregiver and the third used an association between absence and educational attainment, with an estimated impact on future earnings. Andronis et al. (2019) provide a taxonomy of considerations when valuing children's lost time, which included the challenges of measurement, its relevance to the chosen perspective and whether it would have any impact on welfare loss. What remains is consensus about how this time should be measured and whether its inclusion is appropriate. With no value currently available, a pragmatic approach would be required to conduct an economic evaluation from the schools' perspective that includes this value.

Evidence would suggest that a school perspective, as described previously, has yet to be recognised and implemented at least in the topic area systematically reviewed in Chapter 3. This economic evaluation uses the available data from

The Birmingham Daily Mile trial to conduct an exploratory economic evaluation from the schools' perspective, with a comparison with a public sector perspective, which used a CUA approach. Wellbeing was collected using the MDI and CHU9D data used to construct QALYs. The school perspective also included children's leisure time, to reflect the opportunity cost of participating in the Daily Mile. Chapter 4 described the Birmingham Daily Mile trial and Chapter 6 described the MDI and CHU9D measures in depth. This chapter describes the methods used to conduct the economic evaluation and the results. The analysis explicitly considers the preferences of school staff through qualitative interviews as guidance for methods for conducting economic evaluations of public health interventions in school settings.

7.2 **Methods**

As a case study to explore the schools' perspective in economic evaluation, the cost-effectiveness of the Daily Mile compared to usual activities was assessed from school and public sector perspectives. The base-case analysis took the school perspective. The time horizon was one school year, reflecting the duration of the Birmingham Daily Mile trial. Discounting was therefore not conducted on either costs or outcomes. The Daily Mile intervention was described in Chapter 4. The comparator was usual school health and wellbeing activities, with schools requested not to introduce any new initiatives over the course of the year.

The school perspective took the form of a CEA, using wellbeing (MDI) as the outcome. This choice of CEA was driven by the need for wellbeing as an outcome and the school decision maker's interest in the cost per unit of effect. The absence of a preference-based measure of wellbeing meant that a CUA was not feasible and CCA would not satisfy their requirements. Furthermore, CBA is not possible without a monetary value of wellbeing. A CUA from a public sector perspective was chosen as the comparison due to the preference of NICE (NICE, 2014) for this approach and the existence of a QALY WTP threshold which allows an evaluation of cost-effectiveness.

7.2.1 Perspectives

The analysis was undertaken from both a school and public sector perspective. Table 20 provides a summary of what each perspective comprises, and the sensitivity analyses conducted. The school perspective is populated using the most appropriate data available from the trial, therefore is not a comprehensive demonstration of this new approach. The notable differences are that the school perspective includes the opportunity cost of children's education time and uses wellbeing as the outcome, instead of the QALY. The QALY is NICE's preferred outcome (NICE, 2014) and the costs of teacher time reflect the cost to the education sector of providing the intervention. The costs and outcomes are described in the sections that follow.

Table 20 Summary of the school and public sector perspectives and sensitivity analysis conducted

		Perspective	
		School	Public sector
Inputs	Costs	Teacher time Opportunity cost of lost classroom time	Teacher time
	Outcomes	Wellbeing (MDI)	QALYs
Sensitivity analysis	Subgroup	Gender	Gender
	Costs	No opportunity cost included	Children's opportunity cost included
	Outcomes	Alternative measure of wellbeing (SWLS-C)	

7.2.2 Participants

The trial inclusion and exclusion criteria were described in Chapter 4. Briefly, the sample comprised children in years 3 and 5 (aged 7-8 and 9-10) attending schools in Birmingham. As noted in Chapter 6, there was a large amount of missing data for the CHU9D. The clinical effectiveness dataset comprised 2,280 children, whereas the number included in the economic evaluation was much fewer for this reason.

7.2.3 Outcomes

The analyses are based on children's self-reported HRQL and wellbeing. Data were collected using the CHU9D (HRQL) and MDI (wellbeing) measures, described in detail in Chapter 6 (Section 6.2.2). The MDI produces a number of different scores, and for this analysis the Wellbeing Sum Score was chosen as a summary measure of wellbeing. This score is referred to as MDI for the purpose of the analyses in this chapter. This also performed well in the analyses reported in Chapter 6. The SWLS-C was used in a sensitivity analysis. This is a validated measure of life satisfaction, nested within the MDI. CHU9D data were used to generate utilities and calculate QALYs, for which the area under the curve method was used (Glick et al., 2014).

7.2.4 Costs and resource use

It was expected that unanticipated costs of the ‘free’ Daily Mile that arose in the qualitative interviews would be included in the analyses. However, with the exception of one £10,000 running track, these were viewed to be minor (e.g. printing certificates) and a normative judgement was made that they were unlikely to have an effect on the cost-effectiveness. For completeness, the costs identified are listed in Table 21.

Table 21 Resource use identified in qualitative interviews

Resource
Cubes to count playground laps
Enhanced first aid kit
Risk assessment
Entering Daily Mile data
Classroom displays
Certificates

Therefore, the costs included in the economic evaluation were the teacher time to conduct the Daily Mile and children's lost classroom time. Costs reported are the average cost to conduct the Daily Mile for one child for one year. Local authority primary school teacher's salary was calculated over a standard 190-day school year, providing a cost per child, assuming an average primary school class size of 27 children. Costs were estimated based on how often the Daily Mile was conducted (e.g. 5 days a week) and for how long (e.g. 15 minutes), for each school based on interviews with school staff. Where data were unknown, perfect implementation was assumed.

The cost of the comparator arm of usual activities was assumed to be zero. The cost of children's leisure time was £14.78 per hour (16 Euros, 2017 exchange rate) (Verbooy et al., 2018). This value of adult's leisure time was used as a proxy, given the absence of a figure for this at present (Andronis et al., 2019). Verbooy et al. (2018) estimated value of adult leisure time using contingent valuation. Whilst leisure time and lost education are quite different constructs, they are similar in that they are a value of time not involving undertaking employment and receiving a wage. The contingent valuation involved asking the Dutch general population to trade-off between performing 1 hour more of paid work per week and either 1 hour less unpaid work or 1 hour less of leisure time using a WTA payment scale approach. A WTP question examined the trade-off between unpaid work and leisure time. The payment scale ranged from €0-€50 (or higher) and individuals were asked what they would definitely accept/pay and definitely not accept/pay. If the maximum WTA exceeded €50 an open-ended question was asked. While this study provides an indication of the value of leisure time, it is was subject to the same limitations of WTP studies, such as WTP being associated with income and the hypothetical nature of the task. Furthermore, in relation to its application in this exploratory economic evaluation, the study was conducted with adults and it did not value education time explicitly.

Table 22 Costs and assumptions used in the analysis

	Detail/Assumption	Unit Cost (£)	Source
School teacher's time	Average annual salary Average hourly salary assuming 1265 hour working school year (Department for Education, 2017b)	34,300 27.11	(Department for Education, 2017c)
Children's time	Cost of children's lost education time using a proxy value of adult's leisure time	14.78	(Verbooy et al., 2018)
Teacher time	Cost of 15-minutes of teacher's time Average intervention cost per child (assuming average class size of 27 children and a 190 day school year (Department for Education, 1999))	6.77 47.53	

7.2.5 Analysis

All analyses were conducted in STATA (Version 13) (StataCorp, 2013). The difference in costs between arms were estimated using general linear mixed models (GLLAMM command), adjusted for variables used in randomisation, clustering (school), cluster means and demographic characteristics (age, gender, ethnicity,). Differences in QALYs and wellbeing were analysed using mixed multi-level models (MIXED command), adjusted for the aforementioned factors and baseline utility (QALYs) or wellbeing (MDI). The analysis used a modified intention to treat approach, where all available data were used and no data was imputed.

The differences in costs and effects were bootstrapped 1,000 times to produce confidence intervals for the ICERs and plot cost-effectiveness planes. 95% confidence intervals were reported only for analyses where there was a significant difference in effect, due to the unreliability of confidence intervals when the denominator of the ratio crosses the null value. The net benefit approach is an alternative to calculating ICERs, which does not require the calculation of a ratio value (Morris et al., 2007). The net benefit approach was also used to summarize the results of the economic analyses. In order to calculate the net health or monetary benefit of an intervention, the costs and effects must be in the same units. The net monetary benefit (NMB) was estimated in these analyses, requiring that QALY and wellbeing data were converted into costs. This requires a ceiling threshold, which is a value attached to a unit of effect. The incremental NMB was calculated for a range of WTP thresholds, which enabled the comparison of the Daily Mile and usual health and wellbeing activities. This can be plotted on a cost-effectiveness acceptability curve (CEAC) (Fenwick et al., 2004), which presents the probability that the Daily Mile is more cost-effective than usual activities at different WTP thresholds.

A simplified summary of the results from both perspectives was presented, which could be used to inform decision making. In addition to the economic results, it also includes a summary of the primary outcome of the Birmingham Daily Mile trial (BMIz) for the whole trial sample.

7.3 **<u>Results</u>**

The characteristics of the sample are as reported in Chapter 6, Table 9 (complete case sample). The mean age was 8.9 years and there were slightly more males (52.3%) than females (47.7%).

7.3.1 Cost-effectiveness of the Daily Mile from the schools' perspective

Table 23 presents the unadjusted mean costs and outcomes for the Daily Mile and usual activities and Table 24 presents the results of a CEA from the schools' perspective. The base case analysis showed a beneficial effect of the Daily Mile on wellbeing, with a mean difference of 0.72 MDI units. This difference was not significant (p=0.323). There was a significant difference in costs, with the Daily Mile costing an additional £738.74 per child per year, resulting in a base case ICER of £1,019.83 per MDI unit. This means that improving each child's wellbeing by one point would cost an additional £1,019.83 per child over a school year, compared to providing usual health and wellbeing activities. Figure 4 shows the distribution of 1,000 bootstrap replications of incremental costs and effects. The majority are in the upper right quadrant, suggesting that the Daily Mile is more costly and more effective. The CEAC (Figure 5) plots the probability that the Daily Mile is cost-effective at increasing WTP for a unit of MDI for the whole sample and by gender. At a WTP of £5,000, the probability that the Daily Mile is cost-effective compared to usual activities is 78%. At a threshold of £20,000, the probability would be 83% (result not shown on CEAC).

Table 23 Mean costs and effects of the Daily Mile and Usual Activities – School perspective

	Intervention			**Control**		
	Whole sample Mean (SD)	Boys Mean (SD)	Girls Mean (SD)	Whole sample Mean (SD)	Boys Mean (SD)	Girls Mean (SD)
Costs (£)	716.45 (89.37)	717.82 (87.74)	714.95 (91.30)	0	0	0
Wellbeing (MDI)	59.20 (11.21)	59.46 (11.25)	58.90 (10.68)	58.57 (10.97)	57.94 (11.45)	59.28 (10.95)

Deterministic sensitivity analysis is presented in Table 24. Analyses by gender produced a result for girls that was not consistent with the base case as girls suffered a small non-significant loss in wellbeing when compared to usual activities. This results in a negative ICER, which when interpreted alongside the bootstrapped costs and effects suggests that the Daily Mile is more costly and less effective than usual activities in this group, thus is dominated. The cost-effectiveness planes for girls and boys are presented in Appendix K, Figure 8 and Figure 9. The CEAC in Figure 5 shows that the Daily Mile is likely to be most cost-effective for boys from the schools' perspective, although there was no significant difference in effect and the results are very uncertain. Sensitivity analysis that excluded the opportunity cost of children's time reduced the ICER to £66.72 per MDI unit, which increased the probability that the Daily Mile is cost-effective, but it still remains the more costly option. To test the sensitivity of the findings to the measure of wellbeing, the SWLS-C was used as the outcome, instead of the MDI (Table 24). As with the MDI results, there was a small but non-significant difference in effect in favour of the Daily Mile. This resulted in an ICER of

£13,634.44 per SWLS-C unit. The different ranges of the MDI and SWLS-C mean the ICERS are not comparable in their current form.

Table 24 Cost-effectiveness results from the schools' perspective

		Costs (£)		Outcomes		ICER (£)
		MD (95% upper CI : 95% lower CI)	P-value	MD (95% upper CI : 95% lower CI)	P-value	
Base case [a]		738.74 (694.12: 783.37)	<0.001	0.72 (-0.71: 2.16)	0.323	1019.83
Sensitivity analysis	No opportunity cost [b]	48.33 (45.43: 51.23)	<0.001	0.72 (-0.71: 2.16)	0.323	66.72
	Boys [c]	740.61 (683.70: 797.51)	<0.001	1.61 (-0.47: 3.70)	0.130	458.78
	Girls [d]	726.52 (674.42: 778.61)	<0.001	-0.45 (-2.37: 1.46)	0.641	-1436.27
	SWLS-C [e]	738.74 (694.12: 783.37)	<0.001	0.05 (-0.08: 0.19)	0.431	13634.44

[a] Base case – MDI units, includes children's opportunity cost
[b] No opportunity cost – MDI units, teacher time only
[c] Boys – MDI units, includes children's opportunity cost
[d] Girls – MDI units, includes children's opportunity cost
[e] SWLS-C – SWLS-C units, includes children's opportunity cost

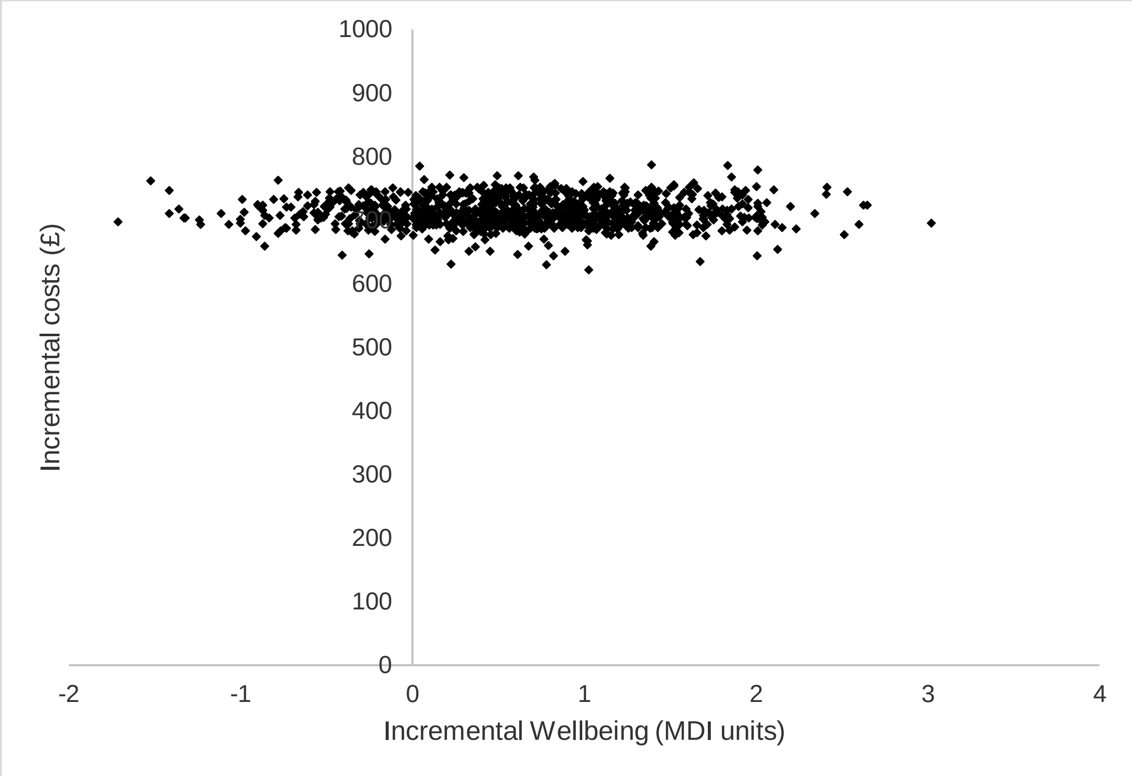

Figure 4 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and effects of the Daily Mile

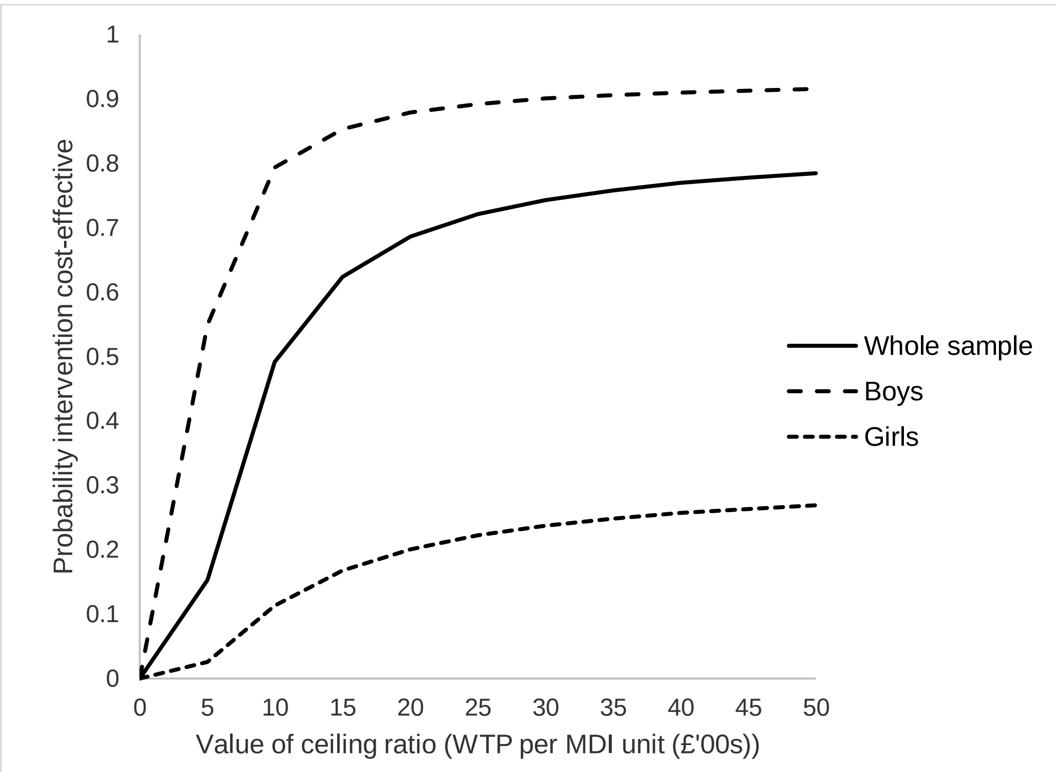

Figure 5 Cost-effectiveness acceptability curve showing the probability that the Daily Mile is cost-effective from the schools' perspective

7.3.2 Cost effectiveness of the Daily Mile from the public sector perspective[2]

Table 25 provides the unadjusted mean costs and effects for the whole sample and separately by gender, to enable a comparison with the school perspective. Table 26 presents the results of the CUA analysis from the public sector perspective. For the whole sample, the incremental cost per QALY was £9,559.52. There was no significant difference in QALYs for the whole sample or boys, although there was a significant difference for girls favouring the Daily Mile. There was a small but non-significant negative effect for boys, favouring usual activities. The Daily Mile was more costly than usual activities, costing an additional £48.33 per child per year. Figure 6 presents the cost-effectiveness plane for the public sector perspective. Like the schools' perspective, the majority of bootstrap replications indicate that the Daily Mile is more costly and more effective. Using the NICE threshold of £20,000 per incremental QALY (NICE, 2014), the probability that the Daily Mile is cost-effective compared to usual activities is 70% (Figure 7).

Sensitivity analysis indicates that the probability that the Daily Mile is cost-effective increases to 98% for girls, using a threshold of £20,000 per incremental QALY. If the opportunity cost of children's lost education time is included, the ICER for the whole sample (£146,121.70) exceeds the NICE threshold, indicating that the Daily Mile would not be cost-effective. The ICER for boys and cost-effectiveness plane (Appendix Figure 11) suggest

[2] The results reported from the public sector perspective differ to those submitted for publication in the article titled: *Effectiveness and cost-effectiveness of the Daily Mile on childhood weight outcomes and wellbeing: a cluster randomised controlled trial*. When subgroup analysis by gender was run from the schools' perspective (MDI outcome), the model would not converge for girls when the IMD covariate was included. To enable comparability between perspectives, the IMD covariate was not included in any analyses reported in this chapter.

that in this group the Daily Mile is dominated, i.e. usual activities are more effective and less costly than the Daily Mile.

Table 25 Mean costs and effects of the Daily Mile and Usual Activities - Public sector perspective

	Intervention			**Control**		
	Whole sample Mean (SD)	Boys Mean (SD)	Girls Mean (SD)	Whole sample Mean (SD)	Boys Mean (SD)	Girls Mean (SD)
Costs (£)	45.44 (5.67)	45.53 (5.56)	45.35 (5.79)	0	0	0
QALYs	0.836 (0.13)	0.833 (0.14)	0.840 (0.13)	0.839 (0.13)	0.850 (0.12)	0.826 (0.13)

Table 26 Cost-effectiveness results from the public sector perspective

		Costs (£)		Outcomes		ICER (£)
		MD (95% lower CI : 95% upper CI)	P-value	MD (95% lower CI : 95% upper CI)	P-value	
Public sector perspective [a]		48.33 (45.43: 51.23)	<0.001	0.005 (-0.005: 0.015)	0.324	9559.52
Sensitivity analysis	Opportunity cost [b]	738.74 (694.12: 783.37)	<0.001	0.005 (-0.005: 0.015)	0.324	146121.70
	Boys [c]	47.08 (44.54: 49.61)	<0.001	-0.009 (-0.022: 0.005)	0.208	-5451.30
	Girls [d]	47.86 (45.73: 49.99)	<0.001	0.018 (0.002: 0.033)	0.023	2671.43[e]

[a] Public sector perspective – QALYs, includes teacher time only
[b] Opportunity cost – QALYs, includes children's opportunity cost
[c] Boys – QALYs, includes teacher time only
[d] Girls – QALYs, includes teacher time only
[e] 95% CIs £1,383.64: £14,267.51
ICER – Incremental cost-effectiveness ratio MD – Mean difference CI – Confidence interval

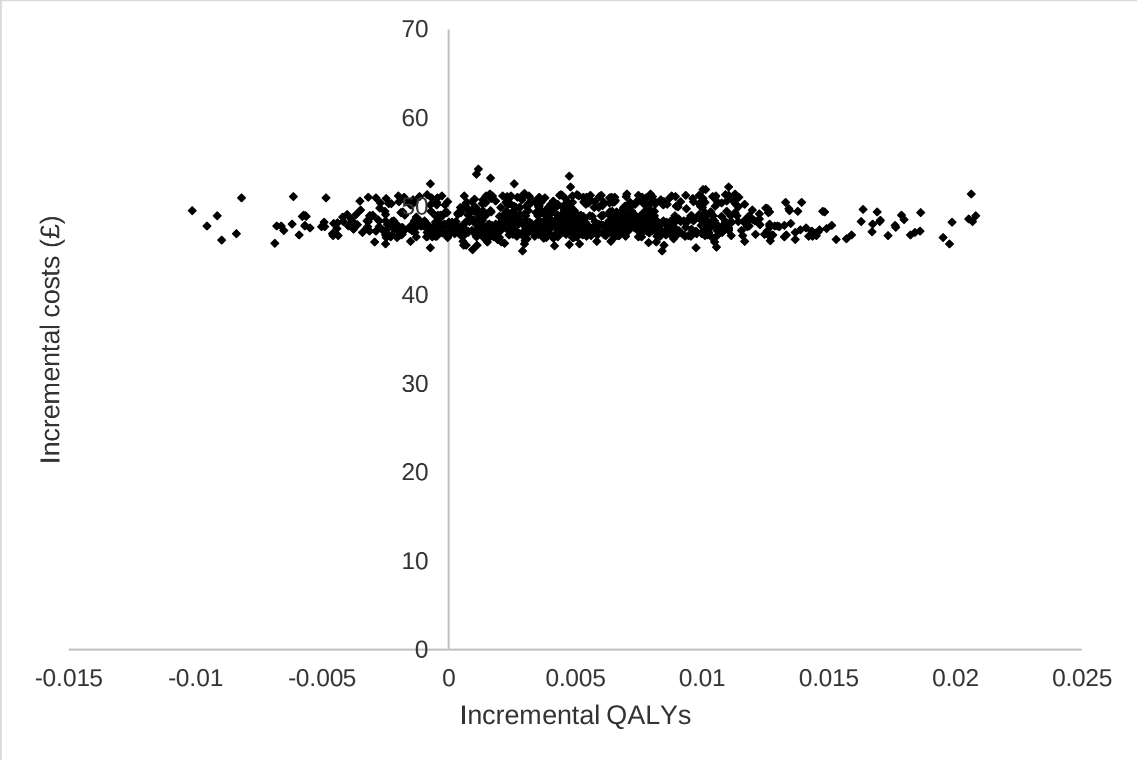

Figure 6 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and QALYs of the Daily Mile

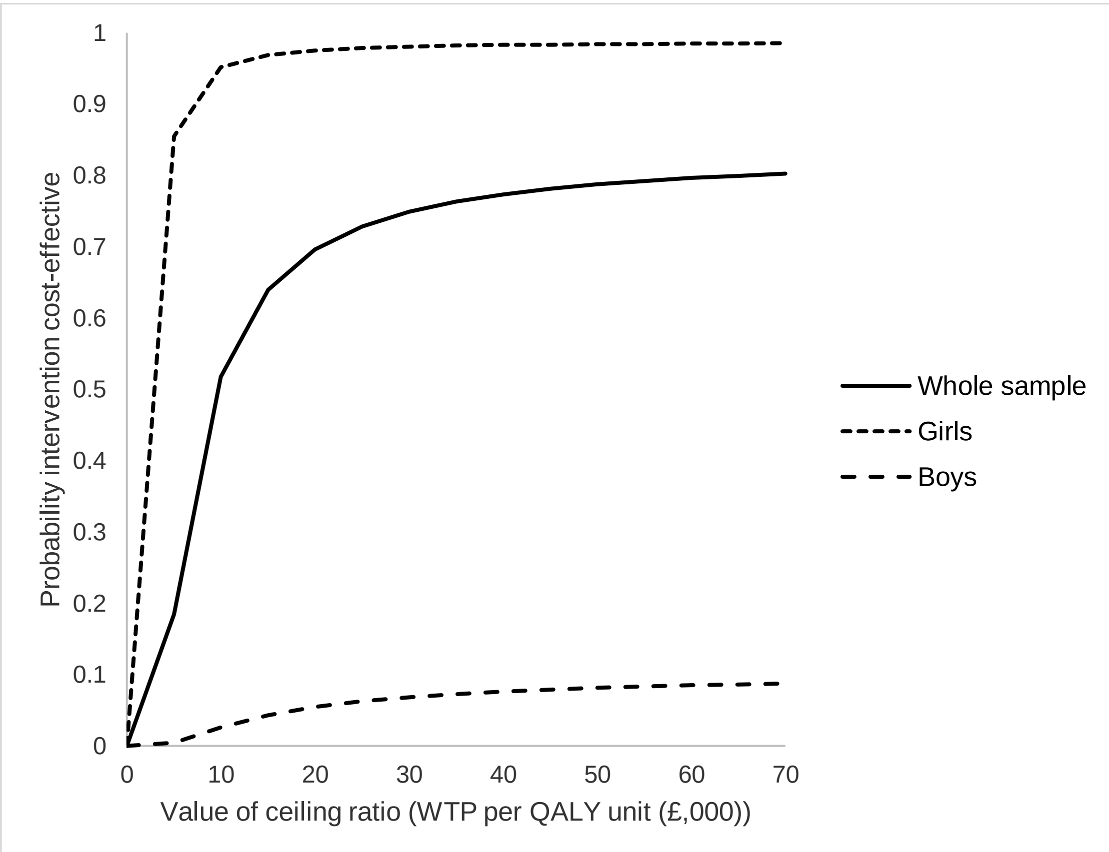

Figure 7 Cost-effectiveness acceptability curve showing the probability that the Daily Mile is cost-effective from the public sector perspective

7.3.3 Comparing the school and public sector perspectives

Table 27 presents a summary of the results from both the school and public sector perspectives. A WTP threshold of £20,000 per MDI unit has been used to communicate the probability that the Daily Mile is cost-effective from the schools' perspective. This is the same cost used for the public sector perspective, which instead reports QALYs. The school portion of the table could be presented to school decision makers to evaluate the value of implementing the Daily Mile. Key findings for consideration are:

- The Daily Mile is beneficial, but not statistically significantly more effective than usual activities for the whole sample, from either perspective
- From the school perspective, the Daily Mile is most likely to be cost-effective in boys
- From the public sector perspective, the Daily Mile is most likely to be cost-effective in girls
- Results from both perspectives would suggest that the Daily Mile is not harmful
- The differences in effectiveness from the public sector perspective (QALY) align with the clinical effectiveness results, with a significant benefit seen in girls only

Table 27 A comparison of the School and Public Sector Perspectives

	Perspective					
	School			Public sector		
	Overall (base case)	Boys	Girls	Overall	Boys	Girls
ICER [a] (£)	1019.83	458.78	-1436.27	9559.52	-5451.30	2671.43
Probability Daily Mile is cost-effective [b]	83%	93%	31%	70%	5%	98%
Difference in effect [cd]	Non-sig	Non-sig	Non-sig	Non-sig	Non-sig	Sig
Cost (£)	716.45	717.82	714.95	48.33	47.08	47.86
Difference in BMIz [d]	Non-sig	Non-sig	Sig	Non-sig	Non-sig	Sig

[a] School perspective- Incremental cost per MDI unit. Public sector perspective – Incremental cost per QALY
[b] At a WTP of £20,000 per MDI unit (school perspective) or QALY (public sector)
[c] School perspective - MDI Public sector perspective – QALY
[d] Non-sig – No significant difference at the 0.05 level. Sig – A significant difference at the 0.05 level, favouring the Daily Mile.

7.4 **Discussion**

7.4.1 Summary of principal findings

This chapter reports an economic evaluation of the Daily Mile, compared to schools' usual health and wellbeing activities, conducted alongside the Birmingham Daily Mile trial. The analyses draw on the findings from Chapters 4-6, presenting analysis from the schools' perspective and a public sector perspective for comparison. The base-case analysis from the schools' perspective found the incremental cost-effectiveness of the Daily Mile is £1,019.83 per MDI unit. The probability of the Daily Mile being cost-effective at a willingness to pay of £20,000 per MDI unit is 83%. Simple equity analysis by gender showed no significant difference in wellbeing for either group, although the Daily Mile was beneficial in boys and had a marginally negative effect in girls. These subgroup differences are contrary to those using the public sector perspective, whereby the Daily Mile was likely to be cost-effective in girls and not boys using the NICE threshold of £20,000 per QALY. It could be that school decision makers are willing to pay £458.78 for a unit of wellbeing in boys, although with no threshold or understanding of the meaning of a unit of MDI, this is speculative.

7.4.2 Results in context

As mentioned in previous chapters, reporting a CCA can be an appropriate method to summarise economic evaluation findings to inform decision making (NICE, 2014). A simple comparison of the perspectives has been presented, although it did not include all evidence obtained from the trial. Academic attainment was highlighted in the qualitative interviews as an important metric for measuring school and teacher performance and the summary of primary school accountability provided in Chapter 1 Section 1.4.9 supports this. It would have been interesting to present the difference in academic attainment alongside the schools' perspective, although there was a high amount of missing data for this outcome (72%). The results of the trial's primary outcome (BMIz) was presented alongside the summary of economic findings, although a CEA was not conducted on this measure. Despite school staff asserting that BMI differences were not of interest to them, presenting this analysis would allow decision makers to assess all of the available evidence from the trial prior to deciding whether to adopt the initiative. Although the opportunity cost of children's time was highlighted as important and thus included in the base case analysis, this cost as currently reported would not come out of school budgets. A sensitivity analysis without this value is therefore also presented for decision makers to consider.

It is interesting that the results from the school perspective align with the insights gained from the qualitative interviews reported in Chapter 5, in which several staff reported girls not participating fully in the Daily Mile, in contrast to boys. It would be anticipated that girls would receive less benefit if they did not engage with the initiative. The clinical effectiveness result affirms the findings of the public sector perspective, but not the school perspective. These findings must be interpreted with caution, however. The summary table does not indicate whether the observed difference in BMIz is clinically meaningful, the meaning of differences in MDI are also unknown and there is great uncertainty in the economic results. Presenting the results in a format that can be interpreted accurately by school decision makers, without a comprehensive understanding of economic evaluation is an area for further research.

The contrasting results for HRQL and wellbeing by gender are difficult to decipher. The analyses reported in Chapter 6 suggests that the CHU9D and MDI are moderately correlated, which indicates they are not measuring precisely the same construct in duplicate. Contrasting results might therefore be expected if the Daily Mile did indeed have different effects by gender in terms of wellbeing and HRQL. The difference could also be attributed to the measure of wellbeing used. The MDI is a long questionnaire, which is perhaps less reliable over time than the CHU9D. It was not possible to assess the test-retest reliability of the MDI using the data available, but this has been noted as a property that should be evaluated in the future.

Gender differences in what contributes to wellbeing, to what extent life experiences impact wellbeing and how individuals interpret questions about wellbeing could exist, explaining the differences in findings. There were no notable gender differences observed in the validation of the MDI, reported in Chapter 6, although this can only be inferred statistically. Qualitative research with children completing the MDI might provide additional insight. Aside from the MDI, other studies have found no gender differences in wellbeing in context-free life satisfaction questionnaires, whereas when domains are specified (e.g. school), differences do emerge (Kaye-Tzadok et al., 2017). Kaye-Tzadok et al. (2017) found that domains contributing to children's wellbeing were consistent between genders, although their importance differed. Satisfaction with school only significantly affected boys' wellbeing for example. It could be that the Daily Mile improved boys' enjoyment of school, which was reflected in their improved wellbeing.

As already discussed, there is neither consensus nor even substantial debate regarding how children's time should be valued. In this analysis a proxy was used, this being a published valuation of adult's leisure time (Verbooy et al., 2018). This is a novel approach and one not identified by Andronis et al. (2019) as being used previously. In the absence of reliable data of the opportunity cost of missed education in primary school, this could be deemed a reasonable approach. There is evidence of the impact of qualifications achieved in secondary education on future earnings and school absence on grades (Department for Education, 2016a). Extending this to younger children, as noted by Andronis et al. (2019) could provide a valuable asset to economic evaluation in a school setting. Disentangling the potential beneficial effects of public health initiatives on education, despite the opportunity cost could complicate any inferences made.

Sensitivity analysis explored the impact of excluding or including children's lost education time for analyses from the school and public sector perspectives, respectively. From the public sector perspective, it reduced the probability that the Daily Mile would be cost-effective to 0%, meaning that the Daily Mile would not be cost-effective at the NICE threshold of £20,000 per QALY.

7.4.3 Strengths and limitations

Assessment of what comprises the schools' perspective was based on the interviews conducted with a sample of school staff. Reflections on the qualitative study have been reported in Chapter 5 section 5.4.1, although pertinent to this analysis is that these views might not be representative of staff at all schools and might not apply to alternative interventions dissimilar to the Daily Mile. Whether this conceptualisation of the school perspective would apply to secondary schools is also uncertain. Regardless, the analysis reported is an exploratory approach that seeks to encapsulate most of the preferences of staff interviewed. Not all aspects could be included, such as the outcome of increased participation and enjoyment and measuring these outcomes are areas for further research. Also, not all costs of the Daily Mile mentioned in the interviews were included in the analyses. One school installed a £10,000 track, although funding was obtained from a competitive application to a charity. Other costs were minimal, so were not included but may have led to a slight reduction in cost-effectiveness of the Daily Mile.

Sensitivity analysis using the SWLS-C demonstrated the challenges of interpreting results of CEA that use outcomes not routinely used in that setting. Firstly, as this is the first use of the MDI and SWLS-C in an economic evaluation, there is no other ICER to which the results could be compared. It is therefore difficult to determine whether the Daily Mile is an efficient use of resources compared to alternative initiatives that could be offered. A meaningful comparison would require an ICER of school-based intervention also using the MDI. Secondly, there is no threshold ICER for either measure. Determining whether the Daily Mile is a cost-effective use of school resources requires some indication of what schools would be willing to pay for a unit of MDI. The ICER for the SWLS-C was understandably much larger than the MDI's, given the much smaller scale (SWLS-C: 1-4, MDI: 12-74), although neither were significantly different between arms. As noted in Chapter 6, the SWLS-C did show some ceiling effects, which may contribute to the observation of no difference between arms.

This is the first economic evaluation of the Daily Mile and the first use of the MDI in a CEA. It was conducted alongside a large randomised study involving a representative sample of UK children, and addressed an important policy question. The trial was pragmatic, allowing schools to autonomously implement the Daily Mile with little interference. The results therefore provide a good indication of the effectiveness and cost-effectiveness of the Daily Mile undertaken in 'real life'. This contrasts with the perceived inflexibility and lack of external validity of RCT designs. The analyses also allow a comparison of the cost-effectiveness of the Daily Mile in terms of wellbeing and QALYs. This provides an additional summary measure of cost-effectiveness that decision makers unfamiliar with QALYs can use to assess the value of implementing the Daily Mile. The absence of a WTP threshold for a unit difference in MDI and understanding of what a meaningful difference in MDI is, makes these results difficult to interpret.

In contrast to the primary outcome of BMIz, a large amount of HRQL, academic attainment and wellbeing data were missing. These were outcomes collected with the assistance of school staff, as opposed to trained researchers. In some cases, just a few participants from each cluster had missing data. This is likely to be due to absence or disinterest. In other clusters, data were missing for all participants. This could be due to technical difficulties with the tablet computers used to record data and/or staff deciding not to or being unable to complete the measures. Imputing missing data from cluster RCTs is computationally challenging. A commonly used and practical method is to conduct multiple imputation, including the cluster variable as a random effect in the imputation model. While multiple imputation is used often, not all statistical commands are compatible with imputed datasets. This creates a situation where the most appropriate method is perhaps not possible. The finding that the Daily Mile seems to be cost-effective in girls is also based on a subgroup analysis, which reflects an even smaller portion of the full trial dataset. Clinical trials are not typically powered for the economic evaluation (Glick et al., 2014), although it could be assumed that the analysis was underpowered.

The rationale for bootstrapping was described in Chapter 2. The cluster RCT design also poses methodological issues for this method. Estimates from cluster trials violate the assumption of independent and identically distributed non-parametrically bootstrapped data (Flynn and Peters, 2005). The two-stage bootstrap procedure has been developed to recognize the clustering within the data. First, clusters are resampled with replacement. This is followed by the resampling of individuals within chosen clusters, to create a bootstrapped sample. This is then repeated to form the distribution (Davison and Hinkley, 1997; Ng et al., 2013). A STATA command has been developed to enable two-stage bootstrapping, although this did would not complete in the Birmingham Daily Mile dataset. A pragmatic solution was to use a standard bootstrap approach, which may have compromised the analysis of uncertainty.

There are several limitations of the trial which have implications for these analyses. Firstly, the trial was only one school year in duration, which means there are no data on the long-term impact of the Daily Mile on health. Costs were limited to the school's costs to implement the Daily Mile. The systematic review in Chapter 3 found that intersectoral costs were rarely incorporated into economic evaluations of similar interventions. The trial was designed quickly, to address an emerging policy issue and with limited funding, therefore collecting healthcare resource use data was unfeasible. This means that the analyses may underestimate the cost-effectiveness from the public health perspective, by not accounting for a potential reduction in NHS resources from improved QALY and wellbeing outcomes. The health benefits of the Daily Mile may not be observed for several years however, so it is possible that no difference would be observed over the short time horizon. As public health in local government is increasingly driven by political cycles (Frew and Breheny, 2019a), evidence provided quickly and demonstrating the short-term benefits that can be expected might be welcomed. Decision makers in schools may also value data allowing them to understand the benefits a child could experience whilst they are attending the school.

A limitation of the Birmingham Daily Mile trial and subsequent economic evaluation was the short follow-up period. It does not enable an understanding of the long-term implications of the Daily Mile on future obesity related illness, quality of life and healthcare resource utilisation. This is an inherent problem with trial based economic evaluations of public health interventions and was discussed in Chapter 2. The school setting offers additional challenges, such as children progressing to new schools and being lost to follow-up. Modelling is an approach that could be used to extend the time horizon of the trial-based analysis and estimate the long-term costs and effects. Modelling is described in Section 2.4.2, and although it can provide a valuable indication of true cost-effectiveness of an intervention, a model is only as good as the data that is available to populate it and results might be highly uncertain. The Early Prevention of Obesity in Childhood (EPOCH) project (Hayes et al., 2019) is intending to develop a more robust model predicting BMI and obesity across childhood and adolescence in addition to healthcare costs, although it can only predict up to age 15 currently due to the ongoing longitudinal study. QALY estimates are not based on primary data, however. Once the model is published the BMIz effect from the Birmingham Daily Mile trial could be inputted, although the model is developed in Australia so would require adaptation. Several model-based analyses were identified in the review reported in Chapter 3, although the majority were microsimulations like the EPOCH model instead of extending trial-based analyses. Estimations of QALYs also tended to be based on adult data. As this was the first use of the MDI in a longitudinal study, there is no data available to predict future effects, although this could be examined in future research.

7.4.4 Recommendations for future research

An immediate extension of this work would be to present the findings to school decision makers in order to determine their usefulness. While a brief summary of the findings was provided in the results, in its present form interpreting the table would still require some understanding of economic evaluation and statistical analysis. Additional work is required to make the results more accessible while maintaining the communication of uncertainty. Establishing whether the results do reflect the school perspective would also be helpful to validate the approach used. This could be done by approaching the individuals who participated in the qualitative study and/or by recruiting a new sample of school staff. A Delphi approach also could be used, taking the interview findings as a starting point and perhaps finding additional components that could be considered in the schools' perspective. Another case study economic evaluation from the schools' perspective could then be conducted.

Future research could address some of the limitations of this analyses. As the Daily Mile has been recommended nationally (HM Government, 2018a), there is an opportunity to conduct further studies utilising existing data sources for both clinical outcomes and healthcare resource use. An economic model could also be developed, to estimate the long-term cost-effectiveness of the Daily Mile. This could appeal to school staff, who were interested in the long-term effectiveness of interventions on health and wellbeing. Currently, modelling would be subject to the same limitations discussed in Chapter 3, with future health benefits predicted based on little evidence. Additional work is required to understand how the MDI can be used to inform decision making. This limitation and future directions have been presented in Chapter 6.

There are different methods of valuing adults' productive time, which have been described in Section 2.5.3. These include the human capital approach (Drummond et al., 2005) and contingent valuation. As noted, the human capital Approach is inappropriate for individuals who do not participate in the labour market, such as children. Approaches for valuing children's time identified by Andronis et al. (2019) included using carer wage rate, school funding lost as a result of non-attendance or an estimation of impact on future earnings. Each of these are pragmatic or spurious estimations based on available data. Estimating the loss of earnings due to absence from education would align more closely with the Human Capital Approach, although obtaining accurate longitudinal data would be challenging. Collecting absence data could be resource intensive and linking this to wages across a lifespan difficult. If linked school attendance data and post-education outcomes are available in the future, this could possibly allow estimation of the value of this time. This would not value the broader benefits of school however, such as developing friendships and personal development.

An alternative approach could be conducting contingent valuation studies to estimate WTP for lost education time, or leisure time. It was briefly discussed in Chapter 2 how valuing outcomes in children, including WTP is challenging. Difficulties include their financial inexperience and cognitive capabilities (Ungar and Gerber, 2009). Several studies have attempted to use parent proxies instead, however issues included parents valuing their children's health 50-100% greater than their own (John et al., 2019). One study estimated parents' WTP for time their children spend receiving medical care (Portrait et al., 2019). Travel, waiting time and treatment time were valued separately. The WTP was a summation of the value they attached to both themselves and their child, as it was assumed the parent would be in attendance. Adults valued children's time as greater than their own and WTP was related to income. If parent proxies are not suitable, WTP has now been attempted with children aged 7-19, despite reservations regarding their capability (Guerriero et al., 2018). Children were able to trade off money and health risks, but WTP was impacted by age, gender and health status. This is methodologically evolving field, but there is potential for children's lost education to be valued using WTP in the future, probably in children aged over 7 and informed by the expertise of experts in child development, health economics and education. A collaborative approach could enable the development of methods that are appropriate and facilitate the most robust data collection possible.

7.4.5 Conclusion

This chapter has presented an exploratory economic evaluation of the Daily Mile from the schools' perspective. The findings are not straightforward, with the Daily Mile not appearing to be harmful, but not potentially effective at reducing BMI in the whole population. Contrasting results in terms of BMI, QALYs and wellbeing mean that the results are difficult to decipher, and decision makers need to identify what their priorities are when using this evidence. The policy message from these results is that in isolation, the Daily Mile is not an effective intervention for obesity prevention. Further research should be designed to identify gender differences, should they exist, and the collection of complete utility and wellbeing data prioritised.

A novel approach to the analyses was used, where qualitative evidence was used to inform the development of the perspective. A measure of wellbeing not previously used in the UK was also included in the analysis and the opportunity cost of children's education time was considered. Comparing the school and public sector perspective, the Daily Mile is evidently more costly from the schools' perspective as it includes the opportunity cost of children's education time. No significant differences in wellbeing overall or by gender might suggest that school decision makers would be unlikely to adopt the Daily Mile, given its additional cost and no apparent wellbeing benefit. It could be that the children's wellbeing was already high, however. The lack of an established threshold WTP for a unit gain in the MDI scale and not understanding what a meaningful difference in MDI scores is, made the interpretation of the results challenging. In contrast, from a public sector perspective the Daily Mile might seem a cost-effective option in girls for NICE decision makers. Implementing the Daily Mile based on this finding might have equity implications however, with boys losing education time while receiving no health benefit. This is reflected upon in the final chapter, which summarises the findings of the thesis and provides an overall discussion of the results and implications.

CHAPTER 8 DISCUSSION

This thesis has examined schools' role in delivering public health interventions and the methods used to conduct economic evaluations in this complex setting. Interviews with school staff enabled an understanding of the decision-making process in regard to prioritising health and wellbeing initiatives and the need for economic evaluation, given similar resource limitations as in health. The competing priority of academic achievement means there is a recognition that providing public health interventions in schools has an opportunity cost. Aside from academic performance, staff were also passionate about observing improvements in wellbeing as a result of new initiatives but did not value outcomes typically reported such as QALYs. Measuring wellbeing using the MDI was shown to be a feasible approach, with the measure demonstrating good psychometric performance and responsiveness to changes in health and HRQL. These findings informed an economic evaluation from the schools' perspective, using wellbeing as the outcome and broadening the perspective to include the opportunity cost of children's time. This chapter summarises the principal findings across all of the empirical work and reflects on the strengths and the limitations of the research. The results are discussed in the context of existing research and directions for future research proposed. Finally, the contributions of this thesis to the methods for conducting economic evaluations of school based public health interventions are presented.

8.1 <u>Summary of principal findings and themes</u>

8.1.1 Decision makers in a school context

The qualitative interviews in Chapter 5 identified that in regard to health and wellbeing initiatives, the primary school decision maker is predominantly the headteacher, supported by senior leadership staff. Teachers have little input into local decision making. The systematic review identified that the majority of economic evaluations did not identify the eventual funder of the initiative, nor the intended audience for the findings. This could be the government, the local authority or the school itself. This suggests that the evidence generated has not been developed to appeal to those expected to utilise the findings. While it is unlikely that headteachers have the time or expertise to find and interpret economic evaluations in the academic literature, they are keen to understand the cost-effectiveness of initiatives. If publications were to be translated into accessible findings, the source literature would need to be appropriately targeted. The decision makers also took into account a range of factors when considering different initiatives. These included the appropriateness for the children's age and the school's needs.

8.1.2 Outcomes from a schools' perspective

The systematic review found that cost-effectiveness outcomes were largely anthropometric or measures of physical activity, e.g. BMIz and MET hours gained. Interviews in Chapter 5 reported that BMI and fitness as outcomes were irrelevant to school staff, with their interests lying in enduring behaviour change and wellbeing. Whilst BMI and METs are objective measures of initial benefits, they are surrogate indicators of future health impact. This could also be said about enjoyment, knowledge and increased participation, which were objectives and outcomes that several staff highlighted as important. None of these were outcomes used in economic evaluations of physical activity interventions identified in the systematic review. It repeatedly arose in interviews that schools' raison d'etre was academic attainment, regardless of staffs' priorities. This is demonstrated in schools' performance framework, overseen by OFSTED and the

DoE. Teacher-reported academic attainment was collected in the Daily Mile trial. In this case, no change was observed and missing data was high, therefore it was not pursued in any further analysis. It was highlighted in interviews that health and academic attainment are perceived as interdependent, and staff expressed an interest in any consequential impacts on academic performance. It could be that this relationship remains assumed. Despite wellbeing being perceived as an important outcome, this was not identified in any existing analyses. As well as academic attainment, wellbeing was also collected in the Daily Mile trial, using the MDI. The MDI had not previously been used in a CEA, or in the UK in fact. This led to the validation of the MDI in Chapter 6 and its inclusion in the economic evaluation in Chapter 7.

8.1.3 Costs from a schools' perspective

The qualitative interviews in Chapter 5 explored the costs associated with the Daily Mile. The tangible costs appeared negligible, such as first aid kit enhancements, with the exception of a £10,000 running track. Staff did however recognise that there was an opportunity cost to providing the Daily Mile. This was spoken about predominantly in relation to the potential impact on English and Maths lessons, which are subjects upon which primary schools' and teachers' performance are evaluated. Teachers appeared to mitigate this by doing the Daily Mile during alternative lessons, such as history. This highlights the tension between prioritising children's health and wellbeing and their academic progress, although staff postulated that these are interdependent. Quantifying this opportunity cost is an area of current uncertainty, therefore in the meantime the economic evaluation of the Daily Mile in Chapter 7 used the cost of adult's leisure time in the interim.

8.1.4 Wellbeing as an outcome in economic evaluation

In Chapter 3, the interviews revealed that school staff thought that improving children's wellbeing was a priority for schools. Nationally this is an emerging priority, with wellbeing being incorporated into a new PSHE curriculum due to be implemented in 2020. Demonstrating wellbeing impact requires the use of a robust measure. In Chapter 6, analyses were conducted to assess the performance of the MDI in a UK sample for the first time. The results were encouraging, with performance not seemingly different between genders and age groups and comparable to previous validation studies. The MDI showed some association with quality of life, suggesting that it could be complementary to measures such as the CHU9D. As the MDI was shown to perform adequately, it was used as an outcome in a CEA of the Daily Mile in Chapter 7. There was no significant difference in wellbeing between the Daily Mile and control. How to contextualise the results of the CEA is challenging, given this was the first use of the MDI in a clinical trial and CEA. Including the MDI in economic evaluation is nevertheless a possibility for future work, particularly as no economic evaluations identified in the systematic review reported wellbeing outcomes.

8.2 Reflections on the research

The empirical work reported in this thesis are the findings of studies nested within one case-study trial. The Birmingham Daily Mile trial aimed to test the presumption that 15 minutes of daily physical activity reduces obesity and improves wellbeing in primary school children. The finding that wellbeing is an important outcome for school staff is not confined to an obesity prevention initiative, however. This was demonstrated in the interviews and could be applied to other health promotion initiatives, such as those promoting mental health. Wellbeing is unlikely to be the definitive outcome in all school based public health interventions, with its relevance to preventative interventions such as smoking campaigns perhaps less obvious in a primary school setting.

Findings of this thesis are most applicable to UK primary schools. Currently the Daily Mile is recommended in primary schools only, thus the Birmingham Daily Mile trial only included such schools. This particular initiative is suited to primary education, where teachers have responsibility for their own class most of the time. In an interview, one teacher reported texting their cover teacher to remind them about the Daily Mile for example. The Daily Mile would be logistically challenging to implement in a secondary school setting, where children are transient between classes and subject teachers each day. The spontaneous nature of the Daily Mile would be unfeasible, with coordination required to balance the opportunity cost between subjects and avoid potentially thousands of children congregating to run a mile at once. As secondary schools have a more structured timetable with specialist teachers delivering one or two subjects to many classes, it could be speculated that the decision-making process is even more dominated by leadership staff in this setting. In secondary schools, identifying where the opportunity cost of delivering public health interventions lies is potentially easier, and perhaps more consistent both between and within schools.

There is extensive literature examining the cost-effectiveness of health promotion in schools. Obesity prevention is a subject of high interest and methodological development, and new economic evaluations are rapidly being published. In hand, syntheses of these publications are increasingly available. The systematic review reported in Chapter 3 focused on interventions for addressing risk factors for obesity, not necessarily weight reduction. Furthermore, in alignment with the aims of the thesis, it was limited to economic evaluations in a school setting. Between the review's conception and production of this thesis, a review exploring a similar question was published (Oosterhoff et al., 2018). While the thesis review identified a broader literature, it also focused on the challenges of evaluating public health interventions. This resulted in a review that contributes to the public health literature in general and that of school-based interventions.

While the systematic review identified the paucity of literature comprehensively addressing the extraordinary requirements of evaluations of public health interventions, for example long-term maintenance of effect and intersectoral costs, the economic evaluation of the Daily Mile was also unable to address several of these issues. The trial was pragmatic and conducted rapidly to address an important policy question. With additional resource, the questions of the effectiveness and cost-effectiveness of the Daily Mile could have been answered more conclusively with a trial incorporating long-term follow up, data linkage between education and the health system and parent participation to triangulate healthcare impact, spill over effects and behaviour change.

Participation in qualitative interviews was limited to school staff, but perhaps could have benefited from the inclusion of children and parents. Expanding the sample to include children may have provided additional insights into the implementation of public health interventions. Children's perspectives could aid the interpretation of the Daily Mile trial results for example or understand their views on the opportunity costs. Children's participation in qualitative research requires methods to be adapted to be age appropriate, and understandably requires substantial additional ethical oversight. This meant that involving children was not pursued, although the potential value of their contribution was recognised. Staff were asked about how they believed children perceived the Daily Mile, therefore a proxy report of their aggregate views was obtained. Staff also commented on the completion of the MDI. Think-aloud interviews with the children could have also strengthened the investigation of the validity of the MDI. Interviewing participating children's parents might also have provided a valuable contribution to some aspects of the study, such as corroborating headteacher's views that parents had a bearing on decision making. Involving parents was also beyond the scope of this thesis, but perhaps an area for future research.

A strength of the MDI evaluation was the large sample size, although there was a significant amount of missing data. If wellbeing was considered an important outcome by staff, one might expect them to be invested in ensuring wellbeing data were collected comprehensively. Interviews with staff provided some insights into why the data quality was not as expected. Reports from staff suggested that technical issues were a cause, with the computer tablets timing out if completion was paused for some time, such as for break time. Recommencing completion meant starting from the beginning and no data was recorded if the questionnaires were abandoned. This could explain why some classes had no MDI data recorded. The consecutive completion of the CHU9D and MDI potentially had implications for the economic evaluation also. Had the data saved periodically, the CHU9D missing data probably would have been reduced, thus providing more reliable estimates of the HRQL impact of the Daily Mile. Another cause of missing data was perhaps the MDI's length and consequent time burden, with some staff perhaps opting to not facilitate completion at 12-months. Staff commented on the length of the MDI, with one explicitly requesting not to have to do it again at follow-up. These issues do not relate to the children's completion of the measure, suggesting the missing data were not entirely attributable to their disinterest or difficulty completing it. Staff's concerns about time could partially explain the challenges experienced recruiting participants for the qualitative study. Despite offering flexible times and telephone interviews, volunteers were not forthcoming. Incentivising participation might have improved response rates, although the opportunity cost of the interview remains.

Whilst not an aim, this thesis has highlighted equity considerations when conducting economic evaluations in a school setting. Boys might experience a small wellbeing gain yet are exposed to the opportunity cost of lost academic time. This is in contrast to girls, who experienced both BMIz and QALY gains. Whether or how this should be addressed in this setting is currently uncertain, although equity is consistently raised as an issue when evaluating public health interventions (Edwards et al., 2013; Weatherly et al., 2009). The role of the health economist is to conduct and present the analyses with its associated uncertainty, which the decision maker will interpret. The case study economic evaluation reported in Chapter 7 explicitly presented the results by gender. The judgement of whether the Daily Mile is a good use of resources when only a subgroup benefits would therefore be up to the decision maker, such as the headteacher. The potential trade-off between academic attainment and health and wellbeing would also be pertinent in this scenario. It is unfortunate but unavoidable that there was not the opportunity to obtain feedback on these issues from the school decision makers interviewed. This challenge could arise in other subgroups, such as ethnicity and investigating how we deal with this is an area for future research.

8.3 **Results in context**

The systematic review highlighted that challenges of conducting economic evaluations of public health interventions described in the literature still remain. Eleven years ago, issues experienced by public health policymakers when applying the NICE recommendations for the evaluation of public health interventions, developed by the Centre of Public Health Excellence (CPHE) were noted (Chalkidou et al., 2008). While the perceptive was described as broader than that permitted for healthcare interventions, a CCA was only acceptable under certain circumstances and had to be presented alongside a CUA. Research priorities for methods development cited in Chalkidou et al. (2008) largely still remain unaddressed, and align with the methodological challenges reported in Chapter 2 (Edwards et al., 2013; Weatherly et al., 2009). A recent review of economic evaluations of targeted physical activity interventions also assessed how the distinguishing features of public health interventions are addressed (Cochrane et al., 2019). In this context it was also found that issues like incorporating intersectoral costs and outcomes and attributing long-term effects are not being adequately addressed.

Broadening the scope of economic evaluation to include wellbeing is an approach gaining traction in public health. A Delphi study with UK local government decision makers found that approaches to include wellbeing in economic evaluation is an area of methodological need (Frew and Breheny, 2019b). Decision makers in the Netherlands (Versteegh et al., 2016) recognised the potential to expand the scope of economic evaluation to measure wellbeing as the outcome for non-curative interventions, in place of the QALY. Challenges identified include the standardisation of instruments, detecting change in wellbeing and obtaining preferences, although the ICECAP was mentioned as a potential measure. Parallels could also be drawn between the inclusion of wellbeing within the evaluation framework for school-based public health interventions and the literature vouching for its use in social care. Wellbeing has been suggested as an alternative outcome when evaluating social care interventions in older adults (Makai et al., 2014) where like public health, health maximisation may not be the primary goal. What wellbeing encompasses, the contribution of health to wellbeing and wellbeing's overlap with HRQL were noted as posing challenges to measuring the concept (Makai et al., 2014). Recent attempts to explore which aspects of wellbeing were important to decision makers from heath, public health and social care using qualitative methods was unsuccessful (Peasgood et al., 2019), with no agreed consensus. Decision makers in social care and public health valued wellbeing more than medical decision makers, although they agreed that it is not adequately measured currently.

An important consideration for decision making is that all the relevant costs and outcomes of an intervention should be measured (Drummond et al., 2005). As wellbeing is a complex concept, there might be several domains that need to be captured in order to provide an adequate evaluation of an intervention's impact on wellbeing. A Delphi study with adults in the general (adult) population identified five core domains of health-related subjective wellbeing, out of a potential 21 that could be used in economic evaluation (de Vries et al., 2016). These included physical independence, positive affect/happiness, negative affect/feeling lost and lonely, autonomy, and personal growth. Domains of wellbeing important to children receiving children's services have also been explored, with the intention of developing a measure for economic evaluation (Holder et al., 2011). Using literature reviews and discussions with children, eight domains were identified. Several were similar to domains in ICECAP measures of capability-wellbeing (Coast et al., 2008a), such as feeling safe and secure, being able to make choices and be understood. Making friends and use of free time after school are domains more distinct from adult conceptualisations of wellbeing. A review identified several candidate measures of adult wellbeing for use in economic evaluation, with the ICECAP-O and ASCOT selected as most promising (Makai et al., 2014).

Neither measure is suitable for use with children, although a capability-wellbeing measure for children is being explored (Coast, 2019).

8.4 Implications for policy making and research

This thesis provides important contributions to the methods used to conduct economic evaluations of public health interventions provided in schools. While the results reported in Chapter 7 might have implications for implementation of the Daily Mile, the wider implications relate to methods that could be used to develop evidence to inform future policy. Implications could be at a school, local authority or national level.

8.4.1 Choice of outcomes and assessment frameworks

The inclusion of appropriate outcomes could improve prioritisation of new public health initiatives at all levels of decision making. An advantage of the QALY is that it enables the comparison of interventions for different conditions. School-based initiatives might be aimed at promoting either/both physical and mental health, which would suggest that a generic measure would be useful to decision makers if they wanted to compare interventions. Findings from this thesis suggest that decision makers did not value outcomes indicative of health changes (e.g. BMI), or the QALY. Alternative generic outcomes that can capture wellbeing outcomes could appeal to policymakers when making recommendations, however. This would be most appropriate at a school-level, where preference-based outcomes are not likely to be demanded. As schools are funded by the public, like the health service, it could be argued that preference-based outcomes should be used in this setting. Given the absence of a wellbeing measure for the primary school population currently, there is the opportunity to use the MDI also.

This thesis has affirmed the tension between improving health, wellbeing and academic attainment in schools. Initiatives that positively impact attainment as well as children's health and wellbeing might have greater potential to be adopted and successfully implemented. Policymakers or those responsible for the assessment of schools such as OFSTED should consider outcomes broader than academic attainment when evaluating school performance. Staff resented that they were judged on children's attainment and progress, noting that academic attainment was dependent on good health and wellbeing. Using economic evaluation to identify the most efficient use of school resources could result in investment in other areas, such as curriculum resources. This could have a positive effect on academic performance in addition to improving health and wellbeing. This could also be looked upon favourably by those assessing schools.

8.4.2 Presentation and tailoring of evidence

The transfer of public health from the NHS to local authorities inspired a number of studies examining the decision-making process and use of evidence. Qualitative research suggests that the presentation of evidence has changed to be targeted at decision-makers and framed as a persuasive document for action (Kneale et al., 2019), as opposed to evidence syntheses applicable to a national or global setting. Kneale et al. (2019) advocated developing accessible evidence for those working in local government without public health training, such as councillors. Parallels could be drawn with school staff, who have neither economics or public health training, yet also want evidence that is interpretable and relevant. Participants from local government were interested in the applicability of evidence to their local population, for example the demographic composition. This was not found in the interviews with school staff, with overall impact and cost-effectiveness being preferred with no reference to the need to tailor this to their local situation (e.g. level of deprivation).

School decision makers should be involved in the generation of evidence. Nutbeam (2003) has highlighted the need for policy-relevant evidence in public health. Evidence that is fit for purpose should evaluate practical interventions, assessing their costs and benefits, and producing timely results that can be effectively communicated. Nutbeam (2003) argues that the responsibility for making appropriate evidence available is with those generating it and those that will use it (e.g. headteachers). The involvement of stakeholders in public health evaluations is recognised as an asset (Rychetnik et al., 2002). Rychetnik et al. (2002) widen the definition of stakeholders to those who will experience the intervention, particularly in relation to outcome measures. This could be likened to Patient and Public Involvement (PPI), which is a concept increasingly embedded within current research, and often a consideration for funding bodies. In the case of public health interventions in schools, stakeholders could include headteachers, teachers, local government, children, parents/guardians and siblings. This thesis has identified who would use economic evidence at a school level, and what costs and outcomes are desirable include.

8.4.3 Tailored evidence tools for public health

PHE provide several resources to assist commissioners' decision making that are informed by health economics (PHE, 2018a). Predominantly targeted at public health teams within local authorities, their suitability for use by school decision makers might be limited at present. Findings from this thesis could be used to tailor these tools for school-based interventions. The PHE Prioritisation Framework (PHE, 2018a) uses MCDA to compare across potential programmes using a transparent framework to inform whether funding should be increased, maintained or reduced. It involves political and strategic considerations and can involve many stakeholders. As school decision makers would not be comparing between other sectors, this framework in its current form is unlikely to be helpful, although it potentially could be adapted for use in prioritisation at a school level. The findings from this thesis could inform which criteria are considered, such as the inclusion of academic attainment, health outcomes and the opportunity cost of delivering the intervention.

The Spend and Outcomes Tool (SPOT) and Health Economics Evidence Resource (HEER) (PHE, 2018a) provide data and summaries of published evidence that can be used to populate the PHE Prioritisation Framework. The HEER is a database of economic evidence that PHE believes to be of sufficient quality to inform local decision making. Selected research is organised into areas of prevention funded by the Public Health Grant local authorities receive, such as Obesity and Physical Activity. Schools and local authorities could use this tool to identify evidence for new initiatives, although currently there are few school-based interventions reviewed. The findings of this thesis could contribute to the refinement of the data extraction to suit school decision makers' needs, as opposed to the local authorities'. The SPOT enables local authorities and CCGs to compare spending across different topics against outcomes achieved in other local authorities to identify variation and areas for improvement. While this data cannot be used for prioritisation alone, it can be incorporated into other analyses. PHE also produces interactive ROI resources for specific topics, drawing on evidence included in the SPOT for example (PHE, 2018a). Local authorities can populate the tool with local data to estimate the value of investing in different services such as the prevention of musculoskeletal conditions. This tool could be adapted to include data sources identified as important for schools.

8.4.4 Equity considerations

The findings of this thesis raise questions about the use of economic evidence when some groups benefit, and others do not. The gender subgroup effect was not anticipated at the study's conceptualisation, and in fact the qualitative research in Chapter 5 would suggest that implementation fidelity was inferior in girls. In 2018, there was an attainment 'gender gap' of 8% in reading, writing and maths, with girls aged 11 outperforming boys in year 6 (Department for Education, 2018b). Furthermore, education and academic attainment in childhood are notable social determinants of health (Cutler and Lleras-Muney, 2006). 2018 NCMP data showed 32.2% of girls being overweight or obese, compared to 36.4% of boys (NHS Digital, 2018). The Government's Childhood Obesity Plan is accompanied by an Equality Assessment, which considers the impact of the policy on the protected characteristics such as sex (HM Government, 2018b). As both sexes would undertake the Daily Mile and there was no published evidence of gender differences in effect, the equality impact was judged as neutral. If the evidence in this thesis is considered, the Daily Mile could be seen to potentially widen weight and academic inequalities. As already noted, it would be up to the decision maker to assess the implications of these results and potentially make value judgements regarding opportunity costs and health benefits.

8.5 Directions for future research

This thesis uses varying methodologies to explore several research questions and from these many areas for future research have arisen, both methodological and experimental. These are in the context of school-based interventions, public health economics and child health. They are now described.

8.5.1 The opportunity cost of children's time

School staff recognised the opportunity cost of delivering public health interventions in terms of lost education. The valuation of children's time is generating research interest and is recognised as a consideration that has been overlooked previously (Andronis et al., 2019). Of the three approaches identified by Andronis et al. (2019) (caregiver wage rate, school funding and future earnings), one could prove informative if school absence and educational attainment in the primary school years could be linked to future earnings. Such data are only available in the latter years of compulsory education currently. Nevertheless, the 'time' lost when children receive a healthcare intervention could be spent in a number of ways, and the impact on income might not be an appropriate proxy. Spending time in hospital might be to the detriment of developing secure attachments with family members, going to school or participating in extra-curricular activities for example. When public health interventions are delivered in school, the opportunity cost could be a history lesson, developing friendships at playtime or receiving acknowledgement of your achievements in an assembly. A distinction is that the public health intervention is not treating an existing illness, therefore might not be considered a necessity. In the case of a physical activity intervention, differing 'doses' in terms of duration will have correspondingly variable opportunity costs. This could have implications for the cost-effectiveness, perhaps leading to additional interest in the 'doses' required to achieve the desired effect. Valuing this time in a public health context is therefore important, although quantifying this opportunity cost is conceptually and methodological complex.

8.5.2 Wellbeing as an outcome in economic evaluation

For a wellbeing initiative the 'dose' of the intervention might need to be varied to achieve the optimal level or change in wellbeing. For the MDI, the optimal wellbeing score is currently unknown. To strengthen the MDI's evidence base for assessing change over time in clinical trials, additional research could be conducted to establish what a meaningful change in wellbeing is. A current limitation of using the MDI to inform decision making is that its use in economic evaluation is currently limited to CCA and CEA. Developing a preference-based algorithm to enable its use in CUA is also a potential avenue of research. Development of a children's preference based wellbeing measure for children in care has previously been initiated, but nothing produced beyond the identification of important domains of wellbeing (Holder et al., 2011). There is also ongoing work to develop a children's measure of capability-wellbeing (Coast, 2019).

There is now evidence that the Grade 4 MDI performs well in a UK population and in light of this, researchers may want to use the Grade 7 version to assess wellbeing in older children. As this version is yet to be used in the UK, it would be appropriate to conduct similar analyses to what has been reported in this thesis. In addition to this replication, examination of the MDI's content validity could be conducted concurrently, as this is a property largely unexplored and a current gap in the literature. It seems that as the MDI comprises questions from existing instruments, its content validity has often been assumed.

8.5.3 School stakeholder engagement

Policy relevant research and dissemination is an important aspect of the research process (Nutbeam, 2003). This seems to have been neglected in the school-based literature, related to health economics. This thesis has found that schools do not seem to fit within the decision-making structure observed in the NHS, which is overseen by NICE for example (NICE, 2014). Despite this, demonstrating an evidence-based case for implementing a new public health intervention appears to be critical for its successful adoption. The development of a framework for the early engagement of school stakeholders might help to generate the appropriate evidence and methods of dissemination to best inform decision making. Producing findings accessible to an audience not trained in healthcare nor health economics and identifying outcomes that are meaningful in this setting appear to be important factors. This process could explore both clinical effectiveness and cost-effectiveness evidence and perhaps use Delphi methods. An extension could develop a framework to be used during the intervention design stage, as per the MRC guidance (Craig et al., 2008). The involvement of health economists could enable the identification of the potential intersectoral cost implications and value of information analyses could aid the case for ceasing or pressing ahead with intervention design.

8.5.4 Evidence to inform economic models

Facilitating enduring behaviour change was identified as an important outcome by school staff. This was seen as more indicative of benefit than a change in BMI for example. A component of this was enjoyment of an activity, which is perhaps a way that this could be captured in the short-term. There is a measure of children's physical activity enjoyment available (Moore et al., 2009), which could be readily implemented in trials. Future research could examine the feasibility of conducting longitudinal research which associates this data with behavioural intentions and long-term behaviour change, perhaps linking to existing surveillance measures of population physical activity (e.g. Health Survey for England) (Strain et al., 2019). This could then be used in economic models to estimate the cost-effectiveness of interventions over longer time horizons. The first cohort of year 6 children with their NHS number linked to their reception NCMP measurement will be measured in 2019/2020 (PHE, 2017). Further linkage to SATs data and physical activity participation could also be used to generate economic model inputs. Academic attainment is the main priority of schools, evidenced both in the literature and in the interviews reported in this thesis. Being able to show the degree to which health benefits translate into an improvement in academic attainment would be a valuable asset for demonstrating effectiveness and cost-effectiveness.

8.6 **Contributions of the thesis**

8.6.1 The first use of qualitative methods to explore the schools' perspective

Education is a complex context for economic evaluation, with its incongruence with the healthcare perspective now documented. This thesis reports the first qualitative investigation of the schools' perspective, generating new knowledge regarding who the decision maker is, the costs of school-based interventions and important outcomes in this context. While NICE has indicated some flexibility when evaluating public health interventions, such as the use of broader outcomes; methods currently used when evaluating interventions delivered in schools are seemingly irrelevant to those tasked with implementing them. CUA for example is neither comprehensible, nor helpful. These interviews with school staff will enable future economic evaluations to produce evidence suitable for schools, and inspire methodological advances informed by robust qualitative research.

8.6.2 The first assessment of the MDI's suitability in the UK and in a trial context

Demonstrating the validity of an outcome measure in the intended population (e.g. UK children) and setting (e.g. school-based clinical trials) is important to support its suitability for use in developing reliable evidence (Streiner et al., 2015). This thesis reports the first use of the MDI Grade 4 in a UK population. Validation work, which went beyond what has been conducted previously suggested that where comparisons allow, the measure performs as well, if not better than in Canadian, Australian and Swiss children. While previous validation studies used associations between MDI scales to test construct validity, this thesis used an external measure, the CHU9D to examine this property. The CHU9D is a robust and widely used measure in economic evaluations of children's interventions. Their concurrent use allowed the first comparison between these two measures, finding that the MDI does not replicate data obtained using the CHU9D. Being conceptually distinct indicates their use could be complementary, showing impacts of interventions on both wellbeing and HRQL. Another critical domain of evidence previously undocumented for the MDI was its responsiveness. Previously, the MDI had only been used for population monitoring of wellbeing, with no evidence of individual level longitudinal change or use in a clinical trial. This thesis provides evidence that it is a feasible measure in a school-based clinical trial, indicating that scales particularly in the SED domain are responsive to change in self-reported health and HRQL.

8.6.3 A comparison between the public sector and schools' perspective

For economic evidence to inform prioritisation and make a difference, it needs to be presented from the viewpoint of the decision maker. Primary qualitative research enabled an understanding of what a schools' perspective comprises, which informed an exploratory economic evaluation of the Daily Mile conducted from two perspectives, namely the schools' and the public sector's. Using primary qualitative evidence to inform an analysis that is informative for the decision maker is an approach not used in a school setting previously. It was ascertained that maximising wellbeing was a framework important in this context, thus a CEA using wellbeing as the outcome and including children's opportunity cost of lost education was presented, in addition to a CUA. While economic evaluations frequently present both a CUA and CEA, for example cost per QALY and cost per BMI unit in physical activity interventions, neither of these were identified as of interest to school staff. It is however recognised that CEA is not always useful when attempting to make consistent decisions across contexts. This thesis uses a novel, mixed methods approach to develop economic evidence that is relevant to the decision maker in a school setting.

LIST OF REFERENCES

Abbott, S., Florin, D., Fulop, N. and Gillam, S. (2002). The meaning of 'health improvement'. *Health Education Journal,* 61**,** 299-308.

Academies Act. (2010) Available at: http://www.legislation.gov.uk/ukpga/2010/32/contents (Accessed: 8 August 2019).

Acheson, D. (1988). *The report of the committee of inquiry into the future development of the public health function,* London.

Adab, P., Barrett, T., Bhopal, R., Cade, J., Canaway, A., Cheng, K., Clarke, J., Daley, A., Deeks, J., Duda, J., Ekelund, U., Frew, E., Gill, P., Griffin, T., Hemming, K., Hurley, K., Lancashire, E., Martin, J., McGee, E., Pallan, M., Parry, J. and Passmore, S. (2018a). A cluster randomised controlled trial of the effectiveness and cost-effectiveness of an obesity prevention intervention in primary school age children *Health Technology Assessment,* 22.

Adab, P., Pallan, M. and Whincup, P. H. (2018b). Is BMI the best measure of obesity? *BMJ,* 360**,** k1274.

Al-Janabi, H., Peters, T., Brazier, J., Bryan, S., Flynn, T., Clemens, S., Moody, A. and Coast, J. (2013). An investigation of the construct validity of the ICECAP-A capability measure. *Quality of Life Research,* 22**,** 1831-1840.

Al Tunaiji, H., Davis, J. C., Mackey, D. C. and Khan, K. M. (2014). Population attributable fraction of type 2 diabetes due to physical inactivity in adults: A systematic review. *BMC Public Health,* 14**,** 1-9.

Andrews, J. (2018) *School performance in academy chains and local authorities - 2017* Available at: https://epi.org.uk/wp-content/uploads/2018/06/EPI-Academy-LA-Performance___.pdf (Accessed: 11 August 2019).

Andronis, L., Maredza, M. and Petrou, S. (2019). Measuring, valuing and including forgone childhood education and leisure time costs in economic evaluation: Methods, challenges and the way forward. *Social Science & Medicine,* 237**,** 112475.

Arnold, J., Bruce-Low, S., Henderson, S. and Davies, J. (2016). Mapping and evaluation of physical activity interventions for school-aged children. *Public Health,* 136**,** 75-79.

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review,* 26**,** 851-883.

Arvidson, M., Lyon, F., McKay, S. and Moro, D. (2010) *The ambitions and challenges of SROI.* Available at: http://epapers.bham.ac.uk/788/1/WP49_Ambitions_and_challenges_of_SROI_-_Arvidson%2C_Lyon%2C_McKay%2C_Moro_Nov_2010.pdf (Accessed: 16th February 2020).

Attema, A., Brouwer, W. and Claxton, K. (2018). Discounting in economic evaluations. *PharmacoEconomics,* 36**,** 745-758.

Babey, S., Wu, S. and Cohen, D. (2014). How can schools help youth increase physical activity? An economic analysis comparing school-based programs. *Preventive Medicine,* 69**,** S55-60.

Balakrishnan, R., Allender, S., Scarborough, P., Webster, P. and Rayner, M. (2009). The burden of alcohol-related ill health in the United Kingdom. *Journal of Public Health,* 31**,** 366-373.

Bambra, C., Gibson, M., Amanda, S., Wright, K., Whitehead, M. and Petticrew, M. (2009). Tackling the wider social determinants of health and health inequalities: Evidence from systematic reviews. *Journal of Epidemiology and Community Health,* 64**,** 284-291.

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology and Health,* 13**,** 623-649.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education and Behavior,* 31**,** 143-164.

Bandura, A. (2005). The primacy of self-regulation in health promotion. *Applied Psychology,* 54**,** 245-254.

Banke-Thomas, A., Aarons, G., Madaj, B., Charles, A. and van den Broek, N. (2015). Social return on investment (SROI) methodology to account for value for money of public health interventions: A systematic review. *BMC Public Health,* 15**,** 582.

Barker, D. and Osmond, C. (1986). Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *The Lancet,* 327**,** 1077-1081.

Bärnighausen, T., Bloom, D., Canning, D., Friedman, A., Levine, O., O'Brien, J., Privor-Dumm, L. and Walker, D. (2011). Rethinking the benefits and costs of childhood vaccination: The example of the haemophilus influenzae type b vaccine. *Vaccine,* 29**,** 2371-2380.

Barrett, J., Gortmaker, S., Long, M., Ward, Z., Resch, S., Moodie, M., Carter, R., Sacks, G., Swinburn, B., Wang, C. and Cradock, A. (2015). Cost effectiveness of an elementary school active physical education policy. *American Journal of Preventive Medicine,* 49**,** 148-59.

Barton, P., Bryan, S. and Robinson, S. (2004). Modelling in the economic evaluation of health care: Selecting the appropriate approach. *Journal of Health Services Research & Policy,* 9**,** 110-118.

Batorova, D. and Sørensen, J. (2019). Methodological review of model-based cost-effectiveness analyses of school-based interventions to increase pupils level of physical activity. *Journal of Physical Education,* 30.

Becerra, V., Gracia, A., Desai, K., Abogunrin, S., Brand, S., Chapman, R., García Alonso, F., Fuster, V. and Sanz, G. (2015). Cost-effectiveness and public health benefit of secondary cardiovascular disease prevention from improved adherence using a polypill in the UK. *BMJ Open,* 5.

Becker, G. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education,* New York, National Bureau of Economic Research.

Belfield, C., Farquharson, C. and Sibieta, L. (2018) *2018 annual report on education spending in England.* Available at: https://www.ifs.org.uk/publications/13306 (Accessed: 9 August 2019).

Bennett, A., Cunningham, C. and Johnston Molloy, C. (2016). An evaluation of factors which can affect the implementation of a health promotion programme under the schools for health in Europe framework. *Evaluation and Program Planning,* 57**,** 50-54.

Bevans, K. and Forrest, C. 2009. The reliability and validity of children's and adolescents' self-reported health. *In:* Ungar, W. (ed.) *Economic evaluation in child health.* Oxford: Oxford University Press.

Birch, S. and Gafni, A. (1992). Cost effectiveness/utility analyses. *Journal of Health Economics,* 11**,** 279-296.

Black, R., Allen, L., Bhutta, Z., Caulfield, L., de Onis, M., Ezzati, M., Mathers, C. and Rivera, J. (2008). Maternal and child undernutrition: Global and regional exposures and health consequences. *The Lancet,* 371**,** 243-260.

Bonell, C., Humphrey, N., Fletcher, A., Moore, L., Anderson, R. and Campbell, R. (2014). Why schools should promote students' health and wellbeing. *BMJ,* 348**,** g3078.

Booth, J. N., Leary, S. D., Joinson, C., Ness, A. R., Tomporowski, P. D., Boyle, J. M. and Reilly, J. J. (2014). Associations between objectively measured physical activity and academic attainment in adolescents from a UK cohort. *British Journal of Sports Medicine,* 48**,** 265.

Bowden, A. and Belfield, C. (2015). Evaluating the talent search trio program: A benefit-cost analysis and cost-effectiveness analysis. *Journal of Benefit-Cost Analysis,* 6**,** 572-602.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health,* 27**,** 281-291.

Boyle, G. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences,* 12**,** 291-294.

Bradley, R. and Corwyn, R. (2002). Socioeconomic status and child development. *Annual Review of Psychology,* 53**,** 371-99.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3**,** 77-101.

Braveman, P. and Gruskin, S. (2003). Defining equity in health. *Journal of Epidemiology and Community Health,* 57**,** 254.

Brazier, J. and Tsuchiya, A. (2015). Improving cross-sector comparisons: Going beyond the health-related QALY. *Applied Health Economics and Health Policy,* 13**,** 557-565.

Brazier, J. E. and Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care,* 42**,** 851-859.

Breheny, K., (2013), 'Explain the fundamental principles of both a welfarist and an extra-welfarist approach to resource allocation decisions in health care. Explain and justify which approach is the most appropriate for decision making in a healthcare context' *Introduction to Health Economics.* University of Birmingham. Unpublished essay.

Breheny, K., Adab, P., Passmore, S., Martin, J., Lancashire, E., Hemming, K. and Frew, E. (2018). A cluster randomised controlled trial evaluating the effectiveness and cost-effectiveness of the Daily Mile on childhood obesity and wellbeing; the Birmingham Daily Mile protocol. *BMC Public Health,* 18**,** 126.

Breheny, K., Passmore, S., Adab, P., Martin, J., Hemming, K., Lancashire, E. R. and Frew, E. (2020). Effectiveness and cost-effectiveness of the Daily Mile on childhood weight outcomes and wellbeing: A cluster randomised controlled trial. *International Journal of Obesity*.

Brennan, A. and Akehurst, R. (2000). Modelling in health economic evaluation. *PharmacoEconomics,* 17**,** 445-459.

Breslin, G., Gossrau-Breen, D., McCay, N., Gilmore, G., MacDonald, L. and Hanna, D. (2012). Physical activity, gender, weight status, and wellbeing in 9- to 11-year-old children: A cross-sectional survey. *Journal of Physical Activity and Health,* 9**,** 394-401.

Breslow, L. (1999). From disease prevention to health promotion. *JAMA,* 281**,** 1030-1033.

Briggs, A. and O'Brien, B. (2001). The death of cost-minimization analysis? *Health Economics,* 10**,** 179-184.

Briggs, A., Sculpher, M. and Buxton, M. (1994). Uncertainty in the economic evaluation of health care technologies: The role of sensitivity analysis. *Health Economics,* 3**,** 95-104.

Briggs, A., Weinstein, M., Fenwick, E., Karnon, J., Sculpher, M. and Paltiel, A. (2012). Model parameter estimation and uncertainty analysis: A report of the ISPOR SMDM modeling good research practices task force working group–6. *Medical Decision Making,* 32**,** 722-732.

Brousselle, A., Benmarhnia, T. and Benhadj, L. (2016). What are the benefits and risks of using return on investment to defend public health programs? *Preventive Medicine Reports,* 3**,** 135-138.

Brouwer, W., Culyer, A., van Exel, J. and Rutten, F. (2008). Welfarism vs. Extra-welfarism. *Journal of Health Economics,* 27**,** 325-338.

Brouwer, W., Niessen, L., Postma, M. and Rutten, F. (2005). Need for differential discounting of costs and health effects in cost effectiveness analyses. *BMJ,* 331**,** 446.

Brown, H., Perez, A., Li, Y., Hoelscher, D., Kelder, S. and Rivera, R. (2007). The cost-effectiveness of a school-based overweight program. *The International Journal of Behavioral Nutrition and Physical Activity,* 4**,** 47.

Brown, T., Moore, T. H. M., Hooper, L., Gao, Y., Zayegh, A., Ijaz, S., Elwenspoek, M., Foxen, S. C., Magee, L., O'Malley, C. and et al. (2019). Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*.

Burke, H., Davis, M., Otte, C. and Mohr, D. (2005). Depression and cortisol responses to psychological stress: A meta-analysis. *Psychoneuroendocrinology,* 30**,** 846-856.

Byford, S., McDaid, D. and Sefton, T. (2003). *Because it's worth it: A practical guide to conducting economic evaluations in the social welfare field,* York, UK, Joseph Rowntree Foundation.

Camfield, L. and Skevington, S. M. (2008). On subjective well-being and quality of life. *Journal of Health Psychology,* 13**,** 764-775.

Carmines, E. and Zeller, R. (1979). *Reliability and validity assessment,* California, USA, Sage publications.

Castelli, L., Marcionetti, J., Crescentini, A. and Sciaroni, L. (2018). Monitoring preadolescents' well-being: Italian validation of the Middle Years Development Instrument. *Journal of Child Indicators Research,* 11**,** 609-628.

Cecchini, M., Sassi, F., Lauer, J. A., Lee, Y. Y., Guajardo-Barron, V. and Chisholm, D. (2010). Tackling of unhealthy diets, physical inactivity, and obesity: Health effects and cost-effectiveness. *The Lancet,* 376**,** 1775-1784.

Centers for Disease Control and Prevention. (2018). *Well-being concepts* [Online]. Available: https://www.cdc.gov/hrqol/wellbeing.htm [Accessed 8th February 2020].

Centre for Workforce Intelligence. (2014) *Mapping the core public health workforce.* Available at: http://www.cfwi.org.uk/publications/mapping-the-core-public-health-workforce (Accessed: 3 October 2016).

Cesaroni, G., Forastiere, F., Agabiti, N., Valente, P., Zuccaro, P. and Perucci, C. A. (2008). Effect of the Italian smoking ban on population rates of acute coronary events. *Circulation,* 117**,** 1183-1188.

Chalkidou, K., Culyer, A., Naidoo, B. and Littlejohns, P. (2008). Cost-effective public health guidance: Asking questions from the decision-maker's viewpoint. *Health Economics,* 17**,** 441-448.

Chesham, R., Booth, J., Sweeney, E., Ryde, G., Gorely, T., Brooks, N. and Moran, C. (2018). The Daily Mile makes primary school children more active, less sedentary and improves their fitness and body composition: A quasi-experimental pilot study. *BMC Medicine,* 16**,** 64.

Claxton, K., Paulden, M., Gravelle, H., Brouwer, W. and Culyer, A. (2011). Discounting and decision making in the economic evaluation of health-care technologies. *Health Economics,* 20**,** 2-15.

Claxton, K., Sculpher, M. and Culyer, A. (2007) *Mark versus luke? Appropriate methods for the evaluation of public health interventions.* Available at: http://eprints.whiterose.ac.uk/140284/ (Accessed: 16th February 2020).

Claxton, K., Sculpher, M., Culyer, A., McCabe, C., Briggs, A., Akehurst, R., Buxton, M. and Brazier, J. (2006). Discounting and cost-effectiveness in nice – stepping back to sort out a confusion. *Health Economics,* 15**,** 1-4.

Coast, J. (2019). Assessing capability in economic evaluation: A life course approach? *The European Journal of Health Economics,* 20**,** 779-784.

Coast, J., Flynn, T., Natarajan, L., Sproston, K., Lewis, J., Louviere, J. and Peters, T. (2008a). Valuing the icecap capability index for older people. *Social Science & Medicine,* 67**,** 874-882.

Coast, J. and Jackson, L. 2017. Understanding primary data analysis. *In:* Coast, J. (ed.) *Qualitative methods for health economics.* London, UK: Rowman & Littlefield.

Coast, J., Peters, T., Natarajan, L., Sproston, K. and Flynn, T. (2008b). An assessment of the construct validity of the descriptive system for the icecap capability measure for older people. *Quality of Life Research,* 17**,** 967-976.

Coast, J., Smith, R. and Lorgelly, P. (2008c). Should the capability approach be applied in health economics? *Health Economics,* 17**,** 667-670.

Cochrane, M., Watson, P. M., Timpson, H., Haycox, A., Collins, B., Jones, L., Martin, A. and Graves, L. E. F. (2019). Systematic review of the methods used in economic evaluations of targeted physical activity and sedentary behaviour interventions. *Social Science & Medicine,* 232**,** 156-167.

Cohen, G. (1995). The Pareto argument for inequality. *Social Philosophy and Policy,* 12**,** p160.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences,* New York, Routledge.

Cohn, E. and Geske, T. (1990). *The economics of education,* Oxford, Pergamon.

Cole, T., Freeman, J. and Preece, M. (1995). Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood,* 73**,** 25-9.

Cole, T. and Pan, H. (2002). Lms growth computer program. 2.

COMET. (2017). *Defining core outcomes for school-based intervention studies on preventing childhood overweight and obesity* [Online]. Available: http://www.comet-initiative.org/studies/details/97 [Accessed 15th February 2020].

Cookson, R., Mirelman, A., Griffin, S., Asaria, M., Dawkins, B., Norheim, O., Verguet, S. and Culyer, A. (2017). Using cost-effectiveness analysis to address health equity concerns. *Value in Health,* 20**,** 206-212.

Corbin, J. and Strauss, A. (2015). *Basics of qualitative research,* California, USA, Sage Publications.

Cradock, A., Barrett, J., Carter, J., McHugh, A., Sproul, J., Russo, E., Dao-Tran, P. and Gortmaker, S. (2014). Impact of the boston active school day policy to promote physical activity among children. *American Journal of Health Promotion,* 28**,** S54-64.

Cradock, A., Barrett, J., Kenney, E., Giles, C., Ward, Z., Long, M., Resch, S., Pipito, A., Wei, E. and Gortmaker, S. (2017). Using cost-effectiveness analysis to prioritize policy and programmatic approaches to physical activity promotion and obesity prevention in childhood. *Preventive Medicine,* 95**,** S17-S27.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. and Petticrew, M. (2008). Developing and evaluating complex interventions: The new Medical Research Council guidance. *BMJ,* 337**,** a1655.

CRD. (2015). *Search strategies* [Online]. York. Available: http://www.crd.york.ac.uk/crdweb/searchstrategies.asp [Accessed 3 February 2017].

Cronbach, L. and Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin,* 52**,** 281-302.

Culyer, A. (1989). The normative economics of health care finance and provision. *Oxford Review of Economic Policy,* 5**,** 34-58.

Culyer, A. and Wagstaff, A. (1993). Equity and equality in health and health care. *Journal of Health Economics,* 12**,** 431-457.

Cummins, R. and Lau, A. (2005). *Personal wellbeing index – school children (pwi-sc),* Melbourne, Australia, School of Psychology, Deakin University.

Currie, G. R., Donaldson, C. and McIntosh, E. (1999). *Cost effectiveness' analysis and the" societal approach": Should the twain ever meet?*

Cutler, D. and Lleras-Muney, A. (2006) *Education and health: Evaluating theories and evidence.* Available at: https://www.nber.org/papers/w12352 (Accessed: 23 September 2019).

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century united states. *Demography,* 42**,** 1-22.

Dahl, E. (1994). Social inequalities in ill-health: The significance of occupational status, education and income-results from a Norwegian survey. *Sociology of Health and Illness,* 16**,** 644-667.

Darlington, E. J., Violon, N. and Jourdan, D. (2018). Implementation of health promotion programmes in schools: An approach to understand the influence of contextual factors on the process? *BMC public health,* 18**,** 163-163.

Davison, A. and Hinkley, D. (1997). *Bootstrap methods and their application,* Cambridge, UK, Cambridge University Press.

Day, R. E., Sahota, P. and Christian, M. S. (2019). Effective implementation of primary school-based healthy lifestyle programmes: A qualitative study of views of school staff. *BMC Public Health,* 19**,** 1239.

de Vries, M., Emons, W., Plantinga, A., Pietersma, S., van den Hout, W., Stiggelbout, A. and Elske van den Akker-van Marle, M. (2016). Comprehensively measuring health-related subjective well-being: Dimensionality analysis for improved outcome assessment in health economics. *Value in Health,* 19**,** 167-175.

Dearden, L., Machin, S. and Vignoles, A. (2009). Economics of education research: A review and future prospects. *Oxford Review of Education,* 35**,** 617-632.

DeBusk, R., Miller, N., Superko, H., Dennis, C., Thomas, R., Lew, H., Berger, W., Heller, R., Rompf, J., Gee, D., Kraemer, H., Bandura, A., Ghandour, G., Clark, M., Shah, R., Fisher, L. and Taylor, C. (1994). A case-management system for coronary risk factor modification after acute myocardial infarction. *Annals of Internal Medicine,* 120**,** 721-729.

Deidda, M., Geue, C., Kreif, N., Dundas, R. and McIntosh, E. (2019). A framework for conducting economic evaluations alongside natural experiments. *Social Science & Medicine,* 220**,** 353-361.

Department for Education. (1999) *Education (school day and school year) (England) regulations* Available at: http://www.legislation.gov.uk/uksi/1999/3181/contents/made (Accessed: 13 September 2019).

Department for Education. (2015) *Pe and sport premium for primary schools.* Available at: https://www.gov.uk/guidance/pe-and-sport-premium-for-primary-schools (Accessed: 16 May 2016).

Department for Education. (2016a) *Absence and attainment at key stages 2 and 4: 2013 to 2014.* Available at: https://www.gov.uk/government/publications/absence-and-attainment-at-key-stages-2-and-4-2013-to-2014 (Accessed: 12 August 2018).

Department for Education. (2016b) *Schools causing concern. Intervening in failing, underperforming and coasting schools.* Available at: https://dera.ioe.ac.uk//29882/ (Accessed: 27 November 2017).

Department for Education. (2017a) *Primary school accountability in 2017.* Available at: https://www.gov.uk/government/publications/primary-school-accountability (Accessed: 27 November 2017).

Department for Education. (2017b) *School teachers' pay and conditions document 2017 and guidance on school teachers' pay and conditions.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/636389/School_teachers__pay_and_conditions_document_2017.pdf (Accessed: 15 May 2018).

Department for Education. (2017c) *School workforce in England. .* Available at: https://www.gov.uk/government/statistics/school-workforce-in-england-november-2017 (Accessed: 13 September 2019).

Department for Education. (2018a) *Mental health and wellbeing provision in schools.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/747709/Mental_health_and_wellbeing_provision_in_schools.pdf (Accessed: 8th February 2020).

Department for Education. (2018b) *National curriculum assessments at key stage 2 in England, 2018* Available at: https://www.gov.uk/government/publications/national-curriculum-assessments-key-stage-2-2018-provisional/national-curriculum-assessments-at-key-stage-2-in-england-2018-provisional--2 (Accessed: 28 August 2019).

Department for Education. (2018c) *Primary school accountability in 2018.* Available at: https://www.gov.uk/government/publications/primary-school-accountability (Accessed: 11 August 2019).

Department for Education. (2019a) *Relationships education, relationships and sex education (rse) and health education.* Available at: https://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education (Accessed: 27 August 2019).

Department for Education. (2019b) *Schools, pupils and their characteristics.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_d ata/file/812539/Schools_Pupils_and_their_Characteristics_2019_Main_Text.pdf (Accessed: 30 August 2019).

Department for Education. (2019c) *Standards for school food in England.* Available at: https://www.gov.uk/government/publications/standards-for-school-food-in-england (Accessed: 27 August 2019).

Department for Education and Employment. (1997) *Excellence in schools.* Available at: http://www.educationengland.org.uk/documents/wp1997/excellence-in-schools.html (Accessed: 23 October 2017).

Department of Health. (2002) *Getting ahead of the curve: A strategy for combating infectious diseases (including other aspects of health protection).* Available at: https://webarchive.nationalarchives.gov.uk/+/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4007697 (Accessed: 23 August 2019).

Department of Health. (2009) *Healthy child programme from 5-19 years old.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/492086/HCP_5_to_19.pdf (Accessed: 20 April 2016).

Department of Health. (2013a) *Exposition book public health allocations 2013-14 and 2014-15: Technical guide* Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213324/Public-Health-Weighted-Capitation-FormulaTechnical-Guide-v0.13.pdf (Accessed: 11 July 2016).

Department of Health. (2013b) *Improving outcomes and supporting transparency: A public health outcomes framework for England, 2013-2016.* Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263658/2901502_PHOF_Improving_Outcomes_PT1A_v1_1.pdf (Accessed: 11 July 2016).

Department of Health. (2014) *Wellbeing and health policy.* Available at: https://www.gov.uk/government/publications/wellbeing-and-health-policy (Accessed: 6 August 2019).

Department of Health. (2015) *Nhs public health functions agreement 2016-17. Public health functions to be exercised by NHS England.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486444/S7A_1617.pdf (Accessed: 5 August 2016).

Department of Health and Social Care. (2015). *The NHS constitution for England* [Online]. Available: https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england [Accessed 22nd January 2020].

Detels, R., Beaglehole, R., Lansang, M. and Gulliford, M. (eds.) 2009. *Oxford textbook of public health,* Oxford: Oxford University Press.

Diener, E. and Suh, E. (1997). Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research,* 40**,** 189-216.

Dobbins, M., Husson, H., DeCorby, K. and LaRocca, R. L. (2013). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database of Systematic Reviews,* 2.

Dodge, R., Daly, A. P., Huyton, J. and Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing,* 2**,** 222-235.

Dolan, P., Gudex, C., Kind, P. and Williams, A. (1995) *A social tariff for EUROQOL: Results from a UK general population survey.* Available at: https://www.york.ac.uk/che/pdf/DP138.pdf (Accessed: 27 July 2017).

Dollahite, J. and Hosig, K. (1998). Impact of a school-based community intervention program on nutrition knwoledge and food choices. *Journal of Nutrition Education,* 30**,** 289.

Donaldson, C. (1990). The state of the art of costing health care for economic evaluation *Community Health Studies,* 14**,** 341-356.

Donaldson, C., Currie, G. and Mitton, C. (2002). Cost effectiveness analysis in health care: Contraindications. *BMJ,* 325**,** 891-894.

Doyle, O., Harmon, C., Heckman, J. and Tremblay, R. (2009). Investing in early human development: Timing and economic efficiency. *Economics and Human Biology,* 7**,** 1-6.

Drummond, M. and Jefferson, T. (1996). Guidelines for authors and peer reviewers of economic submissions to the bmj. *BMJ,* 313**,** 275.

Drummond, M., Sculpher, M., Torrance, G., O'Brien, B. and Stoddart, G. (2005). *Methods for the economic evaluation of health care programmes,* Oxford, Oxford University Press.

Dunlop, D. (1975). Benefit-cost analysis: A review of its applicability in policy analysis for delivering health services. *Social Science & Medicine,* 9**,** 133-139.

Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R. and Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development,* 82**,** 405-432.

Eckermann, S., Dawber, J., Yeatman, H., Quinsey, K. and Morris, D. (2014). Evaluating return on investment in a school based health promotion and prevention program: The investment multiplier for the stephanie alexander kitchen garden national program. *Social Science & Medicine* 114**,** 103-12.

Education Act. (2002) Available at: http://www.legislation.gov.uk/ukpga/2002/32/contents (Accessed: 8 August 2019).

Education and Skills Funding Agency. (2018) *PE and sport premium: Conditions of grant 2018 to 2019.* Available at: https://www.gov.uk/government/publications/pe-and-sport-premium-conditions-of-grant-2018-to-2019 (Accessed: 27 August 2019).

Education Reform Act. (1988) Available at: http://www.legislation.gov.uk/ukpga/1988/40/contents (Accessed: 8 August 2019).

Edwards, R., Charles, J. and Lloyd-Williams, H. (2013). Public health economics: A systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Public Health,* 13**,** 1001.

Edwards, R. T. and McIntosh, E. 2019. Applying methods of economic evaluation to public health: Contempory solutions to traditional challenges. *In:* Edwards, R. & Mcintosh, E. (eds.) *Applied health economics for public health practice and research.* Oxford: Oxford University Press.

Eichler, H., Kong, S., Gerth, W., Mavros, P. and Jonsson, B. (2004). Use of cost-effectiveness analysis in health-care resource allocation decision-making: How are cost-effectiveness thresholds expected to emerge? *Value in Health,* 7**,** 518-28.

Eiser, C. and Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of Life Research,* 10**,** 347-357.

Ekwaru, J., Ohinmaa, A., Tran, B., Setayeshgar, S., Johnson, J. and Veugelers, P. (2017). Cost-effectiveness of a school-based health promotion program in canada : A life-course modeling approach. *PLOS One,* 12.

EuroQol Group (1990). Euroqol-a new facility for the measurement of health-related quality of life. *Health Policy,* 16**,** 199-208.

Evans, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs,* 12**,** 77-84.

Eyles, A. and Machin, S. (2018). The introduction of academy schools to England's education. *Journal of the European Economic Association,* 17**,** 1107–1146.

Faculty of Public Health. (2014) *Functions of the local public health system.* Available at: http://www.fph.org.uk/uploads/Functions%20of%20the%20local%20PH%20system%20FINAL%20200514.pdf (Accessed: 5 August 2016).

Fayers, P. and Machin, D. (2016). *Quality of life: The assessment, analysis and reporting of patient-reported outcomes,* New Jersey, USA, John Wiley & Sons,.

Fegter, S., Machold, C. and Richter, M. (2010). Children and the good life: Theoretical challenges. *Children and the Good Life,* 4**,** 7-13.

Feng, Y., Jiang, R., Kohlmann, T. and Pickard, A. (2019). Exploring the internal structure of the EQ-5D using non–preference-based methods. *Value in Health,* 22**,** 527-536.

Fenwick, E., O'Brien, B. J. and Briggs, A. (2004). Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. *Health Economics,* 13**,** 405-415.

Ferraro, K., Schafer, M. and Wilkinson, L. (2016). Childhood disadvantage and health problems in middle and later life: Early imprints on physical health? *American Sociological Review,* 81**,** 107-133.

Flegal, K., Panagiotou, O. and Graubard, B. (2015). Estimating population attributable fractions to quantify the health burden of obesity. *Annals of Epidemiology,* 25**,** 201-207.

Flores, R. (1995). Dance for health: Improving fitness in African American and Hispanic adolescents. *Public Health Reports,* 110**,** 189-193.

Flynn, T. and Peters, T. (2005). Cluster randomized trials: Another problem for cost-effectiveness ratios. *International Journal of Technology Assessment in Health Care,* 21**,** 403-409.

Frame, P., Berg, A. and Woolf, S. (1997). U.S. Preventive services task force: Highlights of the 1996 report. *American Family Physician,* 55**,** 567-76, 581-2.

Frazer, K., Callinan, J., McHugh, J., van Baarsel, S., Clarke, A., Doherty, K. and Kelleher, C. (2016). Legislative smoking bans for reducing harms from secondhand smoke exposure, smoking prevalence and tobacco consumption. *Cochrane Database of Systematic Reviews*.

Frew, E. and Breheny, K. (2019a). Health economics methods for public health resource allocation: A qualitative interview study of decision makers from an english local authority. *Health Economics, Policy and Law*, 1-13.

Frew, E. and Breheny, K. (2019b). Methods for public health economic evaluation: A Delphi survey of decision makers in English and Welsh local government. *Health Economics,* 28, 1052-1063.

Frew, E., Wolstenholme, J. and Whynes, D. (2004). Comparing willingness-to-pay: Bidding game format versus open-ended and payment scale formats. *Health Policy,* 68, 289-98.

Frieden, T. R. (2017). Evidence for health decision making — beyond randomized, controlled trials. *New England Journal of Medicine,* 377, 465-475.

Friend, S., Flattum, C. F., Simpson, D., Nederhoff, D. M. and Neumark-Sztainer, D. (2014). The researchers have left the building: What contributes to sustaining school-based interventions following the conclusion of formal research support? *Journal of School Health,* 84, 326-333.

Fröhling , S. and Döhner, H. (2008). Chromosomal abnormalities in cancer. *New England Journal of Medicine,* 359, 722-734.

Gafni, A. and Birch, S. (1995). Preferences for outcomes in economic evaluation: An economic approach to addressing economic problems. *Social Science & Medicine,* 40, 767-776.

Gage, H., Kaye, J., Owen, C., Trend, P. and Wade, D. (2006). Evaluating rehabilitation using cost-consequences analysis: An example in parkinson's disease. *Clinical Rehabilitation,* 20, 232-238.

Garber, A., Weinstein, M., Torrance, G. and Kamlet, M. 1996. Theoretical foundations of cost-effectiveness analysis. *In:* Gold, M. & Siegel, J. (eds.) *Cost-effectiveness in health and medicine.* Oxford University Press, USA.

Garber, A. M. and Phelps, C. E. (1997). Economic foundations of cost-effectiveness analysis. *Journal of Health Economics,* 16, 1-31.

Gavin, B., Diana, G.-B., Naomi, M., Gillian, G., Lindsay, M. and Donncha, H. (2012). Physical activity, gender, weight status, and wellbeing in 9- to 11-year-old children: A cross-sectional survey. *Journal of Physical Activity and Health,* 9, 394-401.

Glaser, B. and Strauss, A. (2004). *The discovery of grounded theory: Strategies for qualitative research,* Piscataway, USA, Aldine Transaction.

Glick, H., Doshi, J., Sonnad, S. and Polsky, D. (2014). *Economic evaluation in clinical trials,* Oxford, Oxford University Press

Gold, M. and Siegel 1996. Cost-effectiveness in health and medicine. Oxford, UK: Oxford University Press.

Gordon-Larsen, P., Nelson, M., Page, P. and Popkin, B. (2006). Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics,* 117**,** 417-424.

Gortmaker, S., Wang, Y., Long, M., Giles, C., Ward, Z., Barrett, J., Kenney, E., Sonneville, K., Afzal, A., Resch, S. and Cradock, A. (2015). Three interventions that reduce childhood obesity are projected to save more than they cost to implement. *Health affairs (Project Hope),* 34**,** 1932-9.

Gottesman, M. and Collins, F. (1994). The role of the human genome project in disease prevention. *Preventive Medicine,* 23**,** 591-594.

Graham, H. (2004). Social determinants and their unequal distribution: Clarifying policy understandings. *The Milbank Quarterly,* 82**,** 101-124.

Graziose, M., Koch, P., Wang, Y., Lee Gray, H. and Contento, I. (2017). Cost-effectiveness of a nutrition education curriculum intervention in elementary schools. *Journal of Nutrition Education and Behavior,* 49**,** 684-691.e1.

Greco, G., Lorgelly, P. and Yamabhai, I. (2016). Outcomes in economic evaluations of public health interventions in low and middle income countries: Health, capabilities and subjective wellbeing. *Health Economics,* 25**,** 83-94.

Gregory, T., Engelhardt, D., Lewkowicz, A., Luddy, S., Guhn, M., Gadermann, A., Schonert-Reichl, K. and Brinkman, S. (2018). Validity of the Middle Years Development Instrument for population monitoring of student wellbeing in australian school children. *Child Indicators Research,* 12**,** 873-899.

Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J. and Coast, J. (2006). Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine,* 62**,** 1891-1901.

Guerriero, C., Cairns, J., Bianchi, F. and Cori, L. (2018). Are children rational decision makers when they are asked to value their own health? A contingent valuation study conducted with children and their parents. *Health Economics,* 27**,** e55-e68.

Gutman, M. and Vorhaus, J. (2012) *The impact of pupil behaviour and wellbeing on educational outcomes.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219638/DFE-RR253.pdf (Accessed: 29th August 2019).

Handley, M., Shumway, M. and Schillinger, D. (2008). Cost-effectiveness of automated telephone self-management support with nurse care management among patients with diabetes. *Annals of Family Medicine,* 6**,** 512-518.

Harris, J. (1987). Qalyfying the value of life. *Journal of Medical Ethics,* 13**,** 117-123.

Hayes, A., Tan, E. J., Lung, T., Brown, V., Moodie, M. and Baur, L. (2019). A new model for evaluation of interventions to prevent obesity in early childhood. *Front Endocrinol (Lausanne),* 10**,** 132.

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science,* 312**,** 1900.

Hendy, H., Williams, K. and Camise, T. (2011). Kid's choice program improves weight management behaviors and weight status in school children. *Appetite,* 56**,** 484-94.

Henkel, D. (2011). Unemployment and substance use: A review of the literature (1990-2010). *Current Drug Abuse Reviews,* 4**,** 4-27.

Hillger, C. 2008. Lifestyle and health determinants. *In:* Kirch, W. (ed.) *Encyclopedia of public health.* Dordrecht: Springer Netherlands.

Hirschhorn, K. and Cooper, H. (1961). Chromosomal aberrations in human disease. *The American Journal of Medicine,* 31**,** 442-470.

HM Government. (2018a) *Childhood obesity: A plan for action. Chapter 2.* Available at: https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action/childhood-obesity-a-plan-for-action (Accessed: 11 August 2019).

HM Government. (2018b) *Equality assessment: Childhood obesity: A plan for action. Chapter 2.* Available at: https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action/childhood-obesity-a-plan-for-action (Accessed: 28 August 2019).

HM Revenue and Customs. (2016) *Soft drinks industry levy.* Available at: https://www.gov.uk/government/publications/soft-drinks-industry-levy (Accessed: 27 August 2019).

HM Treasury. (2003) *The green book.* Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf (Accessed: 17 May 2017).

HM Treasury 2018. Sugar tax revenue helps tackle childhood obesity.

Holder, J., Beecham and Knapp, E. (2011) *Developing a wellbeing outcome measure for use in economic evaluations of children's services: Identifying domains important to children and young people.* Available at: http://www.cwrc.ac.uk/resources/documents/Developing_CYPOC_Report_Summary_March_2011.pdf (Accessed: 27 August 2019).

Hollands, F. M., Kieffer, M. J., Shand, R., Pan, Y., Cheng, H. and Levin, H. M. (2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. *Journal of Research on Educational Effectiveness,* 9**,** 30-53.

Horowitz, J. and McConnell, K. (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management,* 44**,** 426-447.

Horsman, J., Furlong, W., Feeny, D. and Torrance, G. (2003). The health utilities index (hui®): Concepts, measurement properties and applications. *Health and Quality of Life Outcomes,* 1**,** 54.

Huebner, E. (1991). Initial development of the student's life satisfaction scale. *Journal of School Psychology International,* 12**,** 231-240.

Huguet, A., Stinson, J. and McGrath, P. (2010). Measurement of self-reported pain intensity in children and adolescents. *Journal of Psychosomatic Research,* 68**,** 329-336.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A. H., Mauskopf, J. and Loder, E. (2013). Consolidated health economic evaluation reporting standards (cheers) statement. 346**,** f1049.

Jain, R., Grabner, M. and Onukwugha, E. (2011). Sensitivity analysis in cost-effectiveness studies. *PharmacoEconomics,* 29**,** 297-314.

Janssens, A., Rogers, M., Thompson Coon, J., Allen, K., Green, C., Jenkinson, C., Tennant, A., Logan, S. and Morris, C. (2015). A systematic review of generic multidimensional patient-reported outcome measures for children, part ii: Evaluation of psychometric performance of english-language versions in a general population. *Value in Health,* 18**,** 334-345.

Jessiman, P. E., Campbell, R., Jago, R., Van Sluijs, E. M. F. and Newbury-Birch, D. (2019). A qualitative study of health promotion in academy schools in England. *BMC Public Health,* 19**,** 1186.

Johannesson, M. and O'Conor, R. (1997). Cost-utility analysis from a societal perspective. *Health Policy,* 39**,** 241-253.

John, J., Koerber, F. and Schad, M. (2019). Differential discounting in the economic evaluation of healthcare programs. *Cost Effectiveness and Resource Allocation,* 17**,** 29.

Johns, B., Baltussen, R. and Hutubessy, R. (2003). Programme costs in the economic evaluation of health interventions. *Cost Effectiveness and Resource Allocation,* 1**,** 1-10.

Joseph, P., Craig, J. and Caldwell, P. (2015). Clinical trials in children. *British Journal of Clinical Pharmacology,* 79**,** 357-369.

Juonala , M., Magnussen, C., Berenson, G., Venn , A., Burns , T., Sabin , M., Srinivasan , S., Daniels , S., Davis , P., Chen , W., Sun , C., Cheung , M., Viikari , J., Dwyer , T. and Raitakari , O. (2011). Childhood adiposity, adult adiposity, and cardiovascular risk factors. *New England Journal of Medicine,* 365**,** 1876-1885.

Jupp, V. (2006). *The sage dictionary of social research methods,* London, Sage.

Kaye-Tzadok, A., Kim, S. S. and Main, G. (2017). Children's subjective well-being in relation to gender — what can we learn from dissatisfied children? *Children and Youth Services Review,* 80**,** 96-104.

Kelly, M., McDaid, D., Ludbrook, A. and Powell, J. (2005) *Economic appraisal of public health interventions.* Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.3820&rep=rep1&type=pdf (Accessed: 28 October 2016).

Kendall, L. (2005) *Excellence in cities: The national evaluation of a policy to raise standards in urban schools 2000-2003.* Available at: https://www.nfer.ac.uk/publications/EIC14/EIC14.pdf (Accessed: 23 October 2017).

Kerr, C., Breheny, K., Lloyd, A., Brazier, J., Bailey, D., Berry-Kravis, E., Cohen, J. and Petrillo, J. (2015). Developing a utility index for the aberrant behavior checklist (abc-c) for fragile x syndrome. *Quality of Life Research,* 24**,** 305-14.

Kesztyus, D., Lauer, R., Kesztyus, T., Kilian, R., Steinacker, J. and Join the Healthy Boat" Study Group (2017). Costs and effects of a state-wide health promotion program in primary schools in Germany - the Baden-Wurttemberg study: A cluster-randomized, controlled trial. *PLOS One,* 12**,** e0172332.

Kesztyus, D., Schreiber, A., Wirt, T., Wiedom, M., Dreyhaupt, J., Brandstetter, S., Koch, B., Wartha, O., Muche, R., Wabitsch, M., Kilian, R. and Steinacker, J. (2013). Economic evaluation of urmel-ice, a school-based overweight prevention programme comprising metabolism, exercise and lifestyle intervention in children. *The European Journal of Health Economics,* 14**,** 185-195.

Khoury, M. (1997). Genetic epidemiology and the future of disease prevention and public health. *Epidemiologic Reviews,* 19**,** 175-80.

Kline, P. (1979). *Psychometrics and psychology,* London, UK, Academic Press.

Knapp, M. (1984). *The economics of social care,* London, UK, Macmillan.

Kneale, D., Rojas-García, A. and Thomas, J. (2019). Obstacles and opportunities to using research evidence in local public health decision-making in England. *Health Research Policy and Systems,* 17**,** 61.

Korber, K. (2015). Quality assessment of economic evaluations of health promotion programs for children and adolescents—a systematic review using the example of physical activity. *Health Economics Review,* 5**,** 35.

Kreimeier, S. and Greiner, W. (2019). EQ-5D-Y as a health-related quality of life instrument for children and adolescents: The instrument's characteristics, development, current use, and challenges of developing its value set. *Value in Health,* 22**,** 31-37.

Kreimeier, S., Oppe, M., Ramos-Goñi, J. M., Cole, A., Devlin, N., Herdman, M., Mulhern, B., Shah, K. K., Stolk, E., Rivero-Arias, O. and Greiner, W. (2018). Valuation of euroqol five-dimensional questionnaire, youth version (EQ-5D-Y) and euroqol five-dimensional questionnaire, three-level version (EQ-5D-3L) health states: The impact of wording and perspective. *Value in Health,* 21**,** 1291-1298.

Kwon, J., Kim, S., Ungar, W., Tsiplova, K., Madan, J. and Petrou, S. (2019). Patterns, trends and methodological associations in the measurement and valuation of childhood health utilities. *Quality of Life Research,* 28**,** 1705-1724.

Ladapo, J., Bogart, L., Klein, D., Cowgill, B., Uyeda, K., Binkle, D., Stevens, E. and Schuster, M. (2016). Cost and cost-effectiveness of students for nutrition and exercise (SNAX). *Academic Pediatrics,* 16**,** 247-53.

Lahelma, E., Martikainen, P., Laaksonen, M. and Aittomäki, A. (2004). Pathways between socioeconomic determinants of health. *Journal of Epidemiology and Community Health,* 58**,** 327-332.

Lamsal, R. and Zwicker, J. (2017). Economic evaluation of interventions for children with neurodevelopmental disorders: Opportunities and challenges. *Applied Health Economics and Health Policy,* 15**,** 763-772.

Landefeld, J. and Seskin, E. (1982). The economic value of life: Linking theory to practice. *American Journal of Public Health,* 72**,** 555-566.

Lang, J. and Connell, C. (2017). Measuring costs to community-based agencies for implementation of an evidence-based practice. *The Journal of Behavioral Health Services & Research,* 44**,** 122-134.

Lawlor, D., Kipping, R., Anderson, E., Howe, L., Chittleborough, C., Moure-Fernandez, A., Noble, S., Rawlins, E., Wells, S. L., Peters, T. J., Jago, R. and Campbell, R. (2016). Active for life year 5: A cluster randomised controlled trial of a primary school-based intervention to increase levels of physical activity, decrease sedentary behaviour and improve diet. *Public Health Research,* 4.

Layard, R. (2016) *Measuring wellbeing and cost-effectiveness analysis.* Available at: https://whatworkswellbeing.org/product/measuring-wellbeing-and-cost-effectiveness-analysis-using-subjective-wellbeing/ (Accessed: 6 August 2019).

Le Grand, J. (1978). The distribution of public expenditure: The case of health care. *Economica,* 45**,** 125-142.

Leck, C., Upton, D. and Evans, N. (2016). Social return on investment: Valuing health outcomes or promoting economic values? *Journal of Health Psychology,* 21**,** 1481-90.

Lee, A. and Maheswaran, R. (2011). The health benefits of urban green spaces: A review of the evidence. *Journal of Public Health,* 33**,** 212-222.

Levin, A., Wang, S., Levin, C., Tsu, V. and Hutubessy, R. (2014). Costs of introducing and delivering hpv vaccines in low and lower middle income countries: Inputs for gavi policy on introduction grant support to countries. *PLOS One,* 9**,** e101114.

Levin, H. (1988). Cost-effectiveness and educational policy. *Educational Evaluation and Policy Analysis,* 10**,** 51-69.

Levin, H. and Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness,* 8**,** 400-418.

Li, X., Jan, S., Yan, L., Hayes, A., Chu, Y., Wang, H., Feng, X., Niu, W., He, F., Ma, J., Han, Y., MacGregor, G. and Wu, Y. (2017). Cost and cost-effectiveness of a school-based education program to reduce salt intake in children and their families in China. *PLOS One,* 12**,** e0183033.

Liddle, I. and Carter, G. (2015). Emotional and psychological well-being in children: The development and validation of the stirling children's well-being scale. *Educational Psychology in Practice,* 31**,** 174-185.

Lippman, L., Moore, K. and McIntosh, H. (2011). Positive indicators of child well-being: A conceptual framework, measures, and methodological issues. *Applied Research in Quality of Life,* 6**,** 425-449.

Littlecott, H., Long, S., Hawkins, J., Murphy, S., Hewitt, G., Eccles, G., Fletcher, A. and Moore, G. F. (2018). Health improvement and educational attainment in secondary schools: Complementary or competing priorities? Exploratory analyses from the school health research network in Wales. *Health Education Behavior,* 45**,** 635-644.

Lloyd, A., Kerr, C., Breheny, K., Brazier, J., Ortiz, A. and Borg, E. (2014). Economic evaluation in short bowel syndrome (sbs): An algorithm to estimate utility scores for a patient-reported sbs-specific quality of life scale (SBS-QOL™). *Quality of Life Research,* 23**,** 449-458.

Lobo, I. (2008). Multifactorial inheritance and genetic disease. *Nature Education,* 1**,** 5.

Local Government Association. (2018) *Improving children and young people's mental health and emotional wellbeing.* Available at: https://www.local.gov.uk/sites/default/files/documents/Improving%20children%20and%20young%20people%E2%80%99s%20mental%20health%20and%20emotional%20wellbeing.pdf (Accessed: 8th February 2020).

Long, K., Brown, J., Jones, S., Aber, J. and Yates, B. (2015). Cost analysis of a school-based social and emotional learning and literacy intervention. *Journal of Benefit-Cost Analysis,* 6**,** 545-571.

Maciosek, M., LaFrance, A., Dehmer, S., McGree, D., Xu, Z., Flottemesch, T. and Solberg, L. (2017). Health benefits and cost-effectiveness of brief clinician tobacco counseling for youth and adults. *Annals of Family Medicine,* 15**,** 37-47.

Makai, P., Brouwer, W., Koopmanschap, M., Stolk, E. and Nieboer, A. (2014). Quality of life instruments for economic evaluations in health and social care for older people: A systematic review. *Social Science & Medicine,* 102**,** 83-93.

Malden, S. and Doi, L. (2019). The Daily Mile: Teachers' perspectives of the barriers and facilitators to the delivery of a school-based physical activity intervention. *BMJ Open,* 9**,** e027169.

Manger, W., Manger, L., Minno, A., Killmeyer, M., Holzman, R., Schullinger, J. and Roccella, E. (2012). Obesity prevention in young schoolchildren: Results of a pilot study. *Journal of School Health,* 82**,** 462-8.

Mangunkusumo, R., Duisterhout, J., De Graaff, N., Maarsingh, E., De Koning, H. and Raat, H. (2006). Internet versus paper mode of health and health behavior questionnaires in elementary schools: Asthma and fruit as examples. *Journal of School Health,* 76**,** 80-86.

Mann, J. (1997). Medicine and public health, ethics and human rights. *Hastings Center Report,* 27**,** 6-13.

Markussen, K., Dalsgaard, S., Obel, C., Wisborg, W., Henriksen, T., Rodriguez, A., Kotimaa, A., Moilanen, I., Thomsen, P., Olsen, J. and Jarvelin, M. (2003). Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: Review of the current evidence. *American Journal of Psychiatry,* 160**,** 1028-1040.

Marmot, M. (2005). Social determinants of health inequalities. *The Lancet,* 365**,** 1099-1104.

Marmot, M. (2010) *Fair society, healthy lives: Strategic review of health inequalities in England post-2010. .* Available at: http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review (Accessed: 11 July 2016).

Marmot, M. and Bell, R. (2012). Fair society, healthy lives. *Public Health,* 126**,** S4-S10.

Marmot, M., Friel, S., Bell, R., Houweling, T. and Taylor, S. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet,* 372**,** 1661-1669.

Marufu, T. C., Ahankari, A., Coleman, T. and Lewis, S. (2015). Maternal smoking and the risk of still birth: Systematic review and meta-analysis. *BMC Public Health,* 15.

Masters, R., Anwar, E., Collins, B., Cookson, R. and Capewell, S. (2017). Return on investment of public health interventions: A systematic review. *Journal of Epidemiology and Community Health,* 71**,** 827-834.

Mathes, T., Walgenbach, M., Antoine, S., Pieper, D. and Eikermann, M. (2014). Methods for systematic reviews of health economic evaluations: A systematic review, comparison, and synthesis of method literature. *Medical Decision Making,* 34**,** 826-840.

Mauskopf, J., Paul, J., Grant, D. and Stergachis, A. (1998). The role of cost-consequence analysis in healthcare decision making. *Pharmacoeconomics,* 13**,** 277-288.

McAuley, K., Taylor, R., Farmer, V., Hansen, P., Williams, S., Booker, C. and Mann, J. (2010). Economic evaluation of a community-based obesity prevention program in children: The apple project. *Obesity* 18**,** 131-6.

McGlothlin, A. and Lewis, R. (2014). Minimal clinically important difference: Defining what really matters to patients. *JAMA,* 312**,** 1342-1343.

McIntosh, E. 2019. Identifying, measuring and valuing outcomes within economic evaluations of public health interventions. *In:* Edwards, R. & Mcintosh, E. (eds.) *Applied health economics for public health practice and research.* Oxford: Oxford University Press.

McIntosh, E. and Luengo-Fernandez, R. (2006). Economic evaluation. Part 2: Frameworks for combining costs and benefits in health care. *J Fam Plann Reprod Health Care,* 32**,** 176-80.

McIntosh, E. and Ryan, M. (2002). Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. *Journal of Economic Psychology,* 23**,** 367-382.

McKeown, T. (1979). *The role of medicine: Dream, mirage or nemesis?,* Oxford, Blackwell.

McLellan, R. and Steward, S. (2015). Measuring children and young people's wellbeing in the school context. *Cambridge Journal of Education,* 45**,** 307-332.

Meng, L., Xu, H., Liu, A., van Raaij, J., Bemelmans, W., Hu, X., Zhang, Q., Du, S., Fang, H., Ma, J., Xu, G., Li, Y., Guo, H., Du, L. and Ma, G. (2013). The costs and cost-effectiveness of a school-based comprehensive intervention study on childhood obesity in china. *PLOS One,* 8**,** e77971.

Mernagh, P., Paech, A., Coleman, K., Weston, A., McDonald, J., Cumming, J. and Green, T. (2010) *Assessing the cost-effectiveness of public health interventions to prevent obesity: Overview report.* Available at: https://pdfs.semanticscholar.org/8e44/72224093e93b993a4c198743e68425a226c4.pdf (Accessed: 21 January 2019).

Meyers, D., Neuberger, J. and He, J. (2009). Cardiovascular effect of bans on smoking in public places a systematic review and meta-analysis. *Journal of the American College of Cardiology,* 54**,** 1249-1255.

Milton, W. (1981). Economic assessments of medical practices and technologies. *Medical Decision Making,* 1**,** 309-330.

Mishan, E. and Quah, E. (1971). *Cost-benefit analysis,* London, Routledge.

Mogyorosy, Z. and Smith, P. (2005) *The main methodological issues in costing health care services: A literature review.* Available at: https://www.york.ac.uk/media/che/documents/papers/researchpapers/rp7_Methodological_issues_in_costing_health_care_services.pdf (Accessed: 27 October 2016).

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M. and de Vet, H. C. (2010). The cosmin study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology,* 63**,** 737-45.

Mons, U., Müezzinler, A., Gellert, C., Schöttker, B., Abnet, C., Bobak, M., de Groot, L., Freedman, N., Jansen, E., Kee, F., Kromhout, D., Kuulasmaa, K., Laatikainen, T., O'Doherty, M., Bueno-de-Mesquita, B., Orfanos, P., Peters, A., van der Schouw, Y., Wilsgaard, T., Wolk, A., Trichopoulou, A., Boffetta, P. and Brenner, H. (2015). Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: Meta-analysis of individual participant data from prospective cohort studies of the chances consortium. *BMJ,* 350.

Moodie, M., Haby, M., Swinburn, B. and Carter, R. (2011). Assessing cost-effectiveness in obesity: Active transport program for primary school children--travelsmart schools curriculum program. *Journal of Physical Activity and Health,* 8**,** 503-15.

Moodie, M., Herbert, J., de Silva-Sanigorski, A., Mavoa, H., Keating, C., Carter, R., Waters, E., Gibbs, L. and Swinburn, B. (2013). The cost-effectiveness of a successful community-based obesity prevention program: The be active eat well program. *Obesity* 21**,** 2072-80.

Mooney, G. (1983). Equity in health care: Confronting the confusion. *Effective Health Care,* 1**,** 179-85.

Mooney, G. (2000). Vertical equity in health care resource allocation. *Health Care Analysis,* 8**,** 203-215.

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D. and Baird, J. (2015). Process evaluation of complex interventions: Medical research council guidance. *BMJ,* 350**,** h1258.

Moore, J., Yin, Z., Hanes, J., Duda, J., Gutin, B. and Barbeau, P. (2009). Measuring enjoyment of physical activity in children: Validation of the physical activity enjoyment scale. *Journal of Applied Sport Psychology,* 21**,** S116-S129.

Mora, T., Llargues, E. and Recasens, A. (2015). Does health education affect BMI? Evidence from a school-based randomised-control trial. *Economics and Human Biology,* 17**,** 190-201.

Morris, S., Devlin, N. and Parkin, D. (2007). *Economic analysis in health care,* Chichester, UK, John Wiley & Sons.

Morton, K. L., Atkin, A. J., Corder, K., Suhrcke, M., Turner, D. and van Sluijs, E. M. F. (2017). Engaging stakeholders and target groups in prioritising a public health intervention: The creating active school environments (CASE) online delphi study. *BMJ Open,* 7**,** e013340.

Mukuria, C., Rowen, D., Peasgood, T. and Brazier, J. (2016) *An empirical comparison of well-being measures used in the UK.* Available at: http://eprints.whiterose.ac.uk/99499/ (Accessed: 23 May 2019).

Natarajan, A. (2002). Chromosome aberrations: Past, present and future. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 504**,** 3-16.

National Research Council (1984). *Development during middle childhood: The years from six to twelve,* Washington, USA, The National Academies Press.

Neumann, P., Ganiats, T., Russell, L., Sanders, G. and Siegel, J. (eds.) 2016. *Cost-effectiveness in health and medicine,* Oxford, UK: Oxford University Press.

Ng, E., Grieve, R. and Carpenter, J. (2013). Two-stage nonparametric bootstrap sampling with shrinkage correction for clustered data. *The Stata Journal,* 13**,** 141-164.

NHS Digital. (2018). National child measurement programme. https://digital.nhs.uk/services/national-child-measurement-programme/

NICE. (2011) *Supporting investment in public health: Review of methods for assessing cost effectiveness, cost impact and return on investment.* Available at: https://www.nice.org.uk/media/default/About/what-we-do/NICE-guidance/NICE-guidelines/Public-health-guidelines/Additional-publications/Cost-impact-proof-of-concept.pdf

NICE. (2014). *Incorporating economic evaluation.* Developing NICE Guidelines: The Manual [PMG20]. https://www.nice.org.uk/process/pmg20/chapter/incorporating-economic-evaluation .

Nicoll, A. and Murray, V. (2002). Health protection - a strategy and a national agency. *Public Health,* 116**,** 129-137.

Nicoll, A., Wilson, D., Calvert, N. and Borriello, P. (2001). Managing major public health crises. *BMJ,* 323**,** 1321.

NIHR Dissemination Centre. (2019) *Moving matters. Interventions to increase physical activity.* Available at: https://www.dc.nihr.ac.uk/themed-reviews/research-into-physical-activity.htm (Accessed: 9 August 2019).

Nutbeam, D. (2003). How does evidence influence public health policy? Tackling health inequalities in England. *Health Promotion Journal of Australia,* 14**,** 154-158.

Office for National Statistics. (2018) *Understanding well-being inequalities: Who has the poorest personal well-being?* Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/understandingwellbeinginequalitieswhohasthepoorestpersonalwellbeing/2018-07-11 (Accessed: 23 May 2019).

OFSTED. (2012) *Not yet good enough: Personal, social, health and economic education in schools.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/413178/Not_yet_good_enough_personal__social__health_and_economic_education_in_schools.pdf (Accessed: 01 February 2016).

OFSTED. (2013) *Beyond 2012 - outstanding physical activity for all.* Available at: https://www.gov.uk/government/publications/beyond-2012-outstanding-physical-education-for-all (Accessed: 19 October 2017).

OFSTED. (2019) *School inspection handbook.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/843108/School_inspection_handbook_-_section_5.pdf (Accessed: 9th February 2020).

Ohinmaa, A., Langille, J., Jamieson, S., Whitby, C. and Veugelers, P. (2011). Costs of implementing and mintaining comprehensive school health : The case of the Annapolis valley health promoting schools program. *Canadian Journal of Public Health,* 102.

Olsen, J. and Smith, R. (2001). Theory versus practice: A review of 'willingness-to-pay' in health and health care. *Health Economics,* 10**,** 39-52.

Oosterhoff, M., Bosma, H., van Schayck, O., Evers, S., Dirksen, C. and Joore, M. (2018). A systematic review on economic evaluations of school-based lifestyle interventions targeting weight-related behaviours among 4–12 year olds: Issues and ways forward. *Preventive Medicine,* 114**,** 115-122.

Palmer, S. and Raftery, J. (1999). Opportunity cost. *BMJ,* 318**,** 1551-1552.

Palmer, S. and Torgerson, D. (1999). Definitions of efficiency. *BMJ,* 318**,** 1136.

Parmenter, K., Waller, J. and Wardle, J. (2000). Demographic variation in nutrition knowledge in England. *Health Education Research,* 15**,** 163-174.

Patalay, P. and Fitzsimons, E. (2016). Correlates of mental illness and wellbeing in children: Are they the same? Results from the UK millennium cohort study. *Journal of the American Academy of Child & Adolescent Psychiatry,* 55**,** 771-783.

Peasgood, T., Brazier, J., Mukuria, C. and Rowen, D. (2014) *A conceptual comparison of well-being measures used in the UK.* Available at: http://eprints.whiterose.ac.uk/99497/1/A%20conceptual%20comparison%20of%20%20-%20026.pdf (Accessed: 8th February 2020).

Peasgood, T., Carlton, J. and Brazier, J. (2019). A qualitative study of the views of health and social care decision-makers on the role of wellbeing in resource allocation decisions in the UK. *Economies,* 7.

Peasgood, T. and E-QALY team 2018. Extending the QALY (e-QALY) project: Stages for developing a new generic measure. *PROMs Conference.* Birmingham.

Petersen, K., Chen, G., Mpundu-Kaambwa, C., Stevens, K., Brazier, J. and Ratcliffe, J. (2018). Measuring health-related quality of life in adolescent populations: An empirical comparison of the CHU9D and the pedsqltm 4.0 short form 15. *The Patient - Patient-Centered Outcomes Research,* 11**,** 29-37.

Petrou, S. and Gray, A. (2011). Economic evaluation using decision analytical modelling: Design, conduct, analysis, and reporting. *BMJ,* 342.

PHE. (2014) *The link between pupil health and wellbeing and attainment: A briefing for head teachers, governors and staff in education settings.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/370686/HT_briefing_layoutvFINALvii.pdf (Accessed: 12th August 2019).

PHE. (2014) *Maximising the school aged children nursing team contribution to the public health of school-aged children.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/303769/Service_specifications.pdf (Accessed: 20 April 2016).

PHE. (2015) *Measuring mental wellbeing in children and young people: Technical appendix.* Available at: https://www.gov.uk/government/publications/measuring-the-mental-wellbeing-of-children-and-young-people (Accessed: 19 September 2019).

PHE. (2016a) *Fit for the future: Public health people. A review of the public health workforce* Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/524599/Fit_for_the_Future_Report.pdf (Accessed: 13 November 2017).

PHE. (2016b) *Sexual and reproductive health in England: A guide to local and national data.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/395207/PHE_Guide_to_National_Local_Sexual_Reproductive_Health_Data.pdf (Accessed: 24 August 2016).

PHE. (2017) *National child measurement programme. Guidance for data sharing and analysis 2015/2016.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624226/PHE_NCMP_analysis_guidance_2015_to_2016.pdf (Accessed: 1 September 2019).

PHE. (2018a) *Guidance. Health matters: Health economics - making the most of your budget.* Available at: https://www.gov.uk/government/publications/health-matters-health-economics-making-the-most-of-your-budget/health-matters-health-economics-making-the-most-of-your-budget

PHE. (2018b) *Health profile for England. Inequalities in health.* Available at: https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health (Accessed: 10 August 2019).

PHE. (2019a) *Phe strategy 2020-25.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/831562/PHE_Strategy_2020-25.pdf (Accessed: 20 August 2019).

PHE. (2019b) *Universal approaches to improving children and young people's mental health and wellbeing.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/842176/SIG_report.pdf (Accessed: 8th February 2020).

Pinto, S., Fumincelli, L., Mazzo, A., Caldeira, S. and Martins, J. (2017). Comfort, well-being and quality of life: Discussion of the differences and similarities among the concepts. *Porto Biomedical Journal,* 2**,** 6-12.

Pirie, K., Peto, R., Reeves, G., Green, J. and Beral, V. (2013). The 21st century hazards of smoking and benefits of stopping: A prospective study of one million women in the UK. *The Lancet,* 381**,** 133-141.

Pollard, E. L. and Lee, P. D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research,* 61**,** 59-78.

Portrait, F., Hoozemans, M. and van den Berg, B. (2019). The total value of time of children undergoing treatment: A contingent valuation from the perspective of parents in the orthopaedic department of a dutch hospital. *Journal of Paediatrics and Child Health,* 55**,** 539-547.

Public Health Agency. (2019) *Public health agency annual report and accounts for the year engded 30 march 2019.* Available at: https://www.publichealth.hscni.net/sites/default/files/2019-07/PHA%20ACCOUNTS%202018-2019.pdf (Accessed: 3 August 2019).

Public Health Wales. (2018) *Our strategic plan 2019-2022.* Available at: https://phw.nhs.wales/about-us/our-priorities/long-term-strategy-documents/public-health-wales-strategic-plan-2019-22/ (Accessed: 2 September 2019).

Puffer, S., Torgerson, D. and Watson, J. (2005). Cluster randomized controlled trials. *Journal of Evaluation in Clinical Practice,* 11**,** 479-483.

Pyke, C. 1998. Costing and pricing in public sector. *In:* Wilson, J. (ed.) *Financial management for the public services.* Oxford, UK: Oxford University Press.

Nvivo,. 11. QSR International

Rafferty, R., Breslin, G., Brennan, D. and Hassan, D. (2016). A systematic review of school-based physical activity interventions on children’s wellbeing. *International Review of Sport and Exercise Psychology,* 9**,** 215-230.

Ramey, C. and Ramey, S. (1998). Early intervention and early experience. *American Psychologist,* 53**,** 109.

Rappange, D., Brouwer, W., Rutten, F. and van Baal, P. (2010). Lifestyle intervention: From cost savings to value for money. *Journal of Public Health,* 32**,** 440-447.

Ratcliffe, J., Flynn, T., Terlich, F., Stevens, K., Brazier, J. and Sawyer, M. (2012). Developing adolescent-specific health state values for economic evaluation: An application of profile case best-worst scaling to the child health utility 9d. *Pharmacoeconomics,* 30**,** 713-727.

Rebok, G., Riley, A., Forrest, C., Starfield, B., Green, B., Robertson, J. and Tambor, E. (2001). Elementary school-aged children's reports of their health: A cognitive interviewing study. *Quality of Life Research,* 10**,** 59-70.

Rechel, B. and McKee, M. (eds.) 2014. *Facets of public health in europe,* Maidenhead, UK: Open University Press.

Rees, G., Goswami, H. and Bradshaw, J. (2010) *Developing an index of children's subjective well-being in England.* Available at: https://www.childrenssociety.org.uk/sites/default/files/tcs/research_docs/Developing%20an%20Index%20of%20Children's%20Subjective%20Well-being%20in%20England.pdf (Accessed: 19 September 2019).

Reznik, M., Wylie-Rosett, J., Kim, M. and Ozuah, P. (2015). A classroom-based physical activity intervention for urban kindergarten and first-grade students: A feasibility study. *Childhood Obesity,* 11**,** 314-324.

Rice, D. (1967). Estimating the cost of illness. *American Journal of Public Health and the Nation's Health,* 57**,** 424-440.

Rice, D., Hodgson, T. and Kopstein, A. (1985). The economic costs of illness: A replication and update. *Health Care Financing Review,* 7**,** 61-80.

Ritzwoller, D., Sukhanova, A., Gaglio, B. and Glasgow, R. (2009). Costing behavioral interventions: A practical guide to enhance translation. *Annals of Behavioral Medicine,* 37**,** 218-227.

Roberts, J. and Haworth, E. (2002). Getting ahead of the curve: A strategy and its implementation. *The Lancet Infectious Diseases,* 2**,** 205-206.

Robinson, R. (1993a). Cost-benefit analysis. *BMJ,* 307**,** 924.

Robinson, R. (1993b). Costs and cost-minimisation analysis. *BMJ,* 307**,** 726-728.

Rose, K., Lake, A. A., Ells, L. J. and Brown, L. (2019). School food provision in England: A historical journey. *Nutrition Bulletin,* 44**,** 283-291.

Rush, E., Obolonkin, V., McLennan, S., Graham, D., Harris, J., Mernagh, P. and Weston, A. (2014). Lifetime cost effectiveness of a through-school nutrition and physical programme: Project energize. *Obesity Research & Clinical Practice,* 8**,** e115-22.

Ryan, M., Watson, V. and Amaya-Amaya, M. (2003). Methodological issues in the monetary valuation of benefits in healthcare. *Expert Review of Pharmacoeconomics and Outcomes Research,* 3**,** 717-727.

Rychetnik, L., Frommer, M., Hawe, P. and Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health,* 56**,** 119-27.

Ryde, G., Booth, J., Brooks, N., Chesham, R., Moran, C. and Gorely, T. (2018). The Daily Mile: What factors are associated with its implementation success? *PLOS One,* 13**,** e0204988.

Sahoo, K., Sahoo, B., Choudhury, A., Sofi, N., Kumar, R. and Bhadoria, A. (2015). Childhood obesity: Causes and consequences. *Journal of Family Medicine and Primary Care,* 4**,** 187-192.

Saldana, L., Chamberlain, P., Bradford, W., Campbell, M. and Landsverk, J. (2014). The cost of implementing new strategies (coins): A method for mapping implementation resources using the stages of implementation completion. *Children and Youth Services Review,* 39**,** 177-182.

Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health,* 18**,** 179-183.

Sanders, G. D., Neumann, P. J., Basu, A. and et al. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA,* 316**,** 1093-1103.

Sanders, R., Han, A., Baker, J. and Cobley, S. (2015). Childhood obesity and its physical and psychological co-morbidities: A systematic review of australian children and adolescents. *European Journal of Pediatrics,* 174**,** 715-746.

Sankar, P., Cho, M. K., Condit, C. M. and et al. (2004). Genetic research and health disparities. *JAMA,* 291**,** 2985-2989.

Sardinha, L. B., Santos, D. A., Silva, A. M., Grøntved, A., Andersen, L. B. and Ekelund, U. (2016). A comparison between BMI, waist circumference, and waist-to-height ratio for identifying cardio-metabolic risk in children and adolescents. *PloS one,* 11**,** e0149351-e0149351.

Scarborough, P., Bhatnagar, P., Wickramasinghe, K., Allender, S., Foster, C. and Rayner, M. (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: An update to 2006–07 NHS costs. *Journal of Public Health,* 33**,** 527-535.

Schonert-Reichl, K., Guhn, M., Gadermann, A., Hymel, S., Sweiss, L. and Hertzman, C. (2012). Development and validation of the Middle Years Development Instrument (mdi): Assessing children's well-being and assets across multiple contexts. *Social Indicators Research,* 114**,** 345-369.

Scottish Government. (2019) *A consultation on the new national public health body 'public health scotland'* Available at: https://www.gov.scot/publications/consultation-new-national-public-health-body-public-health-scotland/pages/2/ (Accessed: 29 August 2019).

Sculpher, M., Claxton, K., Drummond, M. and McCabe, C. (2006). Whither trial-based economic evaluation for health care decision making? *Health Economics,* 15**,** 677-687.

Segal, L., Mortimer, D. and Dalziel, K. (2005) *Risk factor study–how to reduce the burden of harm from poor nutrition, tobacco smoking, physical inactivity and alcohol misuse: Cost utility analysis of 9 multi-risk factor interventions.* Available at: https://doi.org/10.4225/03/5934ce3fc69ed (Accessed: 21 January 2019).

Severens, J. L. and Milne, R. J. (2004). Discounting health outcomes in economic evaluation: The ongoing debate. *Value in Health,* 7**,** 397-401.

Shiell, A., Donaldson, C., Mitton, C. and Currie, G. (2002). Health economic evaluation. *Journal of Epidemiology and Community Health,* 56**,** 85-88.

Sibbald, B. and Roland, M. (1998). Understanding controlled trials: Why are randomised controlled trials important? *BMJ,* 316**,** 201.

Simoens, S. (2009). Health economic assessment: A methodological primer. *International Journal of Environmental Research and Public Health,* 6**,** 2950-2966.

Smith, G., Ebrahim, S., Lewis, S., Hansell, A., Palmer, L. and Burton, P. (2005). Genetic epidemiology and public health: Hope, hype, and future prospects. *The Lancet,* 366**,** 1484-1498.

Solans, M., Pane, S., Estrada, M., Serra-Sutton, V., Berra, S., Herdman, M., Alonso, J. and Rajmil, L. (2008). Health-related quality of life measurement in children and adolescents: A systematic review of generic and disease-specific instruments. *Value in Health,* 11**,** 742-764.

Song, M. and Giovannucci, E. (2016). Estimating the influence of obesity on cancer risk: Stratification by smoking is critical. *Journal of Clinical Oncology,* 34**,** 3237-3239.

Squires, H., Chilcott, J., Akehurst, R., Burr, J. and Kelly, M. (2016). A framework for developing the structure of public health economic models. *Value in Health,* 19**,** 588-601.

Stata statistical software: Release 13. StataCorp LP

Stevens, K. (2009). Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Quality of Life Research,* 18**,** 1105-1113.

Stevens, K. (2012). Valuation of the child health utility 9d index. *PharmacoEconomics,* 30**,** 729-47.

Story, M., Nanney, M. and Schwartz, M. (2009). Schools and obesity prevention: Creating school environments and policies to promote healthy eating and physical activity. *The Milbank Quarterly,* 87**,** 71-100.

Strain, T., Milton, K., Dall, P., Standage, M. and Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. *British Journal of Sports Medicine,* 22.

Streiner, D., Norman, G. and Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use,* Oxford, UK, Oxford University Press.

Sturm, R. and Gresenz, C. (2002). Relations of income inequality and family income to chronic medical conditions and mental health disorders: National survey. *BMJ,* 324**,** 20.

Svensson, M. and Hultkrantz, L. (2017). A comparison of cost-benefit and cost-effectiveness analysis in practice: Divergent policy practices in Sweden. *Nordic Journal of Health Economics,* 5**,** 41-53.

Tavakol, M. and Dennick, R. (2011). Making sense of cronbach's alpha. *International Journal of Medical Education,* 2**,** 53-55.

te Velde, S., Lennert Veerman, J., Tak, N., Bosmans, J., Klepp, K. and Brug, J. (2011). Modeling the long term health outcomes and cost-effectiveness of two interventions promoting fruit and vegetable intake among schoolchildren. *Economics and Human Biology,* 9**,** 14-22.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J. and Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (wemwbs): Development and UK validation. *Health and Quality of Life Outcomes,* 5**,** 63.

The Daily Mile Foundation. (2018a). *The Daily Mile* [Online]. Available: https://thedailymile.co.uk/ [Accessed 10 December 2018].

The Daily Mile Foundation. (2018b). *The Daily Mile - about* [Online]. Available: https://thedailymile.co.uk/about/ [Accessed 9 May].

The Daily Mile Foundation. (2018c). *The Daily Mile - starter pack* [Online]. Available: https://thedailymile.co.uk/wp-content/uploads/2018/04/180404-Daily-MIle-Brochure-Starter-Pack-v2-online.pdf [Accessed 9 May].

The World Bank. (2019). *High income economies* [Online]. Available: https://data.worldbank.org/income-level/high-income [Accessed 25 January 2019].

Thorrington, D. and Eames, K. (2015). Measuring health utilities in children and adolescents: A systematic review of the literature. *PLOS One,* 10**,** e0135672.

Tones, K. and Green, J. (2010). *Health promotion: Planning and strategies,* London, UK, Sage.

Torgerson, D. and Raftery, J. (1999a). Discounting. *BMJ,* 319**,** 914-915.

Torgerson, D. and Raftery, J. (1999b). Measuring outcomes in economic evaluations. *BMJ,* 318**,** 1413-1413.

Torrance, G. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics,* 5**,** 1-30.

Torrance, G. (1987). Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases,* 40**,** 593-600.

Torrance, G. (1997). Preferences for health outcomes and cost-utility analysis. *American Journal of Managed Care,* 3**,** S8-S20.

Tourangeau, R., Rips, L. and Rasinski, K. (2000). *The psychology of survey response,* Cambridge, UK, Cambridge University Press.

Townsend, P. (1979). *Poverty in the united kingdom: A survey of household resources and standards of living,* UK, Penguin.

Tran, B., Ohinmaa, A., Kuhle, S., Johnson, J. and Veugelers, P. (2014). Life course impact of school-based promotion of healthy eating and active living to prevent childhood obesity. *PLOS One,* 9**,** e102242.

Tsiachristas, A., Waters, B. H., Adams, S. A., Bal, R. and M'lken, M. P. R.-v. (2014). Identifying and explaining the variability in development and implementation costs of disease management programs in the netherlands. *BMC Health Services Research,* 14**,** 1-14.

Tulchinsky, T. and Varavikova, E. 2014. Chapter 2 - expanding the concept of public health. *In:* Tulchinsky, T. & Varavikova, E. (eds.) *The new public health* 3rd ed. San Diego: Academic Press.

UK Cabinet Office. (2012) *A guide to social return on investment.* Available at: http://www.socialvalueuk.org/resources/sroi-guide/ (Accessed: 15 May 2017).

Ungar, W. and Gerber, A. 2009. The uniqueness of child health and challenges to measuring costs and consequences. *In:* Ungar, W. (ed.) *Economic evaluation in child health.* Oxford: Oxford University Press.

Ungar, W. and Santos, M. (2003). The pediatric quality appraisal questionnaire: An instrument for evaluation of the pediatric health economics literature. *Value in Health,* 6**,** 584-594.

University of Birmingham. (2019). *Food provision, culture and environment in secondary schools (fuel) study* [Online]. Available: http://www.birmingham.ac.uk/FUEL-study [Accessed 9th February 2019].

University of British Columbia. *Middle years development instrument* [Online]. Available: http://earlylearning.ubc.ca/mdi/ [Accessed 20 February 2019].

Upton, P., Eiser, C., Cheung, I., Hutchings, H., Jenney, M., Maddocks, A., Russell, I. and Williams, J. (2005). Measurement properties of the UK-english version of the pediatric quality of life inventory™ 4.0 (PEDSQL™) generic core scales. *Health and Quality of Life Outcomes,* 3**,** 22.

US Department of Health and Human Services. (2014) *The health consequences of smoking - 50 years of progress: A report of the surgeon general.* Available at: https://www.ncbi.nlm.nih.gov/books/NBK179276/ (Accessed: 3 April 2017).

Ussher, M. H., Owen, C. G., Cook, D. G. and Whincup, P. H. (2007). The relationship between physical activity, sedentary behaviour and psychological wellbeing among adolescents. *Social Psychiatry and Psychiatric Epidemiology,* 42**,** 851-856.

Vale, L., Adamson, A. and Critchley, J. (2012) *Economic evaluation: Evaluating the short-term impacts of the school food policy and experimental modelling of longer term impacts.* Available at: http://phrc.lshtm.ac.uk/papers/PHRC_B5-07A_Final_Report.pdf (Accessed: 21 January 2019).

Vallejo-Torres, L. and Morris, S. (2010). The contribution of smoking and obesity to income-related inequalities in health in England. *Social Science & Medicine,* 71**,** 1189-1198.

Varni, J., Limbers, C. and Burwinkle, T. (2007). How young can children reliably and validly self-report their health-related quality of life?: An analysis of 8,591 children across age subgroups with the pedsql™ 4.0 generic core scales. *Health and Quality of Life Outcomes,* 5**,** 1.

Verbooy, K., Hoefman, R., van Exel, J. and Brouwer, W. (2018). Time is money: Investigating the value of leisure time and unpaid work. *Value Health,* 21**,** 1428-1436.

Versteegh, M., Knies, S. and Brouwer, W. (2016). From good to better: New dutch guidelines for economic evaluations in healthcare. *PharmacoEconomics,* 34**,** 1071-1074.

Wakschlag, L., Pickett, K., Cook, E., Benowitz, N. and Leventhal, B. (2002). Maternal smoking during pregnancy and severe antisocial behavior in offspring: A review. *American Journal of Public Health,* 92**,** 966-974.

Walker, D. G., Hutubessy, R. and Beutels, P. (2010). Who guide for standardisation of economic evaluations of immunization programmes. *Vaccine,* 28.

Wang, H., Li, T., Siahpush, M., Chen, L. and Huberty, J. (2017). Cost-effectiveness of ready for recess to promote physical activity in children. *Journal of School Health,* 87**,** 278-285.

Wang, L. and Michael, S. (2015). Long-term health and medical cost impact of smoking prevention in adolescence. *Journal of Adolescent Health,* 56**,** 160-166.

Wang, L., Yang, Q., Lowry, R. and Wechsler, H. (2003). Economic analysis of a school-based obesity prevention program. *Obesity Research,* 11**,** 1313-24.

Wanless, D. (2004). *Securing good health for the whole population,* London, HM Treasury.

Ware, J. and Sherbourne, C. (1992). The mos 36-item short-form health survey (sf-36) conceptual framework and item selection. *Medical Care,* 30**,** 473-83.

Waters, E., Gibbs, L., Tadic, M., Ukoumunne, O., Magarey, A., Okely, A., de Silva, A., Armit, C., Green, J., O'Connor, T., Johnson, B., Swinburn, B., Carpenter, L., Moore, G., Littlecott, H. and Gold, L. (2017). Cluster randomised trial of a school-community child health promotion and obesity prevention intervention: Findings from the evaluation of fun ‘n healthy in moreland! *BMC Public Health,* 18**,** 92.

Watson, A., Timperio, A., Brown, H., Best, K. and Hesketh, K. (2017). Effect of classroom-based physical activity interventions on academic and physical activity outcomes: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity,* 14**,** 114.

Weatherly, H., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., Rice, N., Sculpher, M. and Sowden, A. (2009). Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health Policy,* 93**,** 85-92.

Whitehead, M. (1991). The concepts and principles of equity and health. *Health Promotion International,* 6**,** 217-228.

Whitehead, M. and Dahlgren, G. (1991). What can be done about inequalities in health? *The Lancet,* 338**,** 1059-1063.

WHO. (1986) *Ottawa charter for health promotion.* Available at: http://www.euro.who.int/__data/assets/pdf_file/0004/129532/Ottawa_Charter.pdf?ua=1 (Accessed: 15 August 2016).

WHO (1994). Development of the whoqol: Rationale and current status. *International Journal of Mental Health,* 23**,** 24-56.

WHO. (2007) *Cost effectiveness and strategic planning (WHO-CHOICE).* Available at: https://www.who.int/choice/cost-effectiveness/inputs/en/ (Accessed: 1 March 2017).

WHO. (2012) *European action plan for strengthening public health capacities and services.* Available at: http://www.euro.who.int/__data/assets/pdf_file/0005/171770/RC62wd12rev1-Eng.pdf?ua=1 (Accessed: 8 August 2016).

Wilkinson, R. and Marmot, M. (eds.) 2003. *Social determinants of health: The solid facts*.

Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., Egmar, A., Greiner, W., Gusi, N., Herdman, M., Jelsma, J., Kind, P., Scalone, L. and Ravens-Sieberer, U. (2010). Development of the EQ-5D-Y: A child-friendly version of the EQ-5D. *Quality of Life Research,* 19**,** 875-886.

Williams, J., Wake, M., Hesketh, K., Maher, E. and Waters, E. (2005). Health-related quality of life of overweight and obese children. *JAMA,* 293**,** 70-76.

Wolfenden, L., Nathan, N. K., Sutherland, R., Yoong, S. L., Hodder, R. K., Wyse, R. J., Delaney, T., Grady, A., Fielding, A., Tzelepis, F., Clinton-McHarg, T., Parmenter, B., Butler, P., Wiggers, J., Bauman, A., Milat, A., Booth, D. and Williams, C. M. (2017). Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Cochrane Database of Systematic Reviews*.

Wolfram, N. and Fuchs, A. 2008. Prevention and health promotion. *In:* Kirch, W. (ed.) *Encyclopedia of public health.* Dordrecht: Springer Netherlands.

Wolstenholme, J., Bargo, D., Wang, K., Harnden, A., Räisänen, U. and Abel, L. (2018). Preference-based measures to obtain health state utility values for use in economic evaluations with child-based populations: A review and UK-based focus group assessment of patient and parent choices. *Quality of Life Research,* 27**,** 1769-1780.

Wu, S., Cohen, D., Shi, Y., Pearson, M. and Sturm, R. (2011). Economic analysis of physical activity interventions. *American Journal of Preventive Medicine,* 40**,** 149-58.

Wyatt, K., Lloyd, J., Creanor, S., Green, C., Dean, S., Hillsdon, M., Abraham, C., Tomlinson, R., Pearson, V., Taylor, R., Ryan, E., Streeter, A., McHugh, C., Hurst, A., Price, L., Crathorne, L., Krägeloh, C., Siegert, R. and Logan, S. (2018). Cluster randomised controlled trial and economic and process evaluation to determine the effectiveness and cost-effectiveness of a novel intervention [healthy lifestyles programme (HELP)] to prevent obesity in school children. *Public Health Research,* 6.

Wyrwich, K., Norquist, J., Lenderking, W. and Acaster, S. (2013). Methods for interpreting change over time in patient-reported outcome measures. *Quality of Life Research,* 22**,** 475-483.

Yeo, S., Lloyd-Williams, H. and Tudor Edwards, R. 2019. The supply and demand of preventative goods and services and the need for economic analysis *In:* Edwards, R. T. & Mcintosh, E. (eds.) *Applied health economics for public health practice and research.* Oxford: Oxford University Press.

Young, N., Varni, J., Snider, L., McCormick, A., Sawatzky, B., Scott, M., King, G., Hetherington, R., Sears, E. and Nicholas, D. (2009). The internet is valid and reliable for child-report: An example using the activities scale for kids (ASK) and the pediatric quality of life inventory (PEDSQL). *Journal of Clinical Epidemiology,* 62**,** 314-320.

Zanganeh, M., Adab, P., Li, B. and Frew, E. (2019). A systematic review of methods, study quality, and results of economic evaluation for childhood and adolescent obesity intervention. *International Journal of Environmental Research and Public Health,* 16**,** 485.

APPENDIX A

Systematic Review Search Strategies

	Medline using Ovid
1	Economics/
2	exp "costs and cost analysis"/
3	Economics, Dental/
4	exp economics, hospital/
5	Economics, Medical/
6	Economics, Nursing/
7	Economics, Pharmaceutical/
8	(economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.
9	(expenditure$ not energy).ti,ab.
10	value for money.ti,ab.
11	budget$.ti,ab.
12	or/1-11
13	((energy or oxygen) adj cost).ti,ab.
14	(metabolic adj cost).ti,ab.
15	((energy or oxygen) adj expenditure).ti,ab.
16	or/13-15
17	12 not 16
18	exp exercise/
19	physical inactivity.mp.
20	physical activity.mp.
21	exp motor activity/
22	(physical education and training).mp.
23	exp "Physical Education and Training"/
24	exp physical fitness/
25	sedentary.ab. or sedentary.ti.
26	exp life style/
27	exp leisure activities/
28	exp walking/
29	exp sports/
30	exp dancing/
31	dancing.mp.
32	exp exercise therapy/
33	(exercise$ adj aerobic$).ti,ab.
34	(physical$ adj5 (fit$ or train$ or activ$ or endur$)).ti,ab.
35	(exercis$ adj5 (train$ or physical$ or activ$)).ti,ab.
36	sport$.ti,ab.
37	walk$.ti,ab.
38	cycle$.ti,ab.
39	(("lifestyle" or life-style) adj5 activ$).ti,ab.
40	(("lifestyle" or life-style) adj5 physical$).ti,ab.

41	or/18-40
42	exp obesity/
43	Weight Gain/
44	exp Weight Loss/
45	obes*.ti,ab.
46	(weight gain or weight loss).ti,ab.
47	(overweight or over weight or overeat* or over eat*).ti,ab.
48	weight change*.ti,ab.
49	((bmi or body mass index) adj2 (gain or loss or change)).ti,ab.
50	(obes* adj2 (prevent* or treat*)).ti,ab.
51	or/42-50
52	exp Diet/
53	nutrition*.ti,ab.
54	healthy eating.ti,ab.
55	Child Nutrition Sciences/
56	fruit*.ti,ab.
57	vegetable*.ti,ab.
58	canteen*.ti,ab.
59	food service*.ti,ab.
60	menu*.ti,ab.
61	calorie*.ti,ab.
62	Energy Intake/
63	energy density.ti,ab.
64	Eating/
65	Feeding Behavior/ or feeding behaviour.mp.
66	dietary intake.ti,ab.
67	Food Habits/
68	Food/
69	Carbonated Beverages/ or soft drink*.mp.
70	soda.ti,ab.
71	(sweetened drink* or sweetened beverage*).ti,ab.
72	Dietary Fats, Unsaturated/ or Dietary Fats/
73	confectionar*.ti,ab.
74	(school adj (lunch* or meal*)).ti,ab.
75	menu plan*.ti,ab.
76	((feeding or food or nutrition*) adj program*).ti,ab.
77	cafeteria*.ti,ab.
78	Nutritional Status/
79	or/52-78
80	schools/
81	school*.ti,ab.
82	kinder*.ti,ab.
83	80 or 81 or 82
84	letter.pt.
85	editorial.pt.
86	historical article.pt.
87	84 or 85 or 86
88	41 or 51 or 79
89	17 and 83 and 88

90	89 not 87
91	limit 90 to english language

	PsychInfo
1	"costs and cost analysis"/
2	"Cost Containment"/
3	(economic adj2 evaluation$).ti,ab.
4	(economic adj2 analy$).ti,ab.
5	(economic adj2 (study or studies)).ti,ab.
6	(cost adj2 evaluation$).ti,ab.
7	(cost adj2 analy$).ti,ab.
8	(cost adj2 (study or studies)).ti,ab.
9	(cost adj2 effective$).ti,ab.
10	(cost adj2 benefit$).ti,ab.
11	(cost adj2 utili$).ti,ab.
12	(cost adj2 minimi$).ti,ab.
13	(cost adj2 consequence$).ti,ab.
14	(cost adj2 comparison$).ti,ab.
15	(cost adj2 identificat$).ti,ab.
16	(pharmacoeconomic$ or pharmaco-economic$).ti,ab.
17	or/1-16
18	(task adj2 cost$).ti,ab,id.
19	(switch$ adj2 cost$).ti,ab,id.
20	(metabolic adj cost).ti,ab,id.
21	((energy or oxygen) adj cost).ti,ab,id.
22	((energy or oxygen) adj expenditure).ti,ab,id.
23	or/18-22
24	17 not 23
25	schools/
26	((primary or elementary or middle or junior) adj (school* or student*)).mp.
27	kinder*.mp.
28	25 or 26 or 27
29	exp Obesity/
30	Weight Gain/
31	exp Weight Loss/
32	obes*.ti,ab.
33	(weight gain or weight loss).ti,ab.
34	(overweight or over weight or overeat* or over eat*).ti,ab.
35	weight change*.ti,ab.
36	((bmi or body mass index) adj2 (gain or loss or change)).ti,ab.
37	primary prevention.ti,ab.
38	(preventive measure* or preventative measure*).ti,ab.
39	(preventive care or preventative care).ti,ab.
40	(obes* adj2 (prevent* or treat*)).ti,ab.
41	29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40
42	exp Exercise/
43	physical activity.ti,ab.
44	physical inactivity.ti,ab.
45	(physical education and training).ti,ab.
46	Physical Fitness/
47	sedentary.ti,ab.
48	exp Sports/

49	dancing.ti,ab.
50	(exercise* adj aerobic*).ti,ab.
51	sport*.ti,ab.
52	((lifestyle* or life style*) adj5 activ*).ti,ab.
53	42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52
54	nutrition*.ti,ab.
55	healthy eating.ti,ab.
56	(fruit* or vegetable* or canteen* or food service* or menu* or calorie*).ti,ab.
57	energy density.ti,ab.
58	Eating/
59	Feeding Behavior/ or feeding behaviour.mp.
60	dietary intake.ti,ab.
61	Food/
62	Carbonated Beverages/ or soft drink*.mp.
63	(soda* or sweetened drink* or sweetened beverage*).ti,ab.
64	unsaturated fat*.ti,ab.
65	confectionar*.ti,ab.
66	(school adj (lunch* or meal*)).ti,ab.
67	menu plan*.ti,ab.
68	((feeding or food or nutrition*) adj program*).ti,ab.
69	cafeteria*.ti,ab.
70	54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69
71	(animal or animals or rat or rats or mouse or mice or hamster or hamsters or dog or dogs or cat or cats or bovine or sheep or ovine or pig or pigs).ab,ti,id,de.
72	editorial.dt.
73	letter.dt.
74	dissertation abstract.pt.
75	72 or 73 or 74
76	41 or 53 or 70
77	24 and 28 and 76
78	77 not 75
79	limit 78 to english language

	Embase
1	Health Economics/
2	exp Economic Evaluation/
3	exp Health Care Cost/
4	pharmacoeconomics/
5	1 or 2 or 3 or 4
6	(econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.
7	(expenditure$ not energy).ti,ab.
8	(value adj2 money).ti,ab.
9	budget$.ti,ab.
10	6 or 7 or 8 or 9
11	5 or 10
12	(metabolic adj cost).ti,ab.
13	((energy or oxygen) adj cost).ti,ab.
14	((energy or oxygen) adj expenditure).ti,ab.
15	12 or 13 or 14
16	11 not 15
17	animal/
18	exp animal experiment/
19	nonhuman/
20	(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab,sh.
21	17 or 18 or 19 or 20
22	16 not 21
23	exp Obesity/
24	Weight Gain/
25	exp Weight Loss/
26	obes*.tw.
27	(weight gain or weight loss).tw.
28	(overweight or over weight or overeat* or over eat*).tw.
29	weight change*.tw.
30	((bmi or body mass index) adj2 (gain or loss or change)).tw.
31	exp Primary Prevention/
32	(primary prevention or secondary prevention).tw.
33	(preventive measure* or preventative measure*).tw.
34	(preventive care or preventative care).tw.
35	(obes* adj2 (prevent* or treat*)).tw.
36	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
37	exp Exercise/
38	physical activity.tw.
39	physical inactivity.tw.
40	exp Motor Activity/
41	(physical education and training).tw.
42	exp "Physical Education and Training"/
43	Physical Fitness/
44	sedentary.tw.
45	exp Life Style/
46	exp Leisure Activities/

47	exp Sports/
48	Dancing/
49	dancing.tw.
50	(exercise* adj aerobic*).tw.
51	sport*.tw.
52	((lifestyle* or life style*) adj5 activ*).tw.
53	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
54	exp Diet/
55	nutrition*.tw.
56	healthy eating.tw.
57	Child Nutrition Sciences/
58	fruit*.tw.
59	vegetable*.tw.
60	canteen*.tw.
61	food service*.tw.
62	menu*.tw.
63	calorie*.tw.
64	Energy Intake/
65	energy density.tw.
66	Eating/
67	Feeding Behavior/ or feeding behaviour.mp.
68	dietary intake.tw.
69	Food Habits/
70	Carbonated Beverages/ or soft drink*.mp.
71	soda.tw.
72	sweetened drink*.tw.
73	Dietary Fats, Unsaturated/ or Dietary Fats/
74	confectionar*.tw.
75	(school adj (lunch* or meal*)).tw.
76	menu plan*.tw.
77	((feeding or food or nutrition*) adj program*).tw.
78	cafeteria*.tw.
79	Nutritional Status/
80	54 or 55 or 56 or 57 or 58 or 59 or 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
81	schools/
82	((primary or elementary or middle or junior) adj (school* or student*)).mp.
83	81 or 82
84	letter.pt.
85	editorial.pt.
86	note.pt.
87	conference abstract.pt.
88	85 or 86 or 87
89	36 or 53 or 80
90	22 and 83 and 89
91	90 not 88
92	limit 91 to english language

	NHS EED
1	(School* or ((primary or elementary or middle or junior) and student*) or kinder*)
2	(Obes* or "Weight Gain" or "Weight Loss" or overweight or "over weight" or overeat* or "over eat*" or "weight change*" or ((bmi or body mass index) and (gain or loss or change)) or "Primary Prevention" or "secondary prevention" or "preventive measure*" or "preventative measure" or "preventive care" or "preventative care" or Exercise or "physical activity" or "physical inactivity" or "Motor Activity" or "physical education and training" or "Physical Fitness" or sedentary or "Life Style" or lifestyle or "Leisure Activities" or sport* or Dancing or aerobic* or diet or nutrition* or "healthy eating" or "Child Nutrition Sciences" or fruit* or vegetable* or canteen* or "food service*" or menu* or calorie* or "Energy Intake" or "energy density" or eating or "Feeding Behavior" or "Feeding Behaviour" or "dietary intake" or food or "Carbonated Beverage*" or "soft drink*" or soda or "sweetened drink*" or “sweetened beverage*” or "Dietary Fats" or confectionar* or "school lunch*" or "school meal*" or "menu plan*" or ((feeding or food or nutrition*) and program*) or cafeteria*)
3	#1 and #2 in Economic Evaluations

	ERIC
S5	S4 Limiters - Journal or Document: Journal Articles (EJ); Educational Level: Early Childhood Education, Elementary Education, Elementary Secondary Education, Grade 1, Grade 2, Grade 3, Grade 4, Grade 5, Grade 6, Grade 7, Grade 8, Grade 9, Grade 10, Grade 11, Grade 12, Intermediate Grades, Junior High Schools, Kindergarten, Middle Schools, Primary Education
S4	S1 AND S2 AND S3
S3	TI ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$)) OR AB ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$))
S2	TI ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)) OR AB ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)
S1	TI ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*)) OR AB ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*))

	SPORTDiscus
S6	S1 AND S2 AND S3 Limiters - Publication Type: Academic Journal Narrow by Language: - english
S4	S1 AND S2 AND S3
S3	TI ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)) OR AB ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)
S2	TI ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*)) OR AB ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*))
S1	TI ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$)) OR AB ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$))

	BEI
S4	S1 AND S2 AND S3
S3	TI ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)) OR AB ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)
S2	TI ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*)) OR AB ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*))
S1	TI ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$)) OR AB ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$))

	EconLit
S4	S1 AND S2 AND S3
S3	TI ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)) OR AB ((Obes* OR "Weight Gain" OR "Weight Loss" OR overweight OR "over weight" OR overeat* OR "over eat*" OR "weight change*" OR ((bmi OR body mass index) AND (gain OR loss OR change)) OR "Primary Prevention" OR "secondary prevention" OR "preventive measure*" OR "preventative measure" OR "preventive care" OR "preventative care" OR Exercise OR "physical activity" OR "physical inactivity" OR "Motor Activity" OR "physical education and training" OR "Physical Fitness" OR sedentary OR "Life Style" OR lifestyle OR "Leisure Activities" OR sport* OR Dancing OR aerobic* OR diet OR nutrition* OR "healthy eating" OR "Child Nutrition Sciences" OR fruit* OR vegetable* OR canteen* OR "food service*" OR menu* OR calorie* OR "Energy Intake" OR "energy density" OR eating OR "Feeding Behavior" OR "Feeding Behaviour" OR "dietary intake" OR food OR "Carbonated Beverage*" OR "soft drink*" OR soda OR "sweetened drink*" OR "Dietary Fats" OR confectionar* OR "school lunch*" OR "school meal*" OR "menu plan*" OR ((feeding OR food OR nutrition*) AND program*) OR cafeteria*)
S2	TI ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*)) OR AB ((School* OR ((primary OR elementary OR middle OR junior) AND student*) OR kinder*))
S1	TI ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$)) OR AB ((economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$))

	CINAHL
S1	MH "Economics+"
S2	MH "Financial Management+"
S3	MH "Financial Support+"
S4	MH "Financing, Organized+"
S5	MH "Business+"
S6	S2 OR S3 or S4 OR S5
S7	S1 NOT S6
S8	MH "Health Resource Allocation"
S9	MH "Health Resource Utilization"
S10	S8 OR S9
S11	S7 OR S10
S12	TI (cost or costs or economic* or pharmacoeconomic* or price* or pricing*) OR AB (cost or costs or economic* or pharmacoeconomic* or price* or pricing*)
S13	S11 OR S12
S14	PT editorial
S15	PT letter
S16	PT commentary
S17	S14 or S15 or S16
S18	S13 NOT S17
S19	MH "Animal Studies"
S20	(ZT "doctoral dissertation") or (ZT "masters thesis")
S21	S18 NOT (S19 OR S20)
S22	(MH "Nutritional Status")
S23	"cafeteria*"
S24	((feeding or food or nutrition*) n1 program*)
S25	"menu plan*"
S26	(school n1 (lunch* or meal*))
S27	(MH "Candy")
S28	"confectionar*"
S29	(MH "Dietary Fats")
S30	"sweetened drink*" Or "sweetened beverage*"
S31	"soda"
S32	(MH "Carbonated Beverages") OR "soft drink*"
S33	(MH "Food")
S34	(MH "Food Habits")
S35	(MH "Food Intake") OR "dietary intake"
S36	(MH "Eating Behavior") OR "feeding behaviour"
S37	(MH "Eating")
S38	(MH "Energy Density")

S39	(MH "Energy Intake")
S40	"calorie*"
S41	"menu*"
S42	(MH "Food Services") OR "food service*"
S43	AB canteen*
S44	fruit*
S45	(MH "Vegetables") OR "vegetable*"
S46	(MH "Fruit+")
S47	"Child Nutrition Sciences" OR (MH "Child Nutrition")
S48	healthy eating
S49	nutrition*
S50	(MH "Diet+")
S51	S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50
S52	((lifestyle* or life style*) n5 activ*)
S53	AB sport*
S54	(exercise* n1 aerobic*)
S55	(MH "Dancing+") OR "Dancing"
S56	(MH "Sports+")
S57	(MH "Leisure Activities+")
S58	(MH "Life Style+") OR (MH "Life Style, Sedentary")
S59	"sedentary"
S60	(MH "Physical Fitness")
S61	(MH "Physical Education and Training") OR "physical education and training"
S62	(MH "Motor Activity+")
S63	physical inactivity
S64	(MH "Physical Activity")
S65	(MH "Exercise+")
S66	S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62 OR S63 OR S64 OR S65
S67	(obes* n2 (prevent* or treat*))
S68	preventive care or preventative care
S69	preventative measure*
S70	preventive measure*
S71	"secondary prevention"
S72	"Primary Prevention"
S73	((bmi or body mass index) n2 (gain or loss or change))
S74	weight change*
S75	(overweight or over weight or overeat* or over eat*)
S76	(weight gain or weight loss)
S77	obes*

S78	(MH "Weight Loss")
S79	(MH "Weight Gain")
S80	(MH "Obesity+")
S81	S67 OR S68 OR S69 OR S70 OR S71 OR S72 OR S73 OR S74 OR S75 OR S76 OR S77 OR S78 OR S79 OR S80

APPENDIX B

Table 28 Characteristics of Included Studies

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
Adab et al. (2018a)	Evaluate the cost-effectiveness of an obesity prevention intervention programme in primary school aged children	UK	WAVES Cooking workshops Signposting of PA opportunities Increased PA Use of role models	No intervention	Full	CUA	Cost per QALY gained	Trial Cluster RCT	Not cost-effective
Babey et al. (2014)	To compare different school based opportunities to increase physical activity	USA	In class physical activity breaks Two daily 1 minute structured PA breaks using exercise videos	N/A	Full	CEA	MET hours gained	Model	Cost-effective
Barrett et al. (2015)	Estimate the cost-effectiveness of an active PE	USA	50% of PE devoted to MVPA, during existing PE lessons	Current practice Estimated 40% MVPA in PE	Full	CEA	Cost per MET hour increase per day	Model	Small benefit

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	policy on PA and BMI			for 97 minutes per week			after 1 year		
Brown et al. (2007)	Assess the cost-effectiveness and net benefit of the CATCH programme	USA	Classroom curriculum PE programme Modification to school food service Family/home based IV	No intervention	Full	CUA	Cost per QALY gained	Model	Cost-effective
Cecchini et al. (2010)	Assess different public health strategies to tackle behavioural risk factors for chronic diseases	England, India, China, Mexico, Brazil, South Africa, Russia	1 hr per week of health education (diet and PA), guest lecture, support of school nurses, indirect education (leaflets etc), environmental changes (e.g. food contracts).	No prevention	Full	CUA	DALY	Model	Not cost-effective
Cradock et al. (2014)	Test the effectiveness of an active	USA	Commit to providing 150mins MVPA per week during existing	No intervention	Partial	Trial reporting costs	Change in MVPA,	Trial Non-	Beneficial

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	school day policy on PA and estimate implementation costs		PE Provide equipment, training to PE educators Wellness champions to increase PA in recess, curriculum and active classroom breaks					randomised	
Cradock et al. (2017)	To evaluate the cost-effectiveness of different policy and programmes to prevent childhood obesity	USA	Active PE - State level policy where 50% of PE must be of moderate to vigorous physical activity	Healthy afterschool - State policy for afterschool learning	Full	CEA	Cost per MET hour per day change	Model	Cost-effective
Cradock et al. (2017)	To evaluate the cost-effectiveness of different policy and programmes to	USA	Active Recess - District level voluntary programme where PA in recess is increased using	Healthy afterschool - State policy for afterschool learning	Full	CEA	Cost per MET hour per day change	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	prevent childhood obesity		structured activities, playground markings and portable equipment						
Cradock et al. (2017)	To evaluate the cost-effectiveness of different policy and programmes to prevent childhood obesity	USA	Active School Day - District level policy requiring schools to provide opportunities for at least 150 minutes of MVPA per day using strategies such as active PE and Recess, classroom breaks and others	Healthy afterschool - State policy for afterschool learning	Full	CEA	Cost per MET hour per day change	Model	Cost-effective
Dollahite and Hosig (1998)	Understand the impact of a school based nutrition programme	USA	Nutrition curriculum Grocery tour New lunch menus Nutrition messages Parent attendance at lunch days	Usual activities	Partial	Trial reporting costs	Success in changing nutrition knowledge and food choice	Trial Cluster non-randomised	Beneficial

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
							behaviour		
Eckermann et al. (2014)	To evaluate the ROI of a school based health promotion programme	Australia	Stephanie Alexander Kitchen Garden National Programme Grant to build kitchen and garden facilities Garden and kitchen specialists support weekly lessons in a kitchen and garden (45mins vegetable garden and 90mins in kitchen) Link lessons to the curriculum	No intervention	Full	ROI	ROI	Trial - Case control	Beneficial
Ekwaru et al. (2017)	To evaluate the cost-effectiveness of APPLE schools in Canada using a lifetime	Canada	Alberta Project Promoting active Living and healthy eating in schools (APPLE) Full time school health facilitator for 2	Usual activities	Full	CUA	Cost per QALY gained	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusio n
	modelling approach		years to implement healthy eating and active living policies whilst engaging students, teachers, parents. School health curriculum with cross curricular links Staff professional development						
Flores (1995)	To determine whether Dance for Health has an effect on cardiovascular health and attitudes towards PA	USA	Aerobic dance 3 times a week for 50 minutes, during usual PE class Health education class twice a week. 25 30 mins long lessons covering nutrition, exercise, smoking etc	No intervention	Partial	Trial reporting costs	Timed mile run BMI Resting heart rate Attitudes to PA	Trial - Cluster RCT	Beneficial
Gortmaker et al. (2015)	To assess the cost-effectiveness of	USA	Federal nutrition standards for school meals for all grade	$0.01/oz SSB tax administered	Full	CEA	Cost per unit of	Model	Cost-saving

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	7 childhood obesity prevention strategies using microsimulation models		levels, including a requirement to increase availability of fruit and vegetables, whole grains, reduce sodium, fats and set minimum and maximum calorie levels	nationally Eliminating the tax deductibility of television (TV) advertising costs for nutritionally poor foods and beverages			BMI reduced		
Gortmaker et al. (2015)	To assess the cost-effectiveness of 7 childhood obesity prevention strategies using microsimulation models	USA	Nutrition Standards for All Foods and Beverages Sold in Schools (Smart Snacks) Intervention. They focus on providing whole grains, fruits and vegetables, key nutrients, while limiting calories, sodium, fats and sugar	advertised to children and adolescents ages 2-19 Restaurant Menu Calorie Labelling The Nutrition and Physical Activity Self-Assessment for Child Care (NAP SACC) Program	Full	CEA	Cost per unit of BMI reduced	Model	Cost-saving

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
				Nationwide four-fold increase in the use of bariatric surgery by eligible adolescents, ages 13 to 19 years old					
Graziose et al. (2017)	Estimate the cost-effectiveness of a nutrition education curriculum intervention in preventing obesity	USA	24 lesson health nutrition education programme delivered during lesson time by teachers over 1 year. Focuses on reducing SSB, fast food, processed food, screen time, PA and increasing fruit and veg	No intervention	Full	CUA	Cost per QALY gained	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
Hendy et al. (2011)	Examine the effectiveness of the Kid's Choice Programme for improving weight management behaviours	USA	Kid's Choice Programme 3 month intervention Children wear name tags at lunch and recess that have holes punched in the when they demonstrate a weight management behaviour (fruit eaten, healthy drink and exercise) Children receive rewards each week, where they trade stars for prizes	Control. Good citizen rewards	Partial	Trial reporting costs	BMI change	Trial - Non-randomised 1 school	Beneficial
Kesztyus et al. (2013)	An economic evaluation of an overweight prevention programme, with metabolism,	Germany	28 teaching units delivered during lessons (targeting PA, sweetened beverages, media use) Activity breaks	No intervention	Full	CEA	Cost per cm (waist circumference) Cost per unit (waist to	Trial Cluster RCT	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	exercise and lifestyle components		Homework Materials for parents				height ratio)		
Kesztyus et al. (2017)	Assess the cost-effectiveness of a school based, state wide health promotion programme	Germany	Teaching delivered during lessons (targeting PA, sweetened beverages, media use) Activity breaks Homework Materials for parents	No intervention	Full	CEA	Cost per case of abdominal obesity avoided	Trial Cluster RCT	Cost-effective
Ladapo et al. (2016)	To examine the costs and CE of implementing a multi-component obesity prevention intervention	USA	Peer leaders to educate other students Peer leaders to promote healthy behaviours Environmental changes e.g. water fountains Signs and posters	No intervention	Full	CEA	Cost per student per day	Trial Cluster RCT	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
Lawlor et al. (2016)	To investigate the effectiveness of a school-based intervention to increase physical activity, reduce sedentary behaviour and increase fruit and vegetable consumption, with a CCA also	UK	Active for Life Year 5 16 lessons, delivered by teachers about PA and diet. Teachers attended a 1 day training event 10 parent-child homework activities Information for schools to put in newsletters Written information from Change4Life	Usual activities	Full	CCA	Time spent in MVPA Time spent in sedentary behaviour Servings of fruit and vegetables	Trial Cluster RCT	Not effective
Li et al. (2017)	Estimate the cost-effectiveness of a salt reduction programme in China	China	Eight lesson salt reduction education programme Homework to educate parents and family	No intervention	Full	CEA/CUA	Systolic blood pressure unit reduction	Model/trial	Cost-effective
Manger et al. (2012)	Assess the effectiveness of an obesity	USA	8 weekly lessons lasting 30mins teaching about	Normal curriculum	Partial	Trial reporting costs	BMI change	Trial - Non randomi	Small benefit

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	prevention programme (VITAL)		healthy eating and PA					sed pre-post design	
McAuley et al. (2010)	Assess the costs and health benefits of a community based obesity prevention initiative	New Zealand	A Pilot Programme for Lifestyle and Exercise Community activity coordinators at each school encouraged a little more PA every day by increasing variety and opportunities Short bursts of activity during class time Water filters Free fruit Resources targeting sugary drinks, fruit and veg	No intervention	Full	CEA	Kilogram of weight gain prevented per child per year	Trial Cluster non-randomised	Beneficial
Meng et al. (2013)	Evaluate the cost-effectiveness of	China	1 year intervention Nutrition education - PA intervention -	No intervention	Full	CEA	BMI change	Trial Cluster RCT	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	a combined PA and nutrition intervention compared to a single intervention to prevent childhood obesity		Happy 10 moderate PA for 10mins, either once, twice or for 20mins in the classroom. Education about PA was given also Comprehensive intervention - Combination of PA and nutrition education						
Mernagh et al. (2010)	Assess the cost-effectiveness of public health interventions to prevent obesity	New Zealand	A Pilot Programme for Lifestyle and Exercise (APPLE) 2 year intervention Community activity coordinators at each school encouraged a little more PA every day by increasing variety and opportunities Short bursts of	No intervention	Full	CUA	Cost per QALY gained	Model	Not cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
			activity during class time Water filters Free fruit Resources targeting sugary drinks, fruit and veg						
Mernagh et al. (2010)	Assess the cost-effectiveness of public health interventions to prevent obesity	New Zealand	Be Active Eat Well 3 year intervention Targeted evidence based behaviour change. Reduction in TV viewing, reduced SSB consumption, reduced energy dense snacks, increased water, fruit and veg consumption.	No intervention	Full	CUA	Cost per QALY gained	Model	Not cost-effective
Mernagh et al. (2010)	Assess the cost-effectiveness of public health	New Zealand	School nutrition policy initiative 2 year intervention School nutrition	No intervention	Full	CUA	Cost per QALY gained	Model	Not cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusio n
	interventions to prevent obesity		education, policy (e.g. reduced price healthy food), social marketing, parent outreach.						
Mernagh et al. (2010)	Assess the cost-effectiveness of public health interventions to prevent obesity	New Zealand	Switch-play 1 year intervention Activity/behaviour modification. (1) through reducing the time spent in sedentary behaviours (e.g. TV viewing, playing electronic games and recreational computer use); (2) through increasing skills and enjoyment of physical activity; or (3) through a combination of these two strategies	No intervention	Full	CUA	Cost per QALY gained	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
Moodie	Examine the cost-effectiveness of the Be Active Eat Well programme in Australia	Australia	Be Active Eat Well Targeted evidence based behaviour change. Reduction in TV viewing, reduced SSB consumption, reduced energy dense snacks, increased water, fruit and veg consumption.	Usual activities	Full	CUA	DALYs averted	Model	Cost-effective
Moodie et al. (2013)	Assess the cost-effectiveness of a school programme to increase active transport as an obesity prevention measure	Australia	TravelSMART (TSS) Programme designed to increase active travel to school behaviours It included meetings and information sessions, classroom activities, professional development for teachers, whole	Usual activities	Full	CUA	DALYs averted	Model	Not cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
			school activities Classroom activities were for 20 hours over 4 weeks						
Mora et al. (2015)	Assess whether short term effects of a health education intervention persist for up to 6 years	Spain	3 hours a week of activities related to healthy eating habits/physical activity, forming part of normal timetabled lessons	No intervention	Full	CEA	BMI change	Trial Cluster RCT	Significant effect at small cost
Ohinmaa et al. (2011)	A cost-analysis of a comprehensive school health programme in Canada	Canada	Annapolis Valley Health Promoting Schools Programme Making healthy choice the easy choice Health policies and practices Creative supporting environment Skill development More physical	N/A	Partial	Cost-analysis	N/A	Cross sectional review of costs	Costs comparable to similar

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
			activity Healthier food choices More educated about health and nutrition						
Reznik et al. (2015)	Evaluate the short term impact of CHAM JAM on PA levels during the school day	USA	CHAM JAM Audio CD of 10 minute education focused aerobic activities. 25 different lessons, with teachers choosing which one to do and doing them 3 times a day	Usual activities	Partial	Trial reporting costs	Number of steps per school day	Trial Cluster RCT	Beneficial
Rush et al. (2014)	Estimate lifetime cost-effectiveness of a multicomponent PA and nutrition programme to	New Zealand	Multicomponent intervention Increases physical activity Encourages healthy eating Delivered by 'Energizers',	Historical comparison of no intervention	Full	CUA	Cost per QALY gained	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	improve health and reduce rate of weight gain								
Segal et al. (2005)	To estimate the cost-effectiveness of school based interventions seeking to modify harmful lifestyle behaviours	Australia	18 lessons of 30-50 min on TV viewing and video game use. 10 day television turn-off followed by 7 hour per week limit Television managers budgeted viewing by controlling the power Newsletters providing advice to parents about reducing viewing	Usual activities	Full	CEA/CUA	Cost per BMI unit reduction	Model/trial	Not cost-effective
Segal et al. (2005)	To estimate the cost-effectiveness of school based interventions seeking to modify harmful	Australia	Focused sessions infused into curriculum lessons and PE, covering PA, high fat foods, fruit and veg, television viewing	Usual activities	Full	CEA/CUA	Cost per hour of TV viewing reduction	Model/trial	Not cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	lifestyle behaviours								
Segal et al. (2005)	To estimate the cost-effectiveness of school based interventions seeking to modify harmful lifestyle behaviours	Australia	Classroom lessons twice a week. Topics included: 'heart healthy foods', the importance of physical exercise, the dangers of smoking, and how to resist pressure to smoke. Physical activity lessons three times a week. Fun aerobic lessons, with warm up and cool down.	Usual activities	Full	CEA/CUA	Cost per BMI unit reduction	Model/trial	Not cost-effective
Tran et al. (2014)	Estimate the life course impact of the APPLE Schools program in terms of future	Canada	Alberta Project Promoting active Living and healthy eating in schools (APPLE) Full time school health facilitator for 2	No intervention	Full	CCA	Costs and overweight prevented	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	body weight status and avoided health care costs		years to implement healthy eating and active living policies whilst engaging students, teachers, parents. School health curriculum with cross curricular links Staff professional development						
Vale et al. (2012)	To estimate the short term impacts of a school food policy (CCA) and exploratory long-term modelling	UK	School food policy that states what cannot be served and limits the amount of times that certain foods can be provided	No change in policy	Full	CCA/CUA	Cost per QALY gained	Model	Judgement required
te Velde et al. (2011)	Assess the cost-effectiveness of 2 school based diet	The Netherlands	PRO children Provision of a free piece of fruit, a carrot/tomato twice a week, a classroom	No intervention	Full	CUA	DALYs averted	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
	interventions, compared to no intervention		curriculum Schoolgruiten Free fruit and veg scheme, encouraged to use a curriculum to increase knowledge and skills relevant to fruit and veg.						
Wang et al. (2003)	Assess the cost-effectiveness and cost-benefit of Planet Health	USA	Focused sessions infused into curriculum lessons and PE, covering PA, high fat foods, fruit and veg, television viewing	No intervention	Full	CUA	Cost per QALY gained	Model	Cost-effective
Wang et al. (2017)	Examine the cost-effectiveness of the Ready for Recess PA programme in the USA	USA	Ready for Recess Activity zones and equipment Activities facilitated using activity cards One component from Active and	No intervention	Full	CEA	MET hours gained	Trial	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
			Healthy School programme						
Waters et al. (2017)	Evaluate the effectiveness of a multicomponent intervention for childhood obesity focusing on inequalities	Australia	Fun and Healthy in Moreland Based on health promoting schools Community Development Workers worked with schools to develop programmes customised for them	Usual activities	Partial	Trial reporting costs	BMI zscore	Trial Cluster RCT	Not effective
Wu et al. (2011)	To calculate cost-effectiveness ratios for physical activity interventions	Not specific but costs in dollars	School based PA interventions. Authors systematically reviewed effectiveness studies, finding 16 publications (26 interventions). From these they estimated costs and effects in MET hours gained.	Point of decision prompts Community campaign Individual adapted behaviour change Social support Creation or enhanced	Full	CEA	Cost per MET-hour gained, per person	Model	Cost-effective

First author and year	Objective	Country	Intervention	Comparator	Full or partial analysis	Type of economic evaluation	Outcome	Study design	Economic conclusion
			They then averaged CERs across 26 interventions	access to environments for PA					
Wyatt et al. (2018)	To estimate the costs and resource use associated with the HeLP intervention and its cost-effectiveness	UK	24 components, mostly delivered by co-ordinators. Components included assemblies, newsletters, poetry, PHSE and drama lessons, goal setting and parents' forum	Usual activities	Partial	No significant effect, so cost analysis	Cost per QALY gained	Trial Cluster RCT	Not effective

APPENDIX C

Middle Years Development Instrument (MDI)

We would like to ask you some questions about how you think and feel about things in your life and about what you like to do.

Here are some things to know before getting started:

This is **not a test!** There are **no right or wrong answers**. Some people think or feel one thing and other people think or feel something else. We want to know what *you* think and how *you* feel. Your answers are VERY IMPORTANT and will help improve activities and programs for children your age.

It is important for you to know that ALL OF YOUR ANSWERS that you put in this survey will be **confidential (private)** and will **not** be shared with your teacher, Headteacher, parents, or your friends.

Please answer each question the best you can. **Thank you for your help!**

INSTRUCTIONS

•If you do not understand a question, please raise your hand and **ask for help**.

•Make sure you **understand** the question and how to mark your answer **before** you answer.

•Only mark **one answer** for each question.

Please tell us a little bit about yourself

What is your first name?

Are you a boy or a girl? □ Boy □ Girl

What year are you in at school? □ Year 3□ Year 4□ Year 5 □ Year 6

What is your birthday? _____________ ________ ___________ (drop down boxes)

Month Day

Which of these adults do you live with most of the time? *(Tick all adults you live with.)*

□ Mother □ Father □ Part time with each parent

□ Grandmother □ Grandfather □ Foster parent(s) or carers

□ Stepmother □ Second mother (two mothers)

□ Stepfather □ Second father (two fathers)

□ Other adults (write in the space below, for example, aunt, uncle, mum's boyfriend or girlfriend, dad’s boyfriend or girlfriend):

____

How many brothers and sisters do you have?

□ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 or more

These questions are about how much you agree or disagree with the statement.

Remember there are no right or wrong answers. Please tick one answer for each question.

	Disagree a lot	Disagree a little	Don't agree or disagree	Agree a little	Agree a lot
I feel sorry for other children who don't have the things that I have.					
When I see someone being mean to someone else it bothers me					
I am a person who cares about the feelings of others.					
I have more good times than bad times.					
I believe more good things than bad things will happen to me.					
I start most days thinking I will have a good day.					
In general, I like being the way I am.					
Overall, I have a lot to be proud of.					
A lot of things about me are good.					
I feel unhappy a lot of the time.					
I feel upset about things.					
I feel that I do things wrong a lot.					
I worry about what other children might be saying about me.					
I worry a lot that other people might not like me.					
I worry about being teased.					
In most ways my life is close to the way I would want it to be.					
The things in my life are excellent.					
I am happy with my life.					
So far I have got the important things I want in life.					
If I could live my life over again, I would have it the same way.					

Since the start of this school year, how often did you do this?

	Not at all in this school year	Once or a few times	About every month	About every week	Many times a week
I cheered someone up who was feeling sad					
I helped someone who was being picked on					
I helped someone who was hurt					

Are there any adults who are IMPORTANT TO YOU at your school

□ No □ Yes

If you answered YES to the question above we would like to know the names for ALL the adults who are important to you in school.

Person 1

Person 2

Person 3

Person 4

Person 5

Person 6

How true is each statement for you?

At my **school** there is a teacher or another adult

	Not at all true	A little true	Pretty much true	Very much true
Who really cares about me				
Who believes that I will be a success				
Who listens to me when I have something to say				

The next four questions are about your parents (or guardians) or other adults who live in your home.

In my **home** there is a parent or another adult

	Not at all true	A little true	Pretty much true	Very much true
Who believes that I will be a success				
Who listens to me when I have something to say				
Who I can talk to about my problems				
I care about what my parents (or guardians) think of me				

In my **area/community** (Not from your school or family) there is an adult

	Not at all true	A little true	Pretty much true	Very much true
Who really cares about me				
Who believes that I will be a success				
Who listens to me when I have something to say				

	No	Yes	Don't know
Are there places in your area/community that provide programs for children your age, like sports, art, dance, music classes and other clubs and activities?			
Are there safe places in your area/community where you feel comfortable to be with your friends, like playgrounds, parks or community centres?			

Please answer the following questions about you and **your friend(s) and your school**

	Disagree a lot	Disagree a little	Don't agree or disagree	Agree a little	Agree a lot
I feel part of a group of friends that do things together.					
I feel that I usually fit in with other children around me.					
When I am with other children my age I feel I belong					
I have at least one really good friend I can talk to when something is bothering me.					
I have a friend I can tell everything to.					
There is somebody my age who really understands me.					
I am certain I can learn the skills taught in school.					
If I have enough time, I can do a good job on all my school work.					
Even if the work in school is hard, I can learn it.					
Teachers and pupils treat each other with respect in this school.					
People care about each other in this school.					
Students in this school help each other, even if they are not friends.					
I feel like I belong in this school.					
I feel like I am important to this school.					
When I grow up, I have goals and plans for the future.					

How important is it to **you** to do the following in school

	Not important at all	Not very important	Somewhat important	Very important
Make friends				
Get good grades				
Learn new things				

Important definition: **Bully -** There are a lot of different ways to bully someone, but a bully has some advantage (stronger, more popular, or something else), wants to hurt the other person (it's not an accident), and does so repeatedly (over and over again) and unfairly. Sometimes a group of pupils will bully another pupil.

The next four questions might make you feel uncomfortable, but it is important for us to know. Please answer the questions honestly.

This school year, how often have you been bullied by other pupils in the following ways?

	Not at all in this school year	Once or a few times	About every month	About every week	Many times a week
Physical bullying. For example, someone hit, shoved, kicked you, spat at you, damaged or took your things without permission					
Verbal bullying (for example, someone called you names, teased, embarrassed, threatened you, or made you do things you didn't want to do).					
Social bullying (for example, someone left you out, excluded you, gossiped and spread rumours about you, or made you look foolish).					
Cyberbullying (for example, someone used the computer or text messages to exclude, threaten, embarrass you, or to hurt your feelings).					

The next questions are about your physical health.

In general, how would you describe your health?

□ Poor
□ Fair
□ Good
□ Excellent

Do you have a physical or health condition that keeps you from doing some things other children your age do? (For example, school activities, sports, or getting together with friends).
□ No
□ Yes, a physical disability (for example, deafness, wheelchair, or something else)
□ Yes, a long term illness (for example, diabetes, or something else)
□ Yes, overweight
□ Yes, something else (please specify)
____

How do you rate your body weight?

□ Very underweight
□ Slightly underweight
□ About the right weight
□ Slightly overweight
□ Very overweight

How often do you like the way you look?

□ Never
□ Hardly ever
□ Sometimes
□ Often
□ Always

In a usual week can you say how often you do these things?

	Never	Once a week	2 times a week	3 times a week	4 times a week	5 times a week	6 times a week	Every day
How often do you eat breakfast?								
How often do your parents or other adult family members eat meals with you?								
How often do you eat food like sweets, crisps, drink pop or something else?								
How often do you get a good night's sleep?								

What time do you usually go to bed during the week

□ Before 9pm
□ Between 9pm and 10pm
□ Between 10pm and 11pm
□ Between 11pm and midnight
□ After midnight

ABOUT YOUR TIME AFTER SCHOOL

On school days, who are you usually with for most of the time from after school time to dinner time (about 3pm – 6pm)

Please tick all of the people you are with after school

□ By myself □ Father (or stepfather, foster father)
□ Friend(s) about my age □ Younger brothers/sisters
□ Mother (or stepmother, foster mother) □ Older brothers/sisters
□ Other adult(s) (for example, grandparent, aunt or uncle, coach, babysitter)
□ Other (describe) ______________________

How many days a week do you go to these places from after school to dinner time (about 3pm to 6pm)

	Never	Once a week	2 times a week	3 times a week	4 times a week	5 times a week (every school day)
I go home						
I stay at school to participate in after school activities (for example, sports, tutoring, clubs).						
I go to an after school program/childcare (in my school or somewhere else).						
I go to a friend's house.						
I go to a park, playground or community centre						
I go to the shops						
I go somewhere else, for example, a family member's home, or other places.						

The next questions are about activities that are organised. So activities that are planned and supervised by a teacher, instructor, adult, coach or volunteer.

During last week from after school to dinner time (about 3pm to 6pm) how many days did you participate in:

	Never	Once a week	2 times a week	3 times a week	4 times a week	5 times a week (every school day)
Educational lessons or activities (for example, tutoring, maths, language school, or something else)?						
Art or music lessons (for example, drawing, painting, playing a musical instrument, or something else)?						
Youth clubs (for example, Scouts, Girl Guides, Boys and Girls Clubs, or something else)?						
Individual sports with a coach or instructor (for example, swimming, dance, gymnastics, tennis, skating, or something else)?						
Team sports with a coach or instructor (for example, basketball, hockey, football, or something else)?						

The next questions ask you about other activities that you might do after school. That is, these questions are about activities that are not planned and usually **<u>not supervised</u>** by a teacher, instructor, adult, coach, or volunteer.

During last week from after school to dinner time (about 3pm to 6pm) how much time did you spend on the following activities on a normal day?:

	I did not do this activity	Less than 30 minutes	30 mins to 1 hour	1 – 2 hours	2 hours or more
sports and/or exercise for fun (for example, football, swimming, yoga, dancing, or something else)?					
do homework?					
watch TV, Netflix, Youtube, streaming videos, or something else?					
play video or computer games (for example, Play Station, Wii, Xbox, multi-user online games, or something else)?					
read for fun?					
practice a musical instrument (for example, drums, clarinet, violin, or something else)?					
do arts & crafts (for example, painting, drawing, or something else)?					
Spend time with friends in person?					
Spend time with friends on the phone, tablet or computer?					

Think about what you WANT to do on school days from after school to dinner time (about 3pm – 6pm)

□ I am already doing the activities I want to be doing.
□ I wish I could do additional activities.
Please list **one** activity you wish you could do:
____
Where would you like this activity to be?
□ School
□ Home
□ Park or playground
□ Community centre
□ Other (describe)___________________

What stops you from participating in the activities that you want to participate in after school? *(Tick all of the things that stop you.)*
□ Nothing stops me.
□ I have to go straight home after school.
□ It is too difficult to get there.
□ The activity that I want is not offered.
□ The activity times do not fit the times that I can attend.
□ It's not safe for me to go.
□ I have too much homework to do.
□ My parents do not approve.
□ It costs too much.
□ I need to take care of brothers or sisters or do other things at home.
□ I am afraid I will not be good enough in that activity.
□ I'm too busy.
□ I don't know what is available.
□ None of my friends are interested or want to go.
□ Other, please describe ___________________________

	Disagree a lot	Disagree a little	Don't agree or disagree	Agree a little	Agree a lot
When I'm sad, I can usually start doing something that will make me feel better.					
After I'm interrupted or distracted, I can easily continue working where I left off.					
I can calm myself down when I'm excited or upset.					
If something isn't going according to my plans, I change my actions to try and reach my goal.					
When I have a serious disagreement with someone, I can talk calmly about it without losing control.					
I work carefully when I know something will be tricky.					

APPENDIX D

CHU9D

These questions ask about how you are **today.** For each question, read all the choices and decide which one is most like you **today**.

1. Worried

I don't feel worried today
I feel a little bit worried today
I feel a bit worried today
I feel quite worried today
I feel very worried today

2. Sad

I don't feel sad today
I feel a little bit sad today
I feel a bit sad today
I feel quite sad today
I feel very sad today

3. Pain

I don't have any pain today
I have a little bit of pain today
I have a bit of pain today
I have quite a lot of pain today
I have a lot of pain today

4. Tired

I don't feel tired today
I feel a little bit tired today
I feel a bit tired today
I feel quite tired today
I feel very tired today

5. Annoyed

I don't feel annoyed today
I feel a little bit annoyed today
I feel a bit annoyed today
I feel quite annoyed today
I feel very annoyed today

6. School Work/Homework (such as reading, writing, doing lessons)

I have no problems with my schoolwork/homework today
I have a few problems with my schoolwork/homework today
I have some problems with my schoolwork/homework today
I have many problems with my schoolwork/homework today
I can't do my schoolwork/homework today

7. Sleep

Last night I had no problems sleeping
Last night I had a few problems sleeping
Last night I had some problems sleeping
Last night I had many problems sleeping
Last night I couldn't sleep at all

8. Daily routine (things like eating, having a bath/shower, getting dressed)

I have no problems with my daily routine today
I have a few problems with my daily routine today
I have some problems with my daily routine today
I have many problems with my daily routine today
I can't do my daily routine today

9. Able to join in activities (things like playing out with your friends, doing sports, joining in things)

I can join in with any activities today
I can join in with most activities today
I can join in with some activities today
I can join in with a few activities today
I can join in with no activities today

APPENDIX E

Ethical approval

Dear Dr Emma Frew & Dr Iestyn Williams

Re: "Public Health Interventions in Schools. A Qualitative Study"
Application for Ethical Review ERN_17-0171

Thank you for your application for ethical review for the above project, which was reviewed by the Science, Technology, Engineering and Mathematics Ethical Review Committee.

On behalf of the Committee, I confirm that this study now has full ethical approval.

I would like to remind you that any substantive changes to the nature of the study as described in the Application for Ethical Review, and/or any adverse events occurring during the study should be promptly bought to the Committee's attention by the Principal Investigator and may necessitate further ethical review.

Please also ensure that the relevant requirements within the University's Code of Practice for Research and the information and guidance provided on the University's ethics webpages (available at https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Links-and-Resources.aspx) are adhered to and referred to in any future applications for ethical review. It is now a requirement on the revised application form (https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Ethical-Review-Forms.aspx) to confirm that this guidance has been consulted and is understood, and that it has been taken into account when completing your application for ethical review.

Please be aware that whilst Health and Safety (H&S) issues may be considered during the ethical review process, you are still required to follow the University's guidance on H&S and to ensure that H&S risk assessments have been carried out as appropriate. For further information about this, please contact your School H&S representative or the University's H&S Unit at healthandsafety@contacts.bham.ac.uk.

Kind regards,

Miss Sam Waldron
Deputy Research Ethics Officer
Research Support Group

Aston Webb Building
University of Birmingham
Edgbaston B15 2TT
Tel:
Email:

APPENDIX F

Information sheet

Public Health Interventions in Schools. A Qualitative Study

PARTICIPANT INFORMATION SHEET

You are being invited to take part in an interview study related to the Daily Mile. The project is exploring experiences of the Daily Mile, the outcomes you think are important to measure in school based initiatives like this and the potential costs of the intervention.

Before you decide whether to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Ask me if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

When initiatives such as the Daily Mile are introduced, they could displace assessed or non-assessed curriculum work. Alternatively they may be incorporated into lessons and contribute to teaching of a subject e.g. a maths lesson. Despite delivering them, schools may not necessarily observe any benefits from such interventions. I am interested in what you think should be the goals of these interventions from a school's perspective. I would also like to understand how the Daily Mile is being conducted in your school and any initial and ongoing costs incurred. This will help evaluate the cost-effectiveness of the Daily Mile.

Why is the study being done?

I am conducting a study examining methods used to conduct cost-effectiveness studies of interventions delivered in schools. Public health initiatives that aim to improve health and wellbeing can be delivered and funded by different providers. These providers can have different aims, for example maximising individual's health, reducing inequalities or achieving maximum learning potential. This interview study will help me determine what outcomes are important to schools participating in the Daily Mile, the appropriateness of the different outcomes we are collecting and how much the Daily Mile costs.

Why have I been chosen?

You have been approached because your school is participating in the Daily Mile.

Do I have to take part?

Participation in the study is entirely voluntary and you are under no obligation take part.

What will happen to me if I take part?

At a time that is convenient I am hoping to have a one-to-one discussion with you so that I can ask you some questions. I am asking your permission to have that discussion. If you agree, I will write down what is being said and with your permission, I will audio record the interview to ensure a more accurate account. I would also like to ask you some questions so I can describe the sample that participated (e.g. years of teaching experience). I will collect this information using a short questionnaire. All questions will be optional.

You do not have to agree to participate. If you want the interview to be stopped, this is not a problem.

Will the information I collect be kept confidential?

All information collected about you from conducting the interviews will be kept strictly confidential and anonymous. Data, transcripts and recordings will be kept in locked filing cabinets and password protected computer storage spaces. Anonymous transcripts will be kept as secure files until 10 years after the study ends. All analysis of data will be anonymous, transcripts and recordings will be indexed solely by the use of study ID numbers so that participants are not identifiable. Anonymised data from the interviews may also be used in conjunction with research data from other studies for academic purposes. While written extracts (verbatim quotations) may be used within publications relating to this study, no participants will be referred to by name in any subsequent publications. All data will be treated in accordance with the Data Protection Act 1998

What if I change my mind after the interview has been conducted?

If you would like to withdraw from the study you can do this at any time up to two weeks after the interview by sending me an email ( ), or telephoning ( ). If you notify us after this time period it is possible the data may already have been analysed and used to inform future interviews. If you withdraw from the research I will ask you if you want your contribution so far to be included. If you would like me to not include your contribution I will destroy your written responses and I will not use quotes from your interview when we write-up the research and when I publish the findings. There will be no consequences for you if you withdraw.

Who has reviewed the study?

This study has been reviewed and been given favourable opinion by the University of Birmingham ethics committee.

What will happen to the results of this study?

The results will contribute to scientific publications, my PhD thesis and publications related to the Daily Mile. These include a cost-effectiveness evaluation of the study and a report of the interview findings. The results will be available in approximately 1 year. No names or other identifying information will be published in any reports.

Who can I talk to for more information or advice about the study?

My name is Katie Breheny, a Research Associate in Health Economics at the University of Birmingham. If you have any queries about this research please do not hesitate to contact me at: Department of Health Economics, Public Health Building, University of Birmingham, Birmingham, B15 2TT. Tel: [redacted]. Email: [redacted].

The supervisor of this project is Dr Emma Frew, a Reader in Health Economics at the University of Birmingham. She can be contacted at: [redacted]

If you wish to make a complaint, any complaint about the way you have been dealt with during the study or any possible harm you might have suffered will be addressed. To make a complaint you should contact me on the above number.

What do I do now?

If you would like to hear more about the study or think you might like to take part then please contact me at [redacted] or call me on [redacted]. Thank you for your time.

APPENDIX G

Participant Consent Form

Public Health Interventions in Schools. A Qualitative Study

Interview Study

PARTICIPANT CONSENT FORM

Please read the following statements and, if you agree, tick the corresponding box to confirm agreement:

	Initials
I confirm that I have read and understood the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
I understand that my participation is voluntary and that I am free to withdraw from the study at any time up to two weeks after this interview. without giving a reason.	
I understand that my data will be treated confidentially and any publication resulting from this work will be anonymised.	
I agree for this interview to be audio recorded	
I freely agree to participate in this study.	

Signatures:

______________________________ ______________________________

Name of participant (block capitals) **Signature** **Date**

______________________________ ______________________________

Researcher (block capitals) **Signature** **Date**

If you would like a copy of this consent form to keep, please ask me, Katie Breheny: [redacted]

If you have any complaints or concerns about this research, you can direct these to me at: Department of Health Economics, Public Health Building, University of Birmingham, Birmingham, B15 2TT. Tel: [redacted] **Email:** [redacted]

APPENDIX H

Participant characteristics form

1. What is your job role (e.g. teacher, head teacher, deputy head, teaching assistant, governor)?

   ____

2. For how many years have you been employed in education?

   _______ (years or months)

3. For how many years have you been employed at or worked with this school?

   _______ (years or months)

4. What year group(s) do you work with currently? (N/A if you do not teach)

   ____

5. Do you have any additional qualifications, training or responsibilities related to physical education? (please state)

   ____

APPENDIX I

Interview Topic Guide

	Main Question	Probes
Warm up	1. Role	- Teaching responsibilities - How long they've worked there - Specialisms - Previous roles
Daily Mile implementation	2. Reasons for participating	- Health and wellbeing strategy - Physical activity targets - OFSTED - Possible impact on attention, behaviour, academic attainment - National Child Measurement Programme results
	3. Experience of the Daily Mile	- Positive and negative characteristics - Challenges - Would they continue
	4. Duration and frequency	- How often they did it - How long did it take - Who supervised it
	5. Impacts on the school day	- Time for curriculum subjects, inclusion in lessons such as science, maths - Time for extra-curricular subjects - Time for PE
Cost	6. Costs of the Daily Mile	- Cost of equipment - Staff time - Costs to pupils and parents - Ongoing costs post-implementation
Priority of health and wellbeing in school	7. Role of the school in improving health and wellbeing	- Responsibility/priority for schools and teachers - Wider impacts of poor wellbeing - Outcomes important for school in health and wellbeing initiatives - For policy, what should the aim be? - How cost-effectiveness analysis could be used by school

	Main Question	Probes
		- How could cost-effectiveness analysis be presented in order to be meaningful
Daily Mile outcomes	8. Influence of the Daily Mile on health and wellbeing	- Impact of Daily Mile on effectiveness measures (wellbeing, academic attainment, fitness) - Appropriateness of Daily Mile effectiveness measures (BMI, subjective academic attainment, MDI, fitness) - Other outcomes not captured in the study
	9. Personal views on physical activity	- Importance - Personal involvement
	10. Any other comments?	

APPENDIX J

Table 29 Analysis of known groups: Gender

Scale	Gender		Mean difference (95% Confidence Intervals)	P-value
	Females	Males		
	Mean (SD)	Mean (SD)		
Social and emotional development				
Empathy	4.37 (0.83)	4.06 (1.01)	-0.31 (-0.40 : -0.22)	0.00
Prosocial behaviour	3.46 (1.18)	3.20 (1.22)	-0.25 (-0.37 : -0.13)	0.00
Optimism	4.08 (0.98)	3.90 (1.11)	-0.18 (-0.29 : -0.08)	0.00
Self-esteem	4.37 (0.85)	4.18 (1.01)	-0.20 (-0.29 : -0.11)	0.00
SWLS-C	4.05 (0.95)	3.98 (1.05)	-0.07 (-0.17 : 0.03)	0.14
Sadness	3.28 (1.11)	3.41 (1.13)	0.13 (0.02 : 0.24)	0.02
Worries/anxiety	3.31 (1.34)	2.94 (1.34)	-0.37 (-0.50 : -0.24)	0.00
Connectedness				
Adults at school	3.37 (0.72)	3.22 (0.79)	-0.15 (-0.23 : -0.08)	0.00
Adults in neighbourhood	3.08 (0.98)	2.99 (1.03)	-0.09 (-0.19 : 0.01)	0.07
Adults at home	3.52 (0.68)	3.39 (0.80)	-0.13 (-0.21 : -0.06)	0.00
Peer belonging	4.14 (0.98)	4.09 (1.08)	-0.06 (-0.16 : 0.04)	0.24
Friendship intimacy	4.38 (0.98)	4.28 (1.03)	-0.18 (-0.28 : -0.08)	0.00
School experiences				
Academic self-concept	4.37 (0.86)	4.27 (0.97)	-0.10 (-0.19 : -0.01)	0.03
School support	4.22 (0.94)	4.13 (1.01)	-0.11 (-0.20 : -0.01)	0.02
Bullying	1.82 (1.08)	2.02 (1.19)	0.20 (0.09 : 0.31)	0.00
School belonging	4.18 (1.09)	4.03 (1.21)	-0.16 (-0.28 : -0.05)	0.01

Wellbeing sum	58.75 (10.30)	57.67 (11.49)	-1.14 (-2.22 : -0.06)	0.04

Table 30 Analysis of Known Groups: Ethnicity

Scale	Ethnicity		Mean difference (95% Confidence Intervals)	P-value
	White	Non-white		
	Mean (SD)	Mean (SD)		
Social and emotional development				
Empathy	4.23 (0.92)	4.17 (0.97)	-0.04 (-0.13 : 0.06)	0.44
Prosocial behaviour	3.32 (1.18)	3.33 (1.23)	0.01 (-0.11 : 0.14)	0.85
Optimism	3.98 (1.03)	3.98 (1.08)	0.01 (-0.10 : 0.13)	0.79
Self-esteem	4.37 (0.85)	4.18 (1.01)	0.05 (-0.05 : 0.14)	0.35
SWLS-C	4.02 (1.00)	4.01 (1.01)	0.01 (-0.09 : 0.12)	0.83
Sadness	3.37 (1.11)	3.33 (1.14)	-0.01 (-0.12 : 0.11)	0.89
Worries/anxiety	3.15 (1.33)	3.07 (1.38)	-0.11 (-0.24 : 0.03)	0.13
Connectedness				
Adults at school	3.29 (0.76)	3.29 (0.76)	0.01 (-0.07 : 0.08)	0.90
Adults in neighbourhood	3.02 (1.02)	3.04 (0.98)	-0.01 (-0.12 : 0.10)	0.85
Adults at home	3.46 (0.75)	3.44 (0.74)	0.00 (-0.08 : 0.08)	1.00
Peer belonging	4.14 (1.02)	4.08 (1.05)	-0.03 (-0.14 : 0.08)	0.58
Friendship intimacy	4.38 (0.98)	4.28 (1.03)	-0.07 (-0.18 : 0.03)	0.18
School experiences				
Academic self-concept	4.30 (0.94)	4.34 (0.90)	0.06 (-0.03 : 0.16)	0.19
School support	4.22 (0.94)	4.13 (1.01)	-0.06 (-0.17 : 0.04)	0.22
Bullying	1.82 (1.08)	2.02 (1.19)	-0.09 (-0.21 : 0.03)	0.14
School belonging	4.10 (1.14)	4.07 (1.16)	0.03 (-0.09 : 0.15)	0.67
Wellbeing sum	58.22 (10.96)	58.13 (10.98)	0.28 (-0.88 : 1.43)	0.64

Table 31 Analysis of Known Groups: Year group

Scale	Year group		Mean difference (95% Confidence Intervals)	P-value
	Year 3	Year 5		
	Mean (SD)	Mean (SD)		
Social and emotional development				
Empathy	4.15 (1.04)	4.25 (0.85)	0.09 (0.00 : 0.18)	0.06
Prosocial behaviour	3.32 (1.24)	3.32 (1.17)	-0.01 (-0.13 : 0.11)	0.91
Optimism	3.96 (1.12)	4.00 (0.99)	0.03 (-0.07 : 0.14)	0.53
Self-esteem	4.20 (1.01)	4.33 (0.88)	0.12 (0.03 : 0.22)	0.01
SWLS-C	3.94 (1.07)	4.07 (0.94)	0.12 (0.02 : 0.22)	0.02
Sadness	3.30 (1.16)	3.39 (1.09)	0.09 (-0.02 : 0.21)	0.10
Worries/anxiety	3.26 (1.35)	2.98 (1.34)	-0.29 (-0.42 : -0.16)	0.00
Connectedness				
Adults at school	3.30 (0.76)	3.28 (0.76)	-0.03 (-0.10 : 0.05)	0.52
Adults in neighbourhood	3.14 (0.95)	2.94 (1.04)	-0.20 (-0.30 : -0.10)	0.00
Adults at home	3.37 (0.79)	3.52 (0.70)	0.15 (0.08 : 0.22)	0.00
Peer belonging	4.10 (1.03)	4.13 (1.03)	0.03 (-0.07 : 0.14)	0.53
Friendship intimacy	4.28 (1.03)	4.38 (0.98)	0.11 (0.01 : 0.21)	0.04
School experiences				
Academic self-concept	4.24 (0.99)	4.39 (0.85)	0.15 (0.06 : 0.24)	0.00
School support	4.26 (0.96)	4.10 (0.98)	-0.15 (-0.25 : -0.06)	0.00
Bullying	2.05 (1.24)	1.81 (1.05)	-0.25 (-0.36 : -0.14)	0.00
School belonging	4.12 (1.16)	4.07 (1.16)	-0.04 (-0.15 : 0.07)	0.50

Wellbeing sum	57.41 (11.22)	58.85 (10.68)	1.46 (0.37 : 2.54)	0.01

Table 32 Analysis of Known Groups: Academic attainment

Subject	Scale	Academic attainment Below expected level Mean (SD)	At expected or above Mean (SD)	Mean difference (95% Confidence Intervals)	P-value
Social and emotional development					
Maths	Empathy	4.03 (1.10)	4.32 (0.84)	0.26 (0.15 : 0.38)	0.00
Writing	Empathy	4.06 (1.08)	4.33 (0.82)	0.22 (0.10 : 0.34)	0.00
Reading	Empathy	4.01 (1.11)	4.33 (0.83)	0.28 (0.16 : 0.39)	0.00
Maths	Prosocial behaviour	3.28 (1.26)	3.36 (1.21)	0.06 (-0.09 : 0.21)	0.41
Writing	Prosocial behaviour	3.32 (1.27)	3.33 (1.19)	-0.06 (-0.21 : 0.09)	0.44
Reading	Prosocial behaviour	3.26 (1.26)	3.38 (1.20)	0.07 (-0.09 : 0.22)	0.40
Maths	Optimism	3.90 (1.14)	4.03 (1.00)	0.09 (-0.04 : 0.23)	0.15
Writing	Optimism	3.90 (1.13)	4.04 (1.00)	0.08 (-0.06 : 0.21)	0.26
Reading	Optimism	3.87 (1.16)	4.05 (0.99)	0.14 (0.01 : 0.27)	0.04
Maths	Self-esteem	4.10 (1.09)	4.35 (0.86)	0.23 (0.11 : 0.35)	0.00
Writing	Self-esteem	4.13 (1.06)	4.36 (0.86)	0.19 (0.06 : 0.31)	0.00
Reading	Self-esteem	4.11 (1.07)	4.34 (0.89)	0.20 (0.08 : 0.32)	0.00
Maths	SWLS-C	3.92 (1.08)	4.04 (0.99)	0.11 (-0.02 : 0.24)	0.10
Writing	SWLS-C	3.94 (1.08)	4.03 (0.97)	0.06 (-0.07 : 0.19)	0.36
Reading	SWLS-C	3.89 (1.09)	4.06 (0.98)	0.18 (0.05 : 0.30)	0.01
Maths	Sadness	3.26 (1.14)	3.36 (1.12)	0.11 (-0.03 : 0.25)	0.12
Writing	Sadness	3.23 (1.14)	3.38 (1.12)	0.16 (0.02 : 0.30)	0.03
Reading	Sadness	3.02 (1.05)	3.68 (1.11)	0.18 (0.03 : 0.32)	0.01
Maths	Worries/anxiety	3.08 (1.36)	3.15 (1.34)	-0.01 (-0.18 : 0.15)	0.89

Subject	Scale	Academic attainment		Mean difference (95% Confidence Intervals)	P-value
		Below expected level Mean (SD)	At expected or above Mean (SD)		
Writing	Worries/anxiety	3.09 (1.34)	3.15 (1.35)	0.02 (-0.14 : 0.18)	0.79
Reading	Worries/anxiety	3.07 (1.35)	3.15 (1.34)	-0.21 (-0.40 : -0.02)	0.03
Connectedness					
Maths	Adults at school	3.20 (0.83)	3.32 (0.71)	0.11 (0.02 : 0.21)	0.02
Writing	Adults at school	3.21 (0.82)	3.32 (0.70)	0.08 (-0.01 : 0.18)	0.09
Reading	Adults at school	3.16 (0.84)	3.35 (0.69)	0.17 (0.08 : 0.26)	0.00
Maths	Adults in neighbourhood	2.98 (1.05)	3.04 (0.97)	0.03 (-0.10 : 0.15)	0.67
Writing	Adults in neighbourhood	3.00 (1.04)	3.03 (0.97)	0.00 (-0.13 : 0.13)	1.00
Reading	Adults in neighbourhood	2.97 (1.04)	3.05 (0.98)	0.05 (-0.07 : 0.18)	0.40
Maths	Adults at home	3.32 (0.84)	3.52 (0.67)	0.20 (0.11 : 0.30)	0.00
Writing	Adults at home	3.34 (0.84)	3.53 (0.65)	0.19 (0.10 : 0.29)	0.00
Reading	Adults at home	3.31 (0.84)	3.53 (0.67)	0.24 (0.14 : 0.33)	0.00
Maths	Peer belonging	3.97 (1.13)	4.17 (0.96)	0.17 (0.05 : 0.30)	0.01
Writing	Peer belonging	4.00 (1.10)	4.17 (0.98)	0.13 (0.00 : 0.25)	0.06
Reading	Peer belonging	3.95 (1.13)	4.19 (0.96)	0.22 (0.09 : 0.34)	0.00
Maths	Friendship intimacy	4.14 (1.16)	4.43 (0.90)	0.29 (0.16 : 0.42)	0.00
Writing	Friendship intimacy	4.19 (1.12)	4.42 (0.92)	0.21 (0.08 : 0.34)	0.00
Reading	Friendship intimacy	4.14 (1.15)	4.42 (0.92)	0.27 (0.14 : 0.40)	0.00

Subject	Scale	Academic attainment Below expected level Mean (SD)	 At expected or above Mean (SD)	Mean difference (95% Confidence Intervals)	P-value
School experiences					
Maths	Academic self-concept	4.11 (1.10)	4.46 (0.74)	0.33 (0.22 : 0.44)	0.00
Writing	Academic self-concept	4.15 (1.09)	4.47 (0.70)	0.31 (0.19 : 0.42)	0.00
Reading	Academic self-concept	4.08 (1.12)	4.47 (0.72)	0.38 (0.27 : 0.49)	0.00
Maths	School support	4.03 (1.07)	4.22 (0.92)	0.14 (0.02 : 0.26)	0.03
Writing	School support	4.09 (1.04)	4.18 (0.94)	0.02 (-0.10 : 0.14)	0.72
Reading	School support	4.02 (1.09)	4.22 (0.90)	0.15 (0.02 : 0.27)	0.02
Maths	Bullying	2.10 (1.22)	1.84 (1.10)	-0.22 (-0.36 : -0.08)	0.00
Writing	Bullying	2.08 (1.23)	1.82 (1.06)	-0.23 (-0.38 : -0.09)	0.00
Reading	Bullying	2.10 (1.22)	1.85 (1.11)	-0.23 (-0.37 : -0.08)	0.00
Maths	School belonging	3.95 (1.22)	4.12 (1.13)	0.13 (-0.02 : 0.27)	0.08
Writing	School belonging	3.99 (1.21)	4.11 (1.14)	0.07 (-0.08 : 0.21)	0.35
Reading	School belonging	3.92 (1.24)	4.14 (1.11)	0.17 (0.03 : 0.31)	0.02
Maths	Wellbeing sum	56.64 (11.43)	58.75 (10.73)	1.93 (0.56 : 3.29)	0.01
Writing	Wellbeing sum	56.86 (11.27)	58.81 (10.81)	1.63 (0.25 : 3.01)	0.02

Subject	Scale	Academic attainment		Mean difference (95% Confidence Intervals)	P-value
		Below expected level Mean (SD)	At expected or above Mean (SD)		
Reading	Wellbeing sum	56.35 (11.38)	58.97 (10.73)	2.55 (1.17 : 3.93)	0.00

Table 33 Analysis of Known Groups: Quality of Life

Scale	Quality of life		Mean difference (95% Confidence Intervals)	P-value
	HRQL poor	HRQL good		
	Mean (SD)	Mean (SD)		
Social and emotional development				
Empathy	4.12 (0.97)	4.28 (0.91)	0.16 (0.07 : 0.25)	0.00
Prosocial behaviour	3.27 (1.19)	3.37 (1.22)	0.10 (-0.02 : 0.22)	0.09
Optimism	3.69 (1.10)	4.27 (0.92)	0.58 (0.48 : 0.68)	0.00
Self-esteem	4.06 (0.97)	4.48 (0.87)	0.42 (0.32 : 0.51)	0.00
SWLS-C	3.78 (1.03)	4.24 (0.92)	0.46 (0.36 : 0.55)	0.00
Sadness	3.02 (1.05)	3.68 (1.11)	0.66 (0.55 : 0.76)	0.00
Worries/anxiety	3.37 (1.29)	2.86 (1.37)	-0.48 (-0.61 : -0.34)	0.00
Connectedness				
Adults at school	3.14 (0.81)	3.43 (0.68)	0.30 (0.22 : 0.37)	0.00
Adults in neighbourhood	2.88 (1.03)	3.17 (0.95)	0.30 (0.20 : 0.40)	0.00
Adults at home	3.31 (0.80)	3.59 (0.66)	0.27 (0.20 : 0.35)	0.00
Peer belonging	3.86 (1.11)	4.37 (0.88)	0.51 (0.41 : 0.61)	0.00
Friendship intimacy	4.20 (1.09)	4.47 (0.88)	0.27 (0.17 : 0.37)	0.00
School experiences				
Academic self-concept	4.12 (1.01)	4.52 (0.77)	0.39 (0.30 : 0.48)	0.00
School support	3.99 (1.06)	4.35 (0.85)	0.37 (0.27 : 0.46)	0.00
Bullying	2.07 (1.16)	1.78 (1.11)	-0.27 (-0.38 : -0.16)	0.00
School belonging	3.85 (1.25)	4.33 (1.00)	0.48 (0.37 : 0.60)	0.00
Wellbeing sum	54.34 (10.75)	61.96 (9.79)	7.52 (6.50 : 8.54)	0.00

Table 34 Analysis of Known Groups: Deprivation

Scale	Deprivation		Mean difference (95% Confidence Intervals)	P-value
	High deprivation	Low deprivation		
	Mean (SD)	Mean (SD)		
Social and emotional development				
Empathy	4.18 (0.96)	4.27 (0.90)	0.08 (-0.05 : 0.22)	0.24
Prosocial behaviour	3.32 (1.21)	3.33 (1.21)	0.10 (-0.02 : 0.22)	0.09
Optimism	3.95 (1.07)	4.08 (1.00)	0.11 (-0.05 : 0.27)	0.17
Self-esteem	4.24 (0.95)	4.37 (0.90)	0.13 (-0.01 : 0.27)	0.07
SWLS-C	3.99 (1.01)	4.09 (0.99)	0.08 (-0.08 : 0.24)	0.34
Sadness	3.33 (1.13)	3.41 (1.10)	0.06 (-0.08 : 0.21)	0.40
Worries/anxiety	3.15 (1.34)	2.98 (1.39)	-0.16 (-0.34 : 0.02)	0.09
Connectedness				
Adults at school	3.27 (0.77)	3.36 (0.73)	0.09 (-0.02 : 0.19)	0.11
Adults in neighbourhood	3.00 (1.02)	3.13 (0.94)	0.12 (-0.06 : 0.29)	0.19
Adults at home	3.44 (0.75)	3.47 (0.74)	0.03 (-0.09 : 0.14)	0.65
Peer belonging	4.08 (1.06)	4.23 (0.93)	0.13 (-0.03 : 0.28)	0.12
Friendship intimacy	4.31 (1.01)	4.41 (0.97)	0.08 (-0.08 : 0.23)	0.32
School experiences				
Academic self-concept	4.30 (0.95)	4.38 (0.84)	0.07 (-0.08 : 0.22)	0.34
School support	4.15 (0.98)	4.24 (0.97)	0.07 (-0.14 : 0.28)	0.53
Bullying	1.95 (1.15)	1.85 (1.14)	-0.08 (-0.31 : 0.15)	0.49
School belonging	4.06 (1.18)	4.22 (1.09)	0.17 (-0.03 : 0.37)	0.10
Wellbeing sum	4.18 (0.96)	4.27 (0.90)	0.08 (-0.05 : 0.22)	0.24

Table 35 Analysis of Known Groups: General health

Scale	General health		Mean difference (95% Confidence Intervals)	P-value
	Poor health Mean (SD)	Good health Mean (SD)		
Social and emotional development				
Empathy	3.89 (1.15)	4.26 (0.90)	0.47 (0.32 : 0.62)	0.00
Prosocial behaviour	3.12 (1.24)	3.36 (1.20)	0.22 (0.05 : 0.40)	0.01
Optimism	3.56 (1.16)	4.05 (1.02)	0.47 (0.32 : 0.62)	0.00
Self-esteem	3.76 (1.12)	4.36 (0.88)	0.56 (0.43 : 0.69)	0.00
SWLS-C	3.53 (1.19)	4.09 (0.95)	0.53 (0.39 : 0.67)	0.00
Sadness	3.03 (1.14)	3.40 (1.11)	0.36 (0.20 : 0.52)	0.00
Worries/anxiety	3.32 (1.39)	3.08 (1.34)	-0.21 (-0.40 : -0.02)	0.03
Connectedness				
Adults at school	3.06 (0.85)	3.33 (0.74)	0.27 (0.16 : 0.38)	0.00
Adults in neighbourhood	2.85 (1.06)	3.06 (0.99)	0.23 (0.09 : 0.37)	0.00
Adults at home	3.20 (0.88)	3.49 (0.71)	0.27 (0.16 : 0.37)	0.00
Peer belonging	3.57 (1.27)	4.21 (0.96)	0.63 (0.48 : 0.77)	0.00
Friendship intimacy	3.89 (1.27)	4.41 (0.93)	0.49 (0.35 : 0.63)	0.00
School experiences				
Academic self-concept	3.86 (1.23)	4.39 (0.84)	0.50 (0.38 : 0.63)	0.00
School support	3.75 (1.20)	4.24 (0.92)	0.48 (0.34 : 0.61)	0.00
Bullying	2.29 (1.27)	1.86 (1.11)	-0.36 (-0.52 : -0.20)	0.00
School belonging	3.42 (1.44)	4.21 (1.06)	0.78 (0.62 : 0.94)	0.00
Wellbeing sum	N/A	N/A	N/A	N/A

APPENDIX K

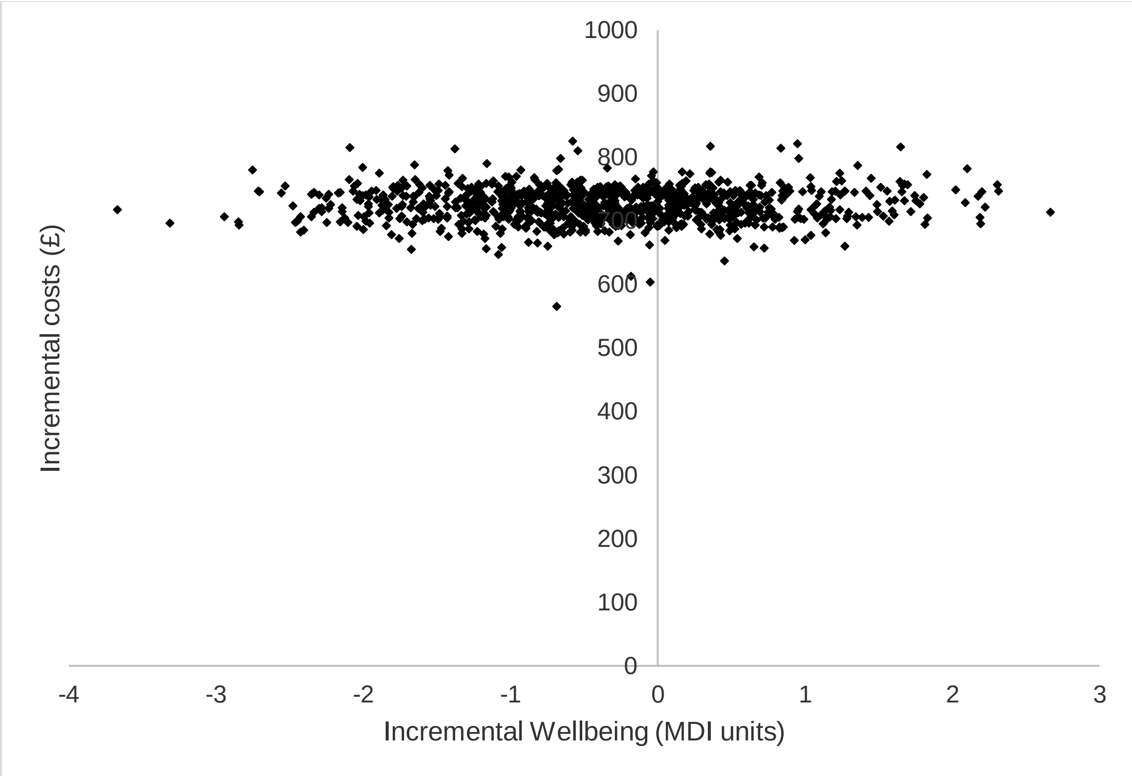

Figure 8 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and effects of the Daily Mile from the school perspective - Girls

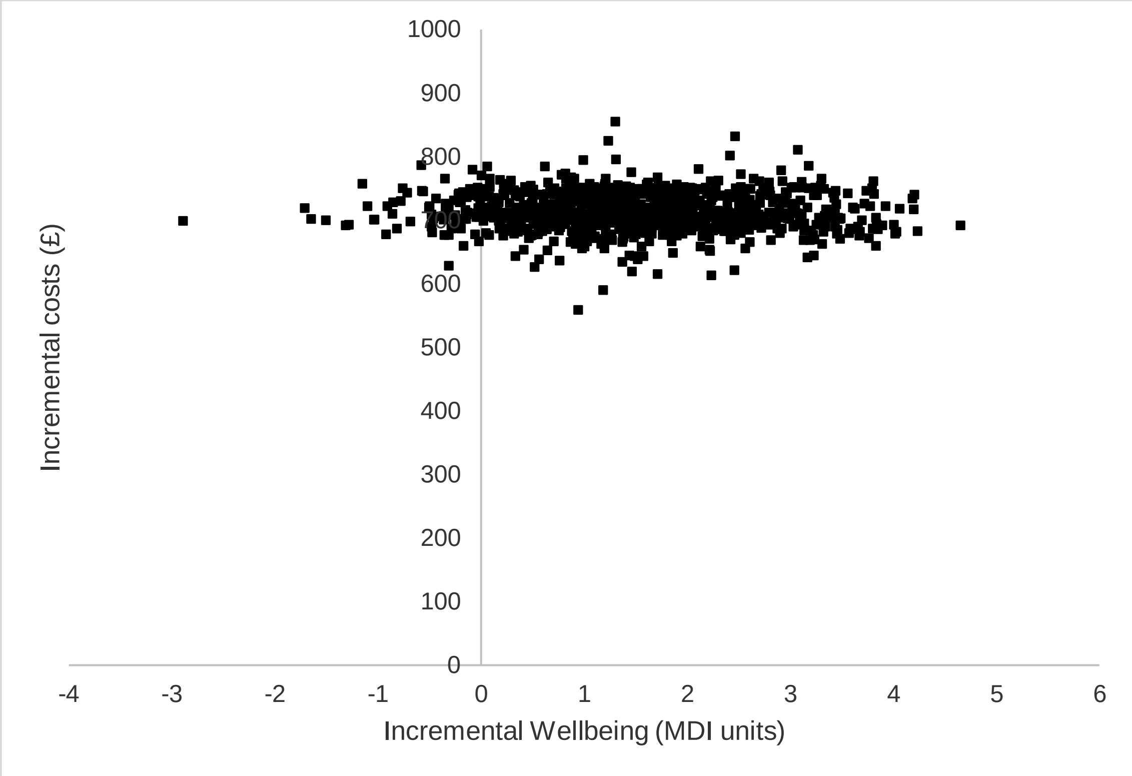

Figure 9 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and effects of the Daily Mile from the school perspective – Boys

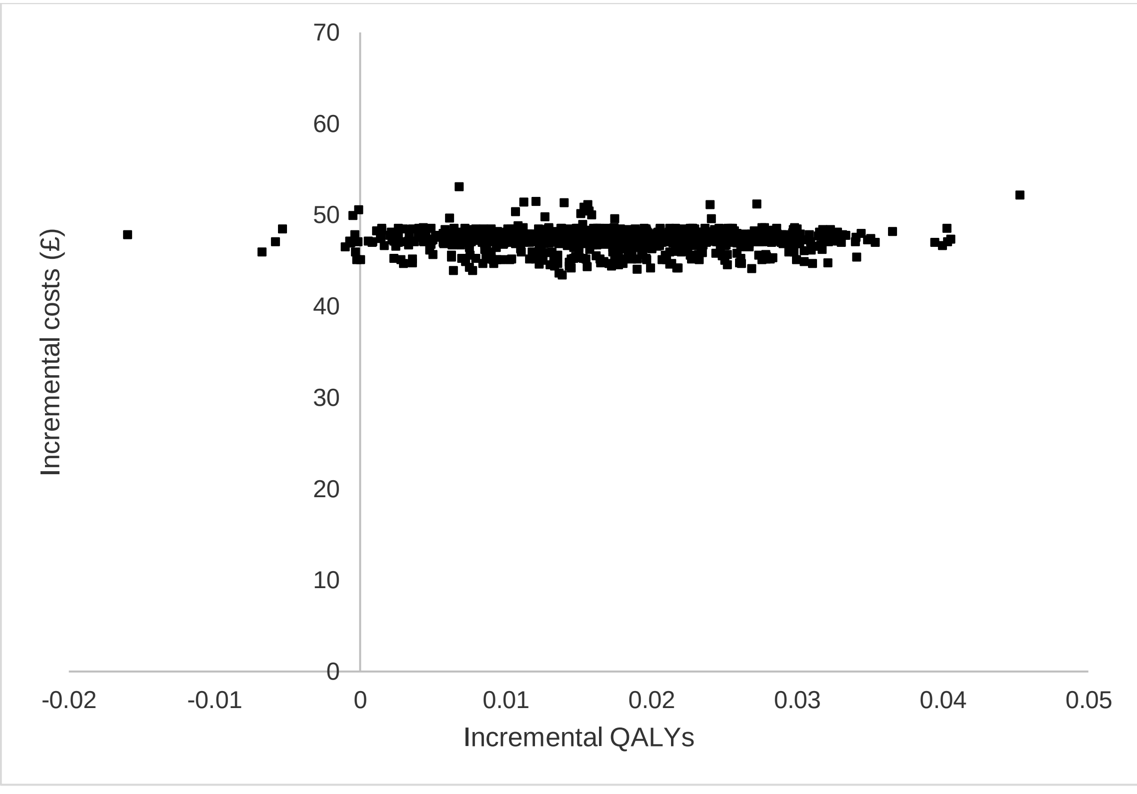

Figure 10 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and effects of the Daily Mile from the public sector perspective – Girls

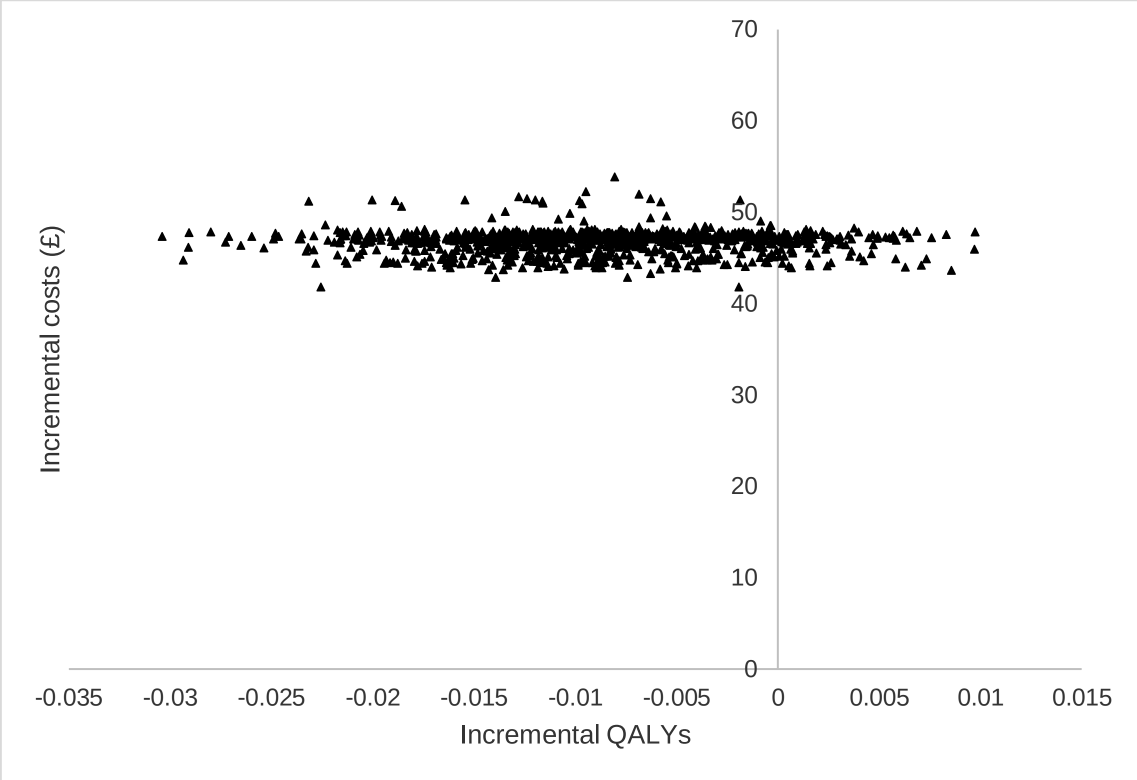

Figure 11 Cost effectiveness plane of 1,000 bootstrap replications of incremental costs and effects of the Daily Mile from the public sector perspective - Boys